Collaborative Conceptual Design Methods in the Context of the Swedish Deregulated Railway Market
- from the perspectives of maintenance, infrastructure management, product provision, and research

Anna Malou Petersson

Operation and Maintenance Engineering
Collaborative conceptual design methods in the context of the Swedish deregulated railway market

- from the perspectives of maintenance, infrastructure management, product provision, and research

Doctoral thesis

Anna Malou Petersson

Luleå University of Technology
Department of Civil, Environmental and Natural Resources Engineering
Division of Operation, Maintenance and Acoustics
Abstract

As a result of the deregulation of the Swedish railway market, the different functions involved during the life cycle of railway products have been spread out over different actors. The trigger for the work presented in this thesis was the insight that, since function-specific knowledge of the railway in Sweden belongs to different organizations, special efforts would be needed to achieve a holistic picture of a railway product during the conceptual design phase. However, there is a lack of methods developed for groups with participants from different organizations possessing different functional knowledge about a product.

The overall aim of the research for this thesis was therefore to develop an ideation method and a concept selection method specifically for groups acting in the context of the Swedish deregulated railway market, with members from different organizations with different functional knowledge. By means of action design research, the methods were developed within the scope of a research project involving four different actors on the Swedish railway market.

The project group tested Method 635, the gallery method and the SIL method. Overall, the gallery method was most popular and the SIL method least popular among the participants. Method 635 was found to have a considerably higher idea and concept generation rate than the gallery and SIL methods.

The developed ideation method combines the rotational viewing of generated ideas in the group and gallery viewing with steps of verbal interaction where the group members take turns to present and comment on generated ideas. The project group regarded the developed method as the best method of all the methods which they had tried, and it outperformed the other methods with regard to the total number of generated concepts and ideas. The views of the project group on the method correlated positively to the views of participants taking part in field tests of the method.

The developed concept selection method combines refined versions of the procedures for topic clarification, goal specification, requirement specification and concept screening developed by Pahl and Beitz with a concept scoring procedure based on the life cycle and societal costs associated with each concept that fulfils the stated requirements.

The most important effect of the composition of the group was found to be the wide range of viewpoints on the topic at hand that the representative group of different relevant actors was able to provide.

Keywords: action design research, collaboration, competitive tendering, concept selection, cross-functional groups, deregulation, field tests, gallery method, ideation, inter-organizational groups, life cycle cost, maintenance, Method 635, methods, performance-based contracts, product development, railway, SIL method, societal cost, Sweden
Sammanfattning på svenska

Som en konsekvens av avregleringen av den svenska järnvägsmarknaden har olika funktioner relaterade till en järnvägsprodukts livscykel spridits ut på olika aktörer. Utgångspunkten för arbetet som utförts inom ramen för denna avhandling var insikten att när funktionsspecifik kunskap fanns hos olika organisationer, skulle speciella åtgärder krävas för att uppnå en helhetsbild av en produkt under den konceptuella designfasen. Dock saknas metoder som är anpassade till grupper där deltagarna kommer från olika organisationer och har olika funktionell kunskap om en produkt.

Det övergripande målet för forskningen som beskrivs i denna avhandling var därför att utveckla en ideäalarmstringsmetod och en konceptvalsmetod specifikt anpassade för grupper på den svenska avreglerade järnvägsmarknaden, med medlemmar från olika organisationer med olika funktionell kunskap. Med hjälp av ”action design research” så utvecklades metoderna inom ett forskningsprojekt där fyra olika aktörer på den svenska järnvägsmarknaden deltog.


Den utvecklade ideäalarmstringsmetoden kombinerar roterande visning (som i Metod 635) med gallerivisning (som i gallerimetoden) med verbal interaktion där gruppmedlemmarna turas om att presentera och kommentera de uppkomna idéerna. Projektdelegtarna tyckte att den utvecklade metoden var den bästa av de metoder de hade prövat, och den överträdde de andra metoderna när det gäller det totala antalet producerade idéer och koncept. Åsikten om metoden hos fälttestdeltagarna som prövat den korrelerade positivt med projektdelegtarnas åsikter.

Den utvecklade konceptvalsmetoden kombinerar förbättrade versioner av problemkartläggning-, målsättning-, kravspecifikation- och konceptvalnings-processer utvecklade av Pahl och Beitz med en konceptpoängsmetod som baseras på livscykel- och samhällskostnader för de koncept som uppfyller samtliga krav.

Den viktigaste effekten av gruppsammansättningen var den omfattande spännvidden av olika synpunkter som deltagarna (med olika funktionell bakgrund från olika organisationer) kunde ge på det aktuella problemet.
To my grandmother Hjördis Petersson
Acknowledgements

First and foremost, I would like to express my sincere gratitude to my main supervisor, Prof. Jan Lundberg, for his guidance, support and encouragement during the research performed for my thesis. I would also like to thank my co-supervisor, Assoc. Prof. Matti Rantatalo, for interesting discussions and fruitful comments on my work.

The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged. The OptiKrea partners participating in the project, the Swedish Transport Administration, Vossloh Nordic Switch Systems and Infranord, are gratefully acknowledged for their contributions to the project.

I am grateful to Björn Lundwall, Roland Bång, Dr Arne Nissen and Jan-Erik Meyer for their commitment to the OptiKrea project. I would like to thank Roland Bång and Björn Lundwall for giving me a practical insight into how turnouts are manufactured, operated and maintained through visits to factories and to turnouts installed in the field. Further, I would like to thank Dr Arne Nissen and Roland Bång for providing the opportunities to perform the field tests reported in this thesis.

I am very grateful to the people whom I have interviewed for giving me their time and for sharing their knowledge and opinions.

I would like to thank Dr Ian Coleman, the leader of In2Rail Work Package 2, and Peter Melander, the UPPSAMT project leader, and all the participants in the field tests reported in this thesis.

The research presented in this thesis has been carried out at the Division of Operation and Maintenance at Luleå University of Technology and I would like to thank Chaired Professor Uday Kumar, the Head of the Division, Assoc. Prof. Roger Johnson, and all my colleagues during my PhD time for their support. A special thanks is extended to my former office mate and office neighbour, Dr Amparo Morant.

I would like to acknowledge gratefully the administrative support extended by several persons at LTU, which has facilitated my research. A special thanks is due to Marie Jakobsson and the IT department, who always provided quick assistance regarding the problems which I encountered when working from Norway.
I would like gratefully to acknowledge Assoc. Prof. Anita Gärling for her valuable comments on earlier versions of this thesis, which have resulted in a significant improvement in the quality.

I would like to thank Paul McMillen, the language reviewer of this thesis, as well as my papers, for his excellent service.

I would also like to thank Gunvor Ekström for her valuable work on the concept drawings shown in Part II of this thesis.

I am very grateful to my family and friends for their support and encouragement. In particular, I would like to thank my late grandmother, Hjördis Petersson, who has been a great inspiration for me and who has not only been very supportive during the different phases of my education, but has also given me insights into the inherent value of education and the importance of everyone having the right to an education and the possibility of acquiring one. I wish you had had the same possibilities as I have had and I wish you had been here to see this thesis.
Appended papers

Several chapters in this monograph are based on material that has been published before, either in journal or conference papers, or in papers that have been submitted to scientific journals. In this part of the thesis, these papers are listed and a description is given of how they have been used in the thesis. Further, the author’s contribution to each paper is clarified. The papers are attached in Part III of the thesis.

List of papers

This thesis is partly based on the following papers, which are referred to by their Roman numerals.


III Petersson, A.M. and Lundberg, J. Developing an ideation method to be used in cross-functional inter-organizational teams by means of action design research. Submitted to Research in Engineering Design.


Reprints were made with the permission of the respective publishers (when necessary).

Use of material from the papers in the thesis

The use of material from each paper is described in Table A.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Chapter(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>4</td>
</tr>
<tr>
<td>II</td>
<td>2, 5, 9, 10</td>
</tr>
<tr>
<td>III</td>
<td>2, 6, 9</td>
</tr>
<tr>
<td>IV</td>
<td>7</td>
</tr>
<tr>
<td>V</td>
<td>1, 2, 4, 8</td>
</tr>
<tr>
<td>VI</td>
<td>11</td>
</tr>
</tbody>
</table>

The author’s contribution to the papers

The research setup and the planning of the studies performed for this thesis were made in a dialogue with the author’s main supervisor, Prof. Jan Lundberg, and in Paper II the author’s co-supervisor, Assoc. Prof. Matti Rantatalo, also contributed significantly to this activity. All the preparation, data collection, data analysis, writing and creation of figures were performed by the author for all the papers listed above. Prof. Jan Lundberg reviewed all the papers and Assoc. Prof. Matti Rantatalo reviewed Paper II.
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADR</td>
<td>action design research</td>
</tr>
<tr>
<td>AHP</td>
<td>analytical hierarchy process</td>
</tr>
<tr>
<td>AMM</td>
<td>associative memory models</td>
</tr>
<tr>
<td>APWETPSS</td>
<td>Act on Procurement within the Water, Energy, Transport and Postal Services Sectors</td>
</tr>
<tr>
<td>BIE</td>
<td>building, intervention and evaluation</td>
</tr>
<tr>
<td>CSM</td>
<td>concept selection method</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>GP</td>
<td>group phase</td>
</tr>
<tr>
<td>IM</td>
<td>infrastructure manager</td>
</tr>
<tr>
<td>IP</td>
<td>individual phase</td>
</tr>
<tr>
<td>IT</td>
<td>information technology</td>
</tr>
<tr>
<td>JVTC</td>
<td>Luleå Railway Research Center</td>
</tr>
<tr>
<td>LCC</td>
<td>life cycle cost</td>
</tr>
<tr>
<td>LTU</td>
<td>Luleå University of Technology</td>
</tr>
<tr>
<td>MCDM</td>
<td>multi-criteria decision making</td>
</tr>
<tr>
<td>MGT</td>
<td>million gross tonnes</td>
</tr>
<tr>
<td>MPB method</td>
<td>modified Pahl and Beitz method</td>
</tr>
<tr>
<td>MSEK</td>
<td>million Swedish crowns</td>
</tr>
<tr>
<td>OKMv1</td>
<td>OptiKrea method version 1</td>
</tr>
<tr>
<td>OKMv2</td>
<td>OptiKrea method version 2</td>
</tr>
<tr>
<td>RAMS</td>
<td>reliability, availability, maintainability and safety</td>
</tr>
<tr>
<td>RQ</td>
<td>research questions</td>
</tr>
<tr>
<td>SC</td>
<td>societal cost</td>
</tr>
<tr>
<td>SSR</td>
<td>Swedish State Railways (Statens Järnvägar)</td>
</tr>
<tr>
<td>SRA</td>
<td>Swedish Rail Administration (Banverket)</td>
</tr>
<tr>
<td>STA</td>
<td>Swedish Transport Administration (Trafikverket)</td>
</tr>
</tbody>
</table>
# Table of contents

Acknowledgements ...................................................................................................... vii
Appended papers .......................................................................................................... ix
  List of papers ........................................................................................................... ix
  Use of material from the papers in the thesis ......................................................... x
  The author’s contribution to the papers ................................................................. x
List of abbreviations ..................................................................................................... xi
Contents ....................................................................................................................... xiii

## PART 1: Railway deregulation in Sweden and concept generation and selection methods developed for that context
1 Introduction ................................................................................................................. 3
  1.1 Aim and research questions ............................................................................... 6
  1.2 Research scope and limitations ......................................................................... 7
2 State of the art ............................................................................................................. 9
  2.1 Diverse groups ................................................................................................. 9
  2.2 Ideation methods ............................................................................................ 10
  2.3 Concept selection methods ............................................................................. 15
  2.4 Design for maintenance ................................................................................ 24
  2.5 Life cycle cost ................................................................................................ 25
3. Research design ....................................................................................................... 27
  3.1 Research approach .......................................................................................... 27
  3.2 Research method ............................................................................................. 28
  3.3 Research process .............................................................................................. 29
  3.4 The OptiKrea Group ...................................................................................... 35
4. Assessment of the Swedish deregulated railway market ...................................... 37
  4.1 Introduction ...................................................................................................... 37
  4.2 Study 1: The impact of procuring maintenance contracts ......................... 38
  4.3 Study 2: Collaboration on the Swedish deregulated railway market .......... 52
  4.4 Study 3: Conditions for product development ............................................. 61
  4.5 Discussion ........................................................................................................ 63
  4.6 Conclusions ..................................................................................................... 70
10.2 The developed methods and the Swedish deregulated railway market........... 245

11. Reflections on using action design research to develop conceptual design methods ....................................................................................................................... 247

11.1 Conclusions ..................................................................................................... 249

12. Implementation of the developed methods at the Swedish Transport Administration ........................................................................................................... 251

12.1 Framework for product development at STA .................................................. 251

12.2 Combining STA routines and the OptiKrea process ........................................ 253

12.3 Adaption of the OptiKrea methods ................................................................. 255

12.4 Roadmap for implementation ........................................................................ 256

References ............................................................................................................. 257

PART II: Overview of generated concepts ................................................................. 269

Introduction ............................................................................................................. 271

Ideation topic 1 ....................................................................................................... 273

Ideation topic 2 ....................................................................................................... 279

Ideation topic 3 ....................................................................................................... 287

Ideation topic 4 ....................................................................................................... 291

Ideation topic 5 ....................................................................................................... 301

PART III: Papers .................................................................................................... 319
Part I: Railway deregulation in Sweden and concept generation and selection methods developed for that context
1 Introduction

In its aim to establish a Single European Railway Area, the European Union (EU) has encouraged competition and the opening up of the railway market through a number of gradual legislative proposals (Alexandersson and Rigas, 2013). Sweden is one of the European countries that have implemented the most far-reaching reforms in the railway sector and can be considered to be several years ahead of current EU policy (Alexandersson and Rigas, 2013). Therefore, Sweden is an interesting example for other countries to learn from and can show the way with regard to addressing different issues related to the liberalization of the market.

The deregulation of the Swedish railway system started in 1988 with the vertical separation of train operations from the railway infrastructure. The deregulation was driven by the need to find new possibilities of financing the railway sector and increasing efficiency (Alexandersson and Hultén, 2008). Railway infrastructure ownership, investments and maintenance responsibility were transferred from the public utility the Swedish State Railways (SSR) to a national authority, the Swedish Rail Administration (SRA). Public procurement of train operations on railway lines began shortly afterwards. In 1998, new policies led to the split-up of SRA into a client and contractor, and competition in the area of railway maintenance was initiated in 2001 with contracts in the low-traffic parts of the network. Again, the purpose was to increase efficiency (Espling, 2007). By 2012, 95% of the maintenance carried out on the State-owned railway infrastructure was being procured publicly (Regeringen, 2012). In 2010, SRA and several other Swedish authorities were merged to form the Swedish Transport Administration (STA). STA is responsible for the strategic planning of roads, railways, ship transportation and aviation, as well as investment in and the operation and maintenance of roads and railways in Sweden.

A consequence of the deregulation process was that the responsibility for the design and production of railway products, which originally belonged to SSR, was transferred to companies and was procured thereafter by SRA and later by STA. For example, the design and production of turnouts were taken over by a company in 1990. In 1993, SRA initiated the public procurement of turnouts through frame agreements.

Although STA does not have any responsibility for the design and production of railway products within their organization, STA participates in the conceptual design phase (the early stages of product development where requirements are formulated, ideas generated and concepts selected) in at least two
different roles. Firstly, if STA recognizes a need to introduce a new or improved product in the infrastructure which STA manages, it can procure this product based on stated requirements. In practice, this often means that the development of this product is outsourced to the company that receives the contract, since these products are typically not “off-the-shelf” products. Secondly, they can take part in development work in research projects. These projects can take place on the national level or, as is increasingly common, on the European level in the form of EU research projects.

A profit-driven company is the most common environment for product development (Ulrich and Eppinger, 2014). In this environment, a successful product development leads to products that are profitable to produce and sell. The existing conceptual design methods are often designed from this perspective. The situation is different, however, from STA’s perspective. STA is to create prerequisites for a transport system that is efficient concerning the public finances, internationally competitive and sustainable in the long term (Näringsdepartementet, 2010). The goal of STA is not to make a profit, but, on the other hand, tax revenues are limited and should be spent efficiently and, therefore, the utility and quality of purchased products must be balanced against their cost. Further, STA must ensure that the available suppliers on the market will be able to provide them with the appropriate products at the appropriate times, which requires foresight, as the development of railway products can take several years and specific requirements differ between different infrastructure managers.

Purchases made by public authorities such as STA are regulated by Swedish legislation. The Swedish Act on Procurement within the Water, Energy, Transport and Postal Services Sectors (APWETPSS) (Lag om Upphandling inom Områdena Vatten, Energi, Transporter och Posttjänster, 2007) applies to purchases concerning railway-related activities (Sundstrand, 2013). The APWETPSS is based on Directive 2004/17/EC of the European Parliament and of the Council (EU, 2004), and corresponding rules apply in other countries belonging to the EU.

The trigger for the research carried out for the present thesis was the insight that, as the different functions involved during the life cycle of railway products were spread out over different actors as a result of deregulation, developing new products would possibly become more difficult and the risk of suboptimization would increase. As the function-specific knowledge belonged to different organizations, it was suspected that special efforts were needed to achieve a holis-

---

1 On 1st January 2017, the APWETPSS (passed in 2007) was replaced by the Act on Procurement within the Supply Sectors (Lag om Upphandling inom Försörjningssektorn, 2016), which is based on Directive 2014/25/EU of the European Parliament and of the Council (EU, 2014). All the research activities reported in the present thesis took place before the replacement of the APWETPSS and, consequently, the findings concerning purchases connected to railway-related activities apply to the APWETPSS and not the new law. It has been out of the scope of the present research to investigate whether or not the new law can be expected to lead to similar findings.
tic picture of a product during the conceptual design phase. By involving representatives from several actors during the conceptual design phase, more viewpoints on the product to be developed can be shared and thus a better product might eventually be designed.

Groups with members from different organizations connected to different functions in relation to a railway product are for the purpose of the present thesis viewed as a special case of cross-functional inter-organizational groups. Such groups are cross-functional, since they possess different functional knowledge about the product, and they are inter-organizational, since their members represent different actors on the railway market. Previous research on groups in the conceptual design phase has focused on cross-functional teams within an organization. However, many technical medium- to large-scale product development projects might benefit from the participation of members from different organizations in the product development team. As outsourcing is becoming increasingly popular and deregulation has opened up previously closed markets in many countries, different functions are being spread out over several actors specializing in different functions related to a product. Despite the fact that previous research has shown examples of the customer’s importance (Hannola et al., 2009) and the subcontractor’s importance (Liker et al., 1998) for the introduction of improvements, studies regarding conceptual design methods with teams consisting of members from different organizations and working in real-life settings have, to the best of the author’s knowledge, been scarce. However, Rexfelt et al. (2011) studied methods for service innovation in two teams with participants from different organizations, and Almefelt and Claesson (2015) studied the application of a design methodology in a large multidisciplinary project within the automobile industry with several participating organizations. Cross-functional inter-organizational teams working in real-life settings differ in several ways from the kind of teams participating in the majority of previous studies. Firstly, members from different organizations are professionals representing several different market actors. According to Straus et al. (2011), little is known about what happens in groups whose members bring different organizational cultures and political agendas. Secondly, such groups typically consider complex engineering topics that require domain-specific knowledge. Highly complex task domains typically require specialists with different expertise and backgrounds to integrate information in order to achieve a high-quality solution (Mesmer-Magnus and DeChurch, 2009). Kavadias and Sommer (2009) found analytically that the nature of the problem matters and that experimental evidence in the literature might have been influenced by simple ideation topics that did not represent real situations. Thirdly, cross-functional inter-organizational groups in real-life settings themselves define and own the topics which are to be the subject of the conceptual design phase. Bolin and Neuman (2006) suggested that differences in the strength of the incentive for high performance would affect the outcome of the process of generating ideas on certain topics.
1.1 Aim and research questions

The overall aim of the present research has been to develop conceptual design methods which are specifically intended for groups acting in the context of the Swedish deregulated railway market, with members from different organizations with different functional knowledge, and which will be implemented after the research project ends. The emphasis is on group members who are engineers or have a technical understanding of the product to be developed, although other group members could be relevant depending on the product. In this thesis, “product” is used in its broadest sense as meaning both physical products and processes, such as the performance of maintenance, or a combination of these.

In Figure 1.1, a simplified generic product development process is presented. The main focus of the research performed for the present thesis was aimed at the development of methods for ideation, i.e. methods that facilitate the generation of suggestions on how to solve the problems associated with the given task. Methods for task definition and the evaluation and selection of concepts have also been developed.

To support the development of the methods, the following research questions were formulated.

**Research question 1**: What are the prerequisites for collaborative product development on the Swedish deregulated railway market?

**Research question 2**: How can an ideation method be designed to facilitate information sharing and idea flow in a cross-functional inter-organizational group in the context of the Swedish deregulated railway market?

**Research question 3**: How can a concept selection method be designed to support STA in the selection of an appropriate concept for further development?

**Research question 4**: What effects arise when using conceptual design methods in cross-functional inter-organizational groups in the context of the Swedish deregulated railway market as a result of the composition of the group?

The search for the effects mentioned in research question 4 has been explorative and broad, and could, for example, concern effects on the ways in which the group members behave and interact, and what type of information is shared in the group.

The research work was divided into different stages that answered different research questions in order to achieve the overall aim. Table 1.1 shows the research stages, their objectives and their connection to the research questions.
Each research stage objective was further broken down into sub-objectives and connected research sub-questions. These sub-objectives and research sub-questions are presented in Chapter 4-8, which deal with Stage 1-5, respectively.

Table 1.1 The research stages and their objectives, and the research questions (RQ) which they relate to.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Objectives</th>
<th>RQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assessment of the prerequisites for product development in the context of the Swedish deregulated railway market</td>
<td>To understand the state of the Swedish deregulated railway market and to clarify the prerequisites for product development</td>
<td>1</td>
</tr>
<tr>
<td>2. Test of established ideation methods</td>
<td>To collect empirical data on ideation in a cross-functional inter-organizational group in the context of the Swedish deregulated railway market and confirm earlier findings in the literature, to provide an understanding of how an ideation method should be designed for that specific context</td>
<td>2,4</td>
</tr>
<tr>
<td>3. Development of an ideation method</td>
<td>To use the findings from Stage 1 and 2 to develop an ideation method for cross-functional inter-organizational groups in the specific context of the Swedish deregulated railway market and evaluate it in several cycles</td>
<td>2</td>
</tr>
<tr>
<td>4. Field tests of the developed ideation method</td>
<td>Evaluations of the developed ideation method in other groups to judge its value in typical use settings</td>
<td>2</td>
</tr>
<tr>
<td>5. Development of a concept selection method</td>
<td>To evaluate and develop further a decision support method intended for the specific context of the Swedish deregulated railway market</td>
<td>3,4</td>
</tr>
</tbody>
</table>

1.2 Research scope and limitations

The research project “OptiKrea” was initiated in 2012 and has been run by Luleå Railway Research Center (JVTC) at Luleå University of Technology. The intention of the project has been to promote the technical development of railway products, especially from the point of view of maintenance and the life cycle cost, through collaboration between different railway market actors. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining, utilizing and researching the railway. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about a railway product that the different actors possess, better solutions would be found. This would be accomplished by the development of conceptual design methods, which would facilitate collaboration.
and innovation, would be tailor-made for the railway sector and could be used in the future. The research work carried out for the present PhD thesis has been part of the OptiKrea project.

Railway infrastructure and particularly track components are expensive assets with long life spans, and the punctuality and effectiveness of train operations heavily depend on their seamless operation. The turnout is a vital part of the railway infrastructure and a failure of a turnout, especially in a critical location, can cause significant delays and societal costs. In 2012, turnout-related failures were among the top ten causes of hours of disruption in Sweden (Trafikverket, 2013). Therefore, the turnout was chosen as a case study object for the OptiKrea project. At the heart of the project has been the so-called “creative team”, which has consisted of representatives from each collaborating actor and, therefore, has been inter-organizational. The team has been cross-functional in the sense that it has represented the different functions that are important when developing a turnout, i.e. research, design, manufacturing, management, maintenance, and disposal. The major part of the research work described in the present thesis has been conducted within this group. Exceptions are Stage 1 and Stage 4 (see Table 1.1), where the research was extended to include other groups.
2 State of the art

In this chapter, prior research relevant to the research performed for this thesis is presented. The research topic is highly multi-disciplinary and, consequently, there are many fields which can contribute relevant findings. To reflect the content of the actual research work carried out, the main emphasis of this chapter is on ideation methods and their use in diverse groups, followed by decision support methods. In addition, some information is given on maintenance and life cycle cost (LCC) calculations, as these concepts are essential to understand the development of the methods applied in the present research.

2.1 Diverse groups

The conceptual design phase is a critical step, since it influences the subsequent steps during product development with regard to cost, quality, and performance (e.g. Rubenstein, 1994; Okudan and Tauhid, 2008). To focus attention on different aspects of the product-to-be as early as possible, in order to avoid becoming locked in an unsuitable solution, working in cross-functional teams has become increasingly popular. Further, as functional diversity is introduced in work groups, more creative and innovative ideas and solutions may be procured (van Knippenberg et al., 2004). Heterogeneous groups are expected to outperform homogeneous groups and non-interacting individuals in fulfilling creative tasks, because heterogeneous groups have a greater range of skills and resources to draw from (Milliken et al., 2003). However, the presence of a diverse group and its potential range of skills and abilities does not guarantee the effective utilization of those skills and abilities. An expert in one domain will often not know what information has value for an expert in another domain (Straus et al., 2011), and information sharing is essential to make use of a group’s available informational resources collectively. For various reasons, the group’s members may be less willing to contribute to the group work, or other processes might inhibit information sharing and therefore decrease group performance and productivity (e.g. Pelled and Adler, 1994; Jassawalla and Sashittal, 1999; van Knippenberg et al., 2004).
2.2 Ideation methods

When a group’s composition is diverse, the use of structured ideation strategies may be able to bolster cognitive activities associated with the creative process by forcing the consideration of many alternatives (Milliken et al., 2003). Several researchers and practitioners have focused on developing practically useful ideation methods that aim to encourage the ideating participants to propose a large amount of ideas on a particular issue during a limited time frame. Shah et al. (2000) classified ideation methods into two categories, intuitive and logical methods. Intuitive methods stimulate the unconscious thought processes of the human mind, whereas logical methods involve systematic decomposition and analysis of the problem (Shah et al., 2000). There is an extensive amount of social psychology literature on brainstorming (Osborn, 1957), the most well-known intuitive ideation method, and the claim by Osborn (1957) that brainstorming results in enhanced efficiency has been refuted by a large body of research. Although it has been found that people feel more productive working in a group than working on their own (Stroebe et al., 1992; Paulus et al., 1993), the combined output of individuals ideating on their own (nominal groups) outperforms that of verbally interactive groups in terms of non-redundant ideas and quality of ideas (e.g. Byron, 2012; Mullen et al., 1991). Since verbal interaction contributes to process losses (Mullen et al., 1991), several so-called brainwriting methods have been suggested. The common feature of these methods is that, instead of communicating ideas verbally, ideas are documented by each participant in silence by writing. Several researchers have highlighted the importance of sketching during engineering design (e.g. Shah et al., 2001; Yang, 2009) and so-called brainsketching techniques communicate ideas through sketches rather than words. Hybrid ideation methods combine brainwriting and/or brainsketching with verbal interaction. Compared to brainstorming, other ideation methods have received comparably less attention and many methods have been suggested, but not studied from a scientific point of view.

2.2.1 Relevant ideation methods

Several ideation methods are relevant to the present thesis. In the original Method 635 (Rohrbach, 1969), a brainwriting technique, six participants write down three solutions to the problem on a sheet of paper during five minutes and then pass the sheet to their neighbour, who reads through the ideas and contributes three further ideas or developments of ideas during five minutes. When the sheet has passed between all the participants, the ideation stops (Rohrbach, 1969). Alternative brainwriting techniques that have features similar to those of Method 635 are the brainwriting pool (Geschka et al., 1973) and the pin card technique (VanGundy, 1984).

The basic brainsketching technique involves each group member individually sketching their ideas on sheets of paper and passing the sheets on, after a few
minutes, to another member, who continues to sketch ideas, using the already generated ideas as a source of inspiration for new ideas (van der Lugt, 2002). Shah et al. (2001) suggested an alternative brainsketching technique called C-sketch. In C-sketch, the participants work independently on developing a sketch (no textual description is allowed) of their proposed solution for the ideation topic, and when the cycle-time ends, the sketch is passed on to their neighbour, who adds, modifies or deletes aspects of the solution during the next cycle. The solutions pass between all the participants, so that when the session is completed, each solution will have been worked upon by every participant and the number of solutions equals the number of participants (Shah et al., 2001).

Examples of hybrid ideation methods are the gallery method (Pahl et al., 2007) and the SIL method (VanGundy, 1984). The gallery method starts with an idea generation step of 15 minutes where the participants use sketches, supported by annotations as necessary, to describe ideas, followed by a step where the sketches are hung on a wall so that all the group members can see and discuss them for 15 minutes. During an additional 15 minute ideation step, the participants use the inspiration from the discussion step to develop ideas further (Pahl et al., 2007). The SIL method also starts with an individual ideation step (VanGundy, 1984). Thereafter, two members read aloud one idea each and all the group members try to combine these two ideas verbally into one idea. During the next step, a third group member reads their idea aloud and the group then tries to integrate this idea with the previous one, and the process continues until an integrated solution is found (VanGundy, 1984).

2.2.2 Related scientific studies

Van der Lugt (2002) compared a version of brainsketching that allowed the participants to explain their ideas briefly after each round of idea sketching to brainstorming. It was found that significantly more ideas were generated during brainstorming than during brainsketching, but more connections with previous ideas were made during brainsketching (van der Lugt, 2002). Comparing C-sketch to Method 635 (allowing only textual description) and the gallery method in an experiment involving mechanical engineering students and practising designers, Shah et al. (2001) found that C-sketch outperformed Method 635 in the three measured areas of quality, novelty, and variety of designs generated and was better than the gallery method with regard to novelty and variety. The gallery method was better than Method 635 concerning these three parameters, suggesting that sketches are a useful means of communication in idea generation in the field of engineering design (Shah et al., 2001). It should be noted that the number of ideas was not compared between the methods, since that quantity was fixed in the ideation method instructions.

Linsey et al. (2011) conducted a study on how two key factors of brainsketching, C-sketch, Method 635 and the individual phases of the gallery method influenced the quantity, quality, novelty and variety of ideas. The two
key factors were as follows: a) how the ideas are displayed to other group members (“rotational view” or “gallery view”), and b) how the ideas are communicated between group members (written words only, sketches only, or a combination of words and sketches), resulting in six different group ideation conditions. During rotational viewing, the participants view different subsets of the entire set of ideas during different time periods, whereas during gallery viewing the participants view the entire set of ideas at once. The experiment involved mechanical engineering students working on the design of a device for shelling peanuts. Linsey et al. (2011) found that the highest number of ideas was generated when the groups used rotational viewing and communication by means of a combination of words and sketches. A larger number of high-quality concepts were generated with gallery viewing in combination with sketches only than were generated using any other condition. There were no significant differences between the results for the different conditions with regard to novelty or variety (Linsey et al., 2011). Linsey and Becker (2010) conducted a complementary study by letting nominal groups ideate on the peanut sheller problem given the same amount of time. They found that real teams using rotational viewing conditions developed a larger number of ideas than equivalent nominal groups. Gallery viewing was better than nominal groups when words and sketches were used to represent the ideas (Linsey and Becker, 2010). This is a very interesting finding, since the extensive research on brainstorming in the field of psychology unambiguously shows that verbally interactive group work decreases the amount of ideas compared to the amount obtained through nominal groups (e.g. Mullen et al., 1991; Byron, 2012). The study by Linsey and Becker (2010) shows that interactive groups can outperform nominal groups under certain conditions. Using ideation methods in an optimal way clearly has a great potential to increase the number of ideas which a team can contribute during concept generation. This is especially important since past research on brainstorming in the psychology field has found that the total number of ideas and the number of ideas of high quality are positively correlated (Dugosh et al., 2000; Paulus et al., 2015), and some supporting evidence for this finding has been found in the area of engineering design (Yang, 2009; Linsey et al., 2011).

Paulus and Yang (2000) have shown that a failure to improve production in interactive groups compared to that in nominal groups through cognitive stimulation may be caused by participants simply not attending to stimulus ideas. Both rotational viewing and gallery viewing encourage the participants to consider the ideas from the other group members. The processing of stimulus ideas is, however, performed within the limited capacity of short-term memory, and therefore it is reasonable to assume that a person can only attend to a limited set of complex ideas simultaneously (Perttula et al., 2006). Consequently, rotational viewing may be more beneficial than gallery viewing because it helps the participant to consider only a limited set of ideas at a time. On the other hand, if the attention paid to stimulus ideas becomes excessive, the performance decreases, as has been shown experimentally by Paulus and Yang (2000) and through sim-
ulations by Brown and Paulus (2002). A possible explanation for the fact that Linsey et al. (2011) found that gallery viewing produced more global product solutions and more high-quality product solutions than rotational viewing, while the average diversity of the product solutions was greater for the rotational viewing condition, may be that participants in the gallery viewing condition focus on a limited set of ideas from the entire pool during most of the ideation session and develop similar solutions with increased quality, while during rotational viewing they are forced to consider a new set of ideas in each successive round, which might enhance the variety by activating other parts of their memory network.

Rexfelt et al. (2011) tested Method 635 and brainstorming in two different cross-functional groups, each consisting of three persons from a manufacturing firm and three persons from a customer firm, as one step in a service development process. The results suggested that the affiliation of the participants did not affect their productivity, whereas their experience of development work and comprehension of the specific method used did have an effect on their productivity (Rexfelt et al., 2011). The group using Method 635 struggled and generated fewer ideas that were suitable for the context compared to the brainstorming group. Rexfelt et al. (2011) explained this finding by the fact that Method 635 was carried out in silence, and thus the participants could not be directed in the right direction, either by the coaches or the other participants in the group, and by the fact that the group applying Method 635 did not have access to an inspirational poster which illustrated a vast range of service categories and which was present when the brainstorming group generated ideas.

Viswanathan and Linsey (2013) found that design methods can have a different impact on novices and experts. Studies performed on ideation methods have, however, mainly involved inexperienced users, predominantly students, working under experimental conditions, although some studies of professionals working under experimental conditions do exist. Chulvi et al. (2012) and Chulvi et al. (2013) examined how the type of ideation method (intuitive or logical) influenced the level of creativity (in terms of novelty and usefulness) of solutions that were generated by a multidisciplinary team. Each team consisted of one designer, one mechanical or industrial engineer, and another designer or engineer. Although Chulvi et al. (2013) reported some interesting results, not least that using any of the methods was better than using no method at all, no explicit conclusions were drawn as to how the multidisciplinary character of the team influenced the results. A few studies of professionals using ideation methods in a real-life context exist (e.g. Sutton and Hargadon (1996) and Jackson and Poole (2003)). Jackson and Poole (2003) found that idea generation in groups in naturally occurring contexts is more complex than typically imagined. The idea generation rate per time unit or member was considerably lower than that in laboratory studies.
2.2.3 Associative memory models in the context of cross-functional inter-organizational groups

Associative memory models (AMMs) used in cognitive psychology (e.g. Collins and Loftus, 1975; Anderson, 1983) offer a framework for understanding the benefits of involving participants with different functional backgrounds and knowledge, and are used to explain and justify theoretically the conceptual design methods and, in particular, the ideation method developed in the research performed for this thesis. In these models, information in the long-term memory is treated as a network of interconnected nodes. The inter-relatedness of two distinct nodes is determined by their associative strength; i.e. closely related pieces of information are connected to each other more strongly than less related pieces of information. A stimulus idea activates (and hence makes more accessible) other nodes that are connected to it, and primarily those with a strong associative connection. Nijstad et al. (2002) argued that “idea generation differs from retrieval, because (new) ideas cannot be directly retrieved from memory”. However, stored knowledge must be used to generate ideas, and therefore idea generation necessarily involves retrieval processes (Nijstad et al., 2002). Therefore, Nijstad et al. (2002) suggested that idea generation is a two-stage process in which a cue activates a set of localized strongly interconnected and semantically related features. The idea production stage follows, where this set is used to generate (new) ideas by combining knowledge, forming new associations or applying knowledge to a new domain. These ideas are added to the search cue to activate more sets in memory, leading to additional idea generation (Nijstad et al., 2002). Since semantically related images have strong mutual ties, successively activated images will often have strong associative connections which lead to a “train of thought” exploring connected ideas (Nijstad et al., 2002). AMMs can explain why productivity loss occurs in group work and provide insight into how to avoid such loss (Linsey et al., 2011). In a group context, the first idea presented in the group might activate a common set of information in all the participants which leads to quick convergence concerning a small set of ideas (Linsey et al., 2011). To achieve a performance gain through idea exchange, an ideating subject must receive ideas belonging to categories which would have a low probability of being surveyed if the subject were to work alone (Perttula et al., 2006). Brown and Paulus (2002) represented a brainstormer’s knowledge of a given problem as a matrix of category transition probabilities. Simulations of the so-called associative memory matrix model showed that, if a brainstormer is presented with ideas belonging to categories of low accessibility, the number of ideas belonging to those categories increases and also the total number of ideas overall; i.e. this makes the individual a more productive brainstormer (Brown and Paulus, 2002). Hence, it would appear that a group of participants with different functional and organizational backgrounds within a relevant field would have an advantage compared to a functionally homogeneous group. A functionally homogeneous group may possess similar
information and connections in the part of their associative memory network that applies to the ideation topic at hand. Then it could be expected that the ideas which they would come up with as a reaction to the ideation topic would be similar, as this topic would activate similar nodes of information. Exchanging ideas with each other to stimulate further activation would not be very helpful, since these ideas would primarily continue to activate ideas similar to those which the individuals would generate on their own. On the other hand, if the associative memory networks of the participants were to have different nodes or a different configuration of associative strength between the nodes, the ideation topic would activate different information depending on the participant. If a participant were to be exposed to ideas from other group members, these ideas might activate parts of their network that would not initially be activated by exposure to the ideation topic, eventually leading to a group set of both more ideas and ideas belonging to more different categories than would be expected from a group with the same number of people with homogeneous backgrounds with respect to the ideation topic at hand. In support of this reasoning, Brown and Paulus (2002) found through simulations that interactive brainwriting (i.e. group members ideating in silence and exchanging ideas in written form) is not universally superior to individual brainstorming, but most effective when the members of the group possess different knowledge of the ideation topic.

AMMs can also explain the phenomenon of design-fixation. Design-fixation involves designers blindly adhering to their initial ideas or presented examples, and restricts the solution space that designers explore in the search for ideas and concepts with respect to the ideation topic (Jansson and Smith, 1991). Viswanathan and Linsey (2013) suggested that design-fixation can be explained by the fact that as designers create their first idea, other nodes in the memory network that are closely associated with this first idea “come to their mind instantly, and this might lead to ideas which are variations of their initial ideas”. Thus, receiving input from others might help designers to move beyond the limited part of their memory network on which they have fixated.

Perttula et al. (2006) found that individuals generated more ideas when they discussed ideas with other individuals. Although there were some doubts about why this effect arose, there is other research (Pelled et al., 1999; Seidel and Fixson, 2013) showing that teams debating ideas come up with more novel innovations, indicating a more thorough exploration of the solution space. One explanation for this might be that, while discussing, new aspects of ideas appear which participants associate with parts of their memory network that have not yet been activated.

2.3 Concept selection methods

The objective of the concept selection phase is to find, for the context of the application concerned (as stipulated by the requirement specification), the most suitable and/or promising concept(s) among all the concepts generated during
the concept generation phase. The concept selection phase is of high importance, as the selection of a poor design concept rarely can be compensated for at a later design stage and incurs a great redesign expense (Okudan and Tauhid, 2008). It is important to note that all product development teams use some method, either explicit or inexplicit, to choose among concepts, and these methods vary concerning efficiency (Ulrich and Eppinger, 2014). Less explicit (and less structured) methods involve, for example, letting one’s intuition steer the process or letting an influential group member make a choice based on their personal preferences (Ulrich and Eppinger, 2014).

During the concept selection stage, it is not only the required functionality that must be considered, but also other criteria, including life cycle issues such as manufacturability, ease of assembly, reliability, and maintainability (Wang, 1997). According to Honkala et al. (2007), there are at least three remarkable challenges in concept selection: a) the nature of the available information is usually based on subjective perceptions and speculations, and accurate calculations are seldom available, b) the stakeholders, users, designers and producers can have conflicting requirements concerning, for example, the product design and manufacturing, or the product performance and sales price, and c) freezing the product concept can have far-reaching effects on the product costs and customer satisfaction which can only be fixed with additional costs and time. The fact that some criteria contradict each other in most design situations defines concept evaluation as a multi-criteria decision-making (MCDM) problem subjected to uncertainty owing to multiple and mostly conflicting criteria and imprecise information during this stage (Okudan and Tauhid, 2008).

MCDM methods comprise a large family of mathematical methods that compare the alternatives in a set of alternatives using multiple criteria. In many real-life MCDM problems, different criteria may be expressed in different dimensions (units). According to Triantaphyllou and Sanchez (1997), this issue of multiple dimensions is what makes the typical MCDM problem a complex one. Dong and Hayes (2012) identified some challenges in the use of MCDM methods: a) the difficulty of deciding how to set the weights for the criteria, and b) the relatively high effort required on the part of the decision maker to create the model of the decision situation and enter the parameters.

Perhaps the greatest challenge during concept selection is the scarce information available at this stage and the large uncertainty connected to the values used to perform the evaluation. However, even when the exact same values are utilised in different MCDM methods, different answers can be given for the same problem, as has been shown, for example, by Triantaphyllou and Mann (1995), Yeh (2002) and Honkala et al. (2007). Considering the fact that different MCDM methods use different procedures to score and rank alternatives, these results are not surprising. This means that the decision maker(s) need(s) to consider carefully what selection method is appropriate for the actual case in question (Okudan and Tauhid, 2008) and that one universal concept selection method will not be possible to find. Even when one is applying the same con-
cept selection method to problems belonging to the same type of problem, this same method might not be the optimal method for every case. Yeh (2002) found that different problem data sets might result in a different method being judged to be optimal for the same problem, in a study which investigated how to choose the most appropriate MCDM method according to how well it reflected the decision information content. This complexity may scare potential users away from applying the methods, and ultimately result in bad decisions due to a misunderstanding of the capability of the different methods and the wrong application of them.

According to Honkala et al. (2007), findings and anecdotal evidence indicate that MCDM methods have not been broadly accepted and used in industry, although they are widely studied by academia. Birkhofer et al. (2005) reported from a German perspective that few scientific design methods were used in industry and argued that scientists have not taken the needs of industry into account when developing such methods. Birkhofer et al. (2005) were of the opinion that scientists often integrate a large number of prerequisites and abstract models, and clutter the methods which they create with new or unclear phrases, rather than meet industry’s need for methods which are simple and pragmatic, which are applicable with simple tools and always available, which are not time-consuming, and whose expense is reasonable in relation to the benefit obtained. Birkhofer et al. (2005) suggested adapting existing methods to the demands of industry. Salonen and Perttula (2005) found that formal concept selection methods were seldom used during actual decision making in Finnish industry, and that, rather, companies used informal methods. On the other hand, companies that reported having used formal methods were more satisfied with their decision-making process than those which reported having made informal decisions (Salonen and Perttula, 2005).

2.3.1 Relevant concept selection methods

Concept selection methods can broadly be categorized according to two sub-phases of concept selection: methods for concept screening and methods for concept scoring. During concept screening, a great number of concepts are quickly screened to sort out a subset of promising concepts. This subset is typically developed further and a more comprehensive concept-scoring method is applied to help the decision maker(s) determine which concept or group of concepts should proceed forward in the product development process after the conceptual design phase. The total score can be viewed as a summary of the decision makers’ opinions on the degree to which the alternative concepts meet the product development goals. In this section, concept-screening and concept-scoring methods that are relevant to the present work are described.
2.3.1.1 Concept screening

As part of their systematic approach to engineering design, Pahl and Beitz developed a systematic selection chart to eliminate unsuitable concepts and assess the preference for those remaining (Pahl et al., 2007). In this chart, each alternative is evaluated and the only alternatives pursued should (Pahl et al., 2007):

a) be compatible with the overall task and with one another,
b) fulfil the demands of the requirements list,
c) be realizable in respect of performance, layout, etc., and
d) be expected to be within the permissible costs.

Further, a preference is justified if, among a large number of alternatives, that preference (Pahl et al., 2007):

e) incorporates direct safety measures or introduces favourable ergonomic conditions, and/or
f) is preferred by the designer’s company; i.e. can be readily developed with the usual know-how, materials and procedures, and under favourable patent conditions.

In addition, the team might consider if the available information is adequate or more information should be collected to enable them to re-evaluate a concept. As soon as a criterion leads to the elimination of a proposal, then the other criteria need not be applied to it. The reason for eliminating any concept is recorded in the chart, which serves as documentation (Pahl et al., 2007).

2.3.1.2 Concept scoring

Okudan and Tauhid (2008) made a literature review of concept-scoring methods from 1980 to 2008 and found that the methods could be classified into categories based on the underlying principle used, i.e. concept-scoring methods based on:

- decision matrices
- the analytical hierarchy process (AHP)
- uncertainty modelling
- utility theory/economic models
- optimization models
- optimization concepts
- heuristics.

The boundaries between these categories are not evident in all cases, and several concept-scoring methods use two or more of the underlying principles. The categories relevant for the present research are concept-scoring methods based on utility theory and the AHP.
Utility theory was originally developed for economic decision-making (Otto and Antonsson, 1993). Pahl and Beitz have incorporated it into their systematic engineering design process (Pahl et al., 2007). The process involves the following stages (Pahl et al., 2007).

1) Objectives are derived from the requirement list and general constraints are defined from which evaluation criteria can be derived. All the criteria must be given a positive formulation, such that a higher value indicates an improvement. This stage can be systematized by the use of an objectives tree, in which the individual objectives are arranged in hierarchical order. The evaluation criteria can be derived from the sub-objectives on the lowest level of the hierarchy.

2) Weighting factors are assigned to each criterion and indicate the relative importance of a particular evaluation criterion. The use of an objectives tree is suggested to facilitate the process. The sum of the factors of all the evaluation criteria (sub-objectives on the lowest level) must be equal to 1, so that a percentage weighting can be attached to all of the sub-objectives. The evaluation of weights proceeds step-by-step from a level of higher complexity to the next lowest level. The weighing factors at any level must add up to 1. The relative weighting of an objective at a particular level with respect to the objective at the highest level is found by multiplication of the weighting factor of the given objective level by the weighting factors of the higher objective levels.

3) A known or analytically determined parameter is assigned to each evaluation criterion. These parameters should either be quantifiable or expressed by statements framed as concretely as possible.

4) The values to be assigned to each evaluation parameter are assessed. A value function connects values and parameter magnitudes. The values are expressed by points, typically chosen from a range from 0 to 10, but other ranges can also be employed. The sub-value for each criterion for each concept is found by multiplying the corresponding value by the determined weighting factor.

5) The overall value for each concept is determined by adding up all its sub-values.

6) There are different possible ways to evaluate and rank the concepts. Typically, the concept with the largest overall value is considered to be the best.

7) Evaluation uncertainties arise due to subjective errors and procedure-inherent shortcomings and should be considered. It is suggested that subjective errors should be avoided, for example, by performing the
evaluation with the involvement of several people, discussing the alternatives in a neutral way, and using suitable and non-dependent evaluation criteria. It is suggested that procedure-inherent shortcomings should be assessed, for example, through estimates of the mean error, using verbal estimates if parameters cannot be expressed in figures with some accuracy, and using fuzzy multi-attribute decision-making methods.

8) The concepts with the highest overall values should be examined for weak spots (weak spot analysis). This is accomplished through graphs of the sub-values, e.g. a value profile. In a value profile, the length of a bar shows the value corresponding to a requirement and the thickness of the bar corresponds to its weight. Consequently, the area will correspond to the weighted sub-value. A line representing the overall weighted value of the concept is drawn and weak spots are identified as bars whose length is below the average bar length. Thick bars are more important to address than narrow bars. Concepts with a similar overall weighted value can be examined for weak spots, and the concept with a more balanced value profile is considered to be better. The search for weak spots can also be used to find out how promising concepts can be improved.

The AHP was developed by Saaty (1980) as a decision-making tool involving multiple attributes. Marsh et al. (1993) developed a version of the AHP which was intended for design decision making and which uses pair-wise comparisons to determine a weight vector for all the alternatives. The vector is calculated as the sum of one row (i.e. the weightings for one concept) divided by the sum of all the rows (i.e. the weighting for all the concepts). This type of approach can be used to determine the weighting factors in other methods, for example as an alternative to stage 2 in the Pahl and Beitz method. Johannesson et al. (2013) suggested a matrix where the evaluation criteria are entered in both the columns and the rows. The criteria are subjected to pair-wise comparison and share a value of 1. The criterion that is more important than the other receives a one and the less important criterion receives a zero. If the criteria are equally important, they each receive 0.5. The entries in each row are added and divided by the total sum of all the rows to give the relative weight of each evaluation criterion.

2.3.1.3 The modified Pahl and Beitz method

The concept selection method applied in the first part of Stage 5 of the research performed for this thesis (see Chapter 8) was proposed by Prof. Jan Lundberg of Luleå University of Technology (LTU) and is briefly described in Lundberg
This method is based on the method developed by Pahl et al. (2007), which incorporates utility theory in a systematic engineering design process (partly described in Section 2.3.1.2), and on the experience gained from applying variants of this method in industry and among students for about 20 years. The variant developed by Prof. Jan Lundberg presented in this section represents an adaption of the method to the needs of the Swedish Rail Administration, which was the predecessor of the Swedish Transport Administration. In the present thesis, this variant is referred to as the “modified Pahl and Beitz” (MPB) method. The method devised by Pahl et al. (2007) has been thoroughly tested in industry and academia and normally appeals to engineers since it is structured and systematic. According to Ulrich and Eppinger (2014, p. 36), structured methods are valuable because they have a clear decision-making process, fill a function as “checklists” and are self-documenting. Moreover, the method devised by Pahl et al. (2007) is able to combine greatly different aspects that are not related to each other.

The MPB method starts with a task clarification that is performed with the objective of a) supporting the collection of relevant information required to prepare a complete and relevant product specification, and b) increasing the competence among the participants involved in the product development group. Practically, this is achieved by answering the questions specified in Table 2.1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What is the core of the problem?</td>
</tr>
<tr>
<td>2</td>
<td>What hidden wishes, demands and exceptions are involved?</td>
</tr>
<tr>
<td>3</td>
<td>What are the tasks which the product has to solve?</td>
</tr>
<tr>
<td>4</td>
<td>Which possibilities are open and which possibilities are closed for the achievement of the new product?</td>
</tr>
<tr>
<td>5</td>
<td>What properties should the product have/not have?</td>
</tr>
<tr>
<td>6</td>
<td>What suitable products exist today?</td>
</tr>
<tr>
<td>7</td>
<td>What are the demands according to standards and regulations?</td>
</tr>
<tr>
<td>8</td>
<td>What are the wishes and demands regarding possibilities for product upgrading?</td>
</tr>
<tr>
<td>9</td>
<td>Are all the demands and wishes really necessary?</td>
</tr>
<tr>
<td>10</td>
<td>What are the future trends regarding technology, environmental aspects and ergonomics?</td>
</tr>
</tbody>
</table>

The next step involves formulating a clear goal for the product development, based on the information that has been revealed by answering the questions in Table 2.1. This goal should reflect the most important aspects of the development work and act as a lodestar during the product development. It is essential that all the participants in the product development team should understand and agree upon this goal before proceeding.
In the next step, the information revealed during the task clarification and the goal specification steps serves as a background for formulating a requirement specification for the product to be developed. To assist them, the participants have the list of headings and examples originally compiled by Pahl et al. (2007, p. 149): geometry, kinematics, forces, energy, material, signals, safety, ergonomics, production, quality control, assembly, transport, operation, maintenance, recycling, costs and schedules. In the list of headings used in the MPB method, “recycling” has been replaced by “environment” to cover a wider range of environmental issues, including pollution. “Design” and “marking and notations” have been added to take any requirements regarding aesthetics into account. Finally, “handling” has been added to the list to cover procedures which must be taken into account during installation and operation, for example, and “operation” has been removed, since it is regarded as being covered by other headings. Table 2.2 shows the list of headings used in the MPB method, together with examples of properties that might be relevant to consider under each heading. All the headings may not be relevant in every case. Under each heading, the group specifies the requirements and the wishes for the product to be developed. The requirements must under all circumstances be fulfilled, whereas wishes do not necessarily have to be fulfilled, but it would be beneficial if they were to be fulfilled. A wish is coupled to a requirement. An example of a requirement is a maximum purchase price of x euros, and an example of a coupled wish is a desired purchase price of less than x euros. The requirements and wishes are stated in a requirements list, where each wish is numbered.

Since the wishes are associated with parameters that typically have different dimensions, they need to be expressed in the same units so that one can calculate a total score for each concept. This is accomplished by transferring the magnitudes of the parameter associated with each wish into a number of points from a range of 0-10, where “0” corresponds to “the value for this property is deemed identical with the requirement limit” and “10” corresponds to “a theoretically perfect solution”. The points in between are assumed to be distributed linearly.

Whereas the requirements are fixed and can only be either fulfilled or unfulfilled, the degree to which the wishes are fulfilled can be used to assign a concept score to those concepts which fulfil all the demands. The wishes are typically not of equal importance and, therefore, to calculate a total score, the wishes need to be assigned a weight that reflects how important they are in comparison to the other wishes. The weights are calculated by means of a matrix where the wishes are compared in pairs, similarly to the method suggested by Johannesson et al. (2013). The wishes are put in consecutive order in rows and columns, and the rows are compared with the columns. If the wish in the row is less important than the wish in the column, a “0” is given in the related matrix element; if the wishes are equally important, a “1” is assigned, and if the wish in the row is more important than the wish in the column, a “2” is given. The weight of the wish given in each row is found by summing the elements in each
### Table 2.2: Headings used to create the specification of requirements. Developed from Pahl et al. (2007, p. 149) by Prof. Jan Lundberg.

<table>
<thead>
<tr>
<th>No.</th>
<th>Heading</th>
<th>Examples of properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Geometry</td>
<td>Dimensions, size, tolerances, space requirements, arrangement, connection, extension</td>
</tr>
<tr>
<td>2</td>
<td>Kinematics</td>
<td>Type and direction of motion, velocity, acceleration, retardation</td>
</tr>
<tr>
<td>3</td>
<td>Forces</td>
<td>Magnitude and direction of force, frequency, weight, load, pressure, deformation, stiffness, inertia forces, resonance, elasticity, stability, durability</td>
</tr>
<tr>
<td>4</td>
<td>Energy</td>
<td>Input and output power, efficiency, loss, friction, ventilation, pressure, temperature, heating, cooling, supply, storage, capacity</td>
</tr>
<tr>
<td>5</td>
<td>Material</td>
<td>Physical and chemical properties of the final product, regulations</td>
</tr>
<tr>
<td>6</td>
<td>Signals</td>
<td>Display, control equipment, signal form, signalling scheme, input and output of measurement signals</td>
</tr>
<tr>
<td>7</td>
<td>Safety</td>
<td>Legislation, direct safety systems, warning systems, operational safety, environmental safety, safety factors</td>
</tr>
<tr>
<td>8</td>
<td>Ergonomics</td>
<td>Man-machine relationship, operating height, clarity of layout, sitting comfort, lighting, grip friendliness, shaping, visual field</td>
</tr>
<tr>
<td>9</td>
<td>Production</td>
<td>Possible aspects related to the production process of the product</td>
</tr>
<tr>
<td>10</td>
<td>Quality control</td>
<td>Possible tests and measurement methods, crack indication, delivery control, measurement parameters that should be quality-controlled</td>
</tr>
<tr>
<td>11</td>
<td>Assembly</td>
<td>Modularization, installation, fitting-in, time aspects, tools</td>
</tr>
<tr>
<td>12</td>
<td>Transport</td>
<td>Limitations due to equipment</td>
</tr>
<tr>
<td>13</td>
<td>Handling</td>
<td>Sound levels, handling forces, work tasks, design of controls and instruments, climate, surroundings, special field of application</td>
</tr>
<tr>
<td>14</td>
<td>Maintenance</td>
<td>Life span, durability, durability with respect to overload, availability, service intervals, inspections, dismountability, cleaning, cost for standstill, maintenance costs, punctuality, painting, repair friendliness</td>
</tr>
<tr>
<td>15</td>
<td>Costs</td>
<td>Investment costs, purchase costs, agreement costs, total costs</td>
</tr>
<tr>
<td>16</td>
<td>Environment</td>
<td>Emissions, recycling, waste, toxicity, environmental stress, prevention of cruelty to animals, energy consumption during manufacturing, environmental stress from manufacturing</td>
</tr>
<tr>
<td>17</td>
<td>Design</td>
<td>Company style, company logo, user style, shape, lifestyle, colour, form, attractiveness</td>
</tr>
<tr>
<td>18</td>
<td>Marking and notations</td>
<td>Standard notation on the product, methods for marking</td>
</tr>
<tr>
<td>19</td>
<td>Schedules</td>
<td>Deadline for delivery, deadline for installation, deadline for operation</td>
</tr>
</tbody>
</table>
row and normalizing by dividing each sum by the sum of all the sums. To check
that the weights are reasonable, the wishes are sorted in a list in descending
order with respect to their weight. The product development group should as-
ssess whether the order of priority is reasonable. Note that a strong requirement
for a certain property will typically give less importance to a wish coupled to
that requirement.

The subscore of each concept for a certain wish is found by multiplying the
 corresponding value by the determined weight. The concept score is the sum of
all the subscores. The concept with the highest score is the most promising con-
cept to develop further. However, if two or more concepts receive a similar
score, a weak-point analysis can be performed. In that case, the subscores of the
different concepts are compared, and the concept that has the least variation
between the subscores is the preferred choice.

2.4 Design for maintenance

Maintenance is defined as follows in the European standard EN 13306 (CEN,
2010): “[The] combination of all technical, administrative and managerial ac-
tions during the life cycle of an item intended to retain it in, or restore it to, a
state in which it can perform the required function.” The actions performed to
maintain an asset are classified into two main categories: preventive mainte-
nance and corrective maintenance. Preventive maintenance includes actions
performed to control the rate of degradation and to reduce the likelihood of
failure occurrence. Corrective maintenance, on the other hand, consists of ac-
tions performed to restore a failed asset to its normal operational state. In other
words, preventive maintenance is performed before the asset diverges from its
normal operational state, in order to avoid operational disturbances, whereas
preventive maintenance occurs after the asset has diverged from its normal op-
erational state. According to EN 13306 (CEN, 2010), preventive maintenance is
divided into the subcategories of condition-based maintenance and pre-
determined maintenance. Condition-based maintenance is initiated by mea-
surements of the actual performance of the asset, whereas pre-determined
maintenance is performed according to a schedule based either on time or us-
age. Because systems and products have become more complex, preventive
maintenance has become increasingly important to avoid sudden breakdowns
that disrupt normal operation, require considerable cost to rectify, and, in the
worst case scenario, cause injuries (Pahl et al., 2007). Condition monitoring, the
process of continuously monitoring a parameter connected to the condition of
an asset, allows preventive maintenance to be performed at the optimal point in
time.

Designers significantly influence maintenance costs and procedures as they
select principle solutions and embodiment features, which strongly affect the
maintainability (Pahl et al., 2007). Technical measures that can reduce mainte-
nance efforts should be considered already during the conceptual design phase. Maintenance requirements should be included in the requirements list (Pahl et al., 2007) and considered during the entire design process. In principle, the aim of developing a technical solution is complete freedom from the need for service by incorporating features in the solution that make maintenance unnecessary or reduce it substantially; and only when such features cannot be realized or are too costly, should maintenance be introduced (Pahl et al., 2007). Consequently, one aim of maintenance optimization is to minimize the life cycle cost (LCC).

2.5 Life cycle cost

The LCC can assist the designer and the decision-maker in finding the right trade-off between the physical design of a product and its required maintenance. The LCC takes into account the total cost of ownership of an asset and includes acquisition costs, operating costs and disposal costs. The aim of life cycle costing is to minimize the costs over the entire life cycle of an asset. By performing LCC calculations, one can determine whether the cost of a more expensive asset will be offset by lower future operating and maintenance costs. To compare costs and benefits that occur in different time periods, discounting is typically used.

The LCC approach is an accepted strategy in the railway community and has been used, for example, in the European project Innotrack as a foundation for making choices (Nissen, 2009). By presenting quantified values, one can gain a better understanding of the type of changes in the design and maintenance strategy that can be planned to lower the life cycle cost (Nissen, 2009). Railway infrastructure and particularly track components are expensive assets with long life spans. For many complex assets, the cost of maintenance plays an important role in the LCC analysis, and is especially important for assets like track infrastructure, whose operation and maintenance phase comprises a major share of the system’s life cycle (Patra et al., 2009). Therefore, it is relevant to consider the cost of ownership per year, which is equal to the LCC divided by the expected lifespan.

All the participants in the process – operators, contractors and vendors – can have a substantial impact on the overall cost of ownership, and it is not until all of them are involved that the benefits sought from the use of life cycle costing will be realized.
3. Research design

This chapter explains the research approach, research method and research process used for the work performed for this thesis. The OptiKrea Group is also described in this chapter, as this is the cross-functional inter-organizational group that has been followed throughout the research work. As described in Section 1.1, the research for this thesis was divided into different stages, although several of them were running in parallel at times. An overview of the stages is to be found in Table 1.1. The specific procedure, data collection and analysis methods used in each stage of the research are detailed in connection with the presentation of the results of each stage to enhance the readability of the thesis. Participants and groups other than the OptiKrea Group are detailed in connection with the presentation of the stage concerned.

3.1 Research approach

Based on the research proposal and objectives, it was found that it would be impossible to separate the studied phenomena from their context. Engineering design is a complex process; the more artificial the evaluation setting is, the more difficult it becomes to assess whether the findings are likely to translate into ideas applicable to practical work environments (Dong and Heyes, 2012). Even if a method is good in theory, it might be too burdensome for people to use it outside a laboratory setting, or people might simply not like it. Preferring one method to another does not necessarily imply that the preferred one will result in better performance, but preferences are important in that they suggest a willingness to accept a tool and implement it in everyday work life. People are unlikely to benefit from an interface which they do not like and are therefore unwilling to use (Dong and Heyes, 2012). Therefore, it is desirable to conduct studies of the application of conceptual design methods in real-life contexts and using real tasks, when the purpose is to develop methods for implementation in a certain context. Studies in real-life contexts, however, create many challenges. The primary challenge is that real problems and natural settings are harder to control and may introduce variability into tests and experiments. However, the benefit is that the results obtained in this way are more likely to reflect what occurs outside the laboratory (Dong and Heyes, 2012). In the present work, a semi real-life context was created through the OptiKrea project. The participants were all professionals working in the intended use context and together they formed a group that was representative of the context which the methods
were supposed to be developed for. Throughout the project the group worked on real-life problems that they had themselves suggested. In a sense, the entire project was a test of the proposed methods of working, which were refined through the course of the project. Performing research within the project gave the opportunity to have some control of the context, but it was more similar to a natural setting than a laboratory setting. Hence, the research is more qualitative than quantitative, although quantitative measures have been used when possible. By employing action design research (ADR), the participants and researchers in the project were all contributing to the development of the conceptual design methods.

3.2 Research method

ADR was used to fulfil the purpose and objectives of the present research. The ADR method was proposed by Sein et al. (2011) in an effort to blend design research with action research. Design research is the study of artefacts in their context (Hevner et al., 2004), whereas action research, interpreted generally, is intervention in a social situation in order to both improve this situation and learn from it (Susman and Evered, 1978). The purpose of ADR is to generate prescriptive design knowledge through learning from the intervention of building and evaluating an information technology (IT) artefact in an organizational setting to address a problem (Sein et al., 2011). Following Sein et al. (2011), artefacts are viewed as ensembles for the purpose of this thesis, meaning that the interaction between design efforts and contextual factors throughout the design process is manifested in the form, structure, goals, and conceptualization of the developed artefact. According to Rogerson and Scott (2014), it is uncommon that a design-based intervention will turn out as planned at the first attempt, especially when addressing a social situation with practitioners involved. Therefore, ADR is problem-driven and aims to build design principles based on iterative cycles in the same context (Wieringa and Morah, 2012).

ADR was initially proposed within the field of information systems and has been applied in several different contexts, e.g. addressing the learning needs of the deaf community (Golding and Tennant, 2013), managing projects funded by a third party (Gröger and Schumann, 2014) and greenhouse gas emission reporting in the meat industry (Hilpert et al., 2013). An IT artefact is a specific bundle of hardware and software that is assembled to fulfil information needs. Therefore, it was considered reasonable to assume that the development of non-IT-based tools for information handling could also profit from ADR. A type of ADR had previously been applied to develop a framework for designing governance structures for major transport corridors (Öberg, 2014). As the present research aimed to develop conceptual design methods for a certain context through interaction between researchers and engineers with different functional knowledge and from different organizations, it was deemed that ADR would be useful in guiding the study, as its strict and explicit principles and iterative cy-
cles allow the researcher to deliver a practical outcome for the involved organization whilst simultaneously meeting academic standards. No previous studies have, to the best of the author’s knowledge, explicitly used ADR to generate prescriptive knowledge on concept generation and selection methods through interaction between researchers and engineers in a real-world context. However, other researchers have applied action research to study developed methods, e.g. a computer support system for systematic design (Almefelt et al., 2003) and methods for service innovation in a multi-disciplinary context (Rexfelt et al., 2011). An important difference between the approaches applied in these studies and those applied in the present research is that in the latter case the target context was allowed to shape the proposed methods through cycles of building, intervention and evaluation (BIE) until the participants felt no need for further refinement.

ADR consists of four stages (Sein et al., 2011): 1) problem formulation; 2) building, intervention (i.e. test of the artefact in the target environment), and evaluation; 3) reflection and learning; and 4) formalization of learning. Each stage is based on certain principles and involves the execution of certain tasks. An overview of the method is shown in Figure 3.1. Section 3.3.1-3.3.4 describe the stages, including their principles and tasks, and their application in the present research. Reflections on the feasibility of using ADR for developing methods in the area of engineering design are to be found in Chapter 11.

3.3 Research process

3.3.1 Problem formulation

The first ADR stage is problem formulation, guided by the principles of practice-inspired research and a theory-ingrained artefact (Sein et al., 2011); i.e. existing knowledge should be used to create an artefact that addresses a practical need interwoven with a context. The first two tasks in this stage are to identify and conceptualize the research opportunity and to formulate initial research questions and objectives.

The trigger for the present study was the insight that, since different functions involved during the life cycle of railway products are spread out over different actors, a lack of suitable collaboration strategies increased the risk of suboptimization during product development. Specifically, the conceptual design phase was of interest, since it has a considerable influence on the subsequent steps during product development with regard to cost, quality and performance (e.g. Rubenstein, 1994; Okudan and Tauhid, 2008). By involving representatives from several actors during the conceptual design phase, it was thought that more viewpoints on the product to be developed would be shared and thus a better product with higher maintainability would eventually be designed. A way to address this issue is to use concept generation and selection methods to structure the interaction between participants representing different
**Figure 3.1** The stages of ADR and their adherent principles and tasks. Adapted from Sein et al. (2011).
actors. The research objectives are to be found in Section 1.1. It was very easy to recruit different railway actors to the project, which was interpreted as a sign that a real need had indeed been identified. Three studies were conducted to explore and understand the current situation and prerequisites for product development in the context of the Swedish deregulated railway market.

The third task of this stage is to cast the problem tentatively as an instance of a class of problems, as this facilitates for the researcher to generate knowledge that can be applied to the class of problems that the specific problem exemplifies (Sein et al., 2011). No concept generation or selection methods had previously been developed specifically for groups with members possessing different functional knowledge and representing different organizations. According to Straus et al. (2011), there is no knowledge of what happens in inter-organizational groups, as different actors bring different cultures and agendas with them to such groups. Nor could the author find any studies on this topic during a literature review. Therefore, it was concluded that design principles for concept generation and selection methods to be used in cross-functional inter-organizational groups would be an interesting research topic, as deregulation and outsourcing have become increasingly common. As a result, the research of this part of the project was initially framed as addressing the following class of field problems: concept generation and concept selection in cross-functional inter-organizational groups.

Task 4 involves identifying contributing theoretical bases and prior (technology) advances. A literature review was performed on idea generation, decision making and methods for the conceptual design phase. A test of three established concept generation methods was executed among the OptiKrea project participants to find empirical evidence as to how a cross-functional inter-organizational group interacts during concept generation.

Task 5 involves securing long-term organizational commitment. The participants were carefully selected in consideration of their background, expertise and function within the organization, so that they could give a representative view of what would be useful for the actor which they represented. Formal letters of intent regarding the contributions from each participating organization were signed, as well as intellectual property agreements. The organizational commitment was followed up during the entire project and secured through adherence to the project plan, keeping the participants updated, and showing them that concrete results were achieved which each organization could benefit from. The last task of this stage involves clarifying roles and assigning responsibilities. On the Swedish deregulated railway market, different stakeholders with respect to the development of infrastructure-related products were identified: the infrastructure manager (the Swedish Transport Administration (STA)), product manufacturers, contractors performing maintenance of the products, and academia. In Sweden, the infrastructure manager typically runs its research and development projects in cooperation with or through research institutes and universities, and therefore academic researchers were relevant stakeholders. All
these stakeholders in the case of turnouts were represented in the project group of seven participants. Two academics were given the role of leading the project and managing the research part, and formed the “scientific team”. The other participants and one of the academics in the scientific team formed the “creative team”, in which issues concerning turnouts would be proposed and addressed by means of the methods proposed by the scientific team. The scientific team would analyse the data from the interventions and propose changes to the method under development. The creative team would give their views on the methods after the interventions and comment on any changes which the scientific team would propose. For the scope of this thesis, the “creative team” is hereafter called the “OptiKrea Group” to distinguish it from the teams employed in the field tests.

3.3.2 Building, intervention and evaluation

Carried out as an iterative process in a target environment, the BIE phase interweaves the building of the artefact, intervention in the organization, and evaluation, and the outcome of the BIE stage is the realized design of the artefact (Sein et al., 2011). This stage is guided by the principles of reciprocal shaping, mutually influential roles, and authentic and concurrent evaluation. The principle of reciprocal shaping implies that increased understanding of the organizational context influences the design of the artefact, and the artefact influences the practices in the organizational context. The principle of mutually influential roles emphasizes the different types of knowledge which the project participants bring with them and the mutual learning among the different participants (Sein et al., 2011). The principle of authentic and concurrent evaluation emphasizes that the evaluation should be on-going and interwoven with the activities throughout the BIE stage (Sein et al., 2011). The evaluation should occur spontaneously in the organizational context, and opportunities should be sought following natural controls when possible, rather than in a controlled environment (Sein et al., 2011; Rogerson and Scott, 2014).

The first task of the BIE stage is to define an initial knowledge-creation target. In the present research, the target was structured conceptual design methods to be used in cross-functional inter-organizational groups on the Swedish deregulated railway market, as was described in detail in Section 1.1.

The next task involves selecting or customizing the form of BIE to be implemented. According to Sein et al. (2011), the form of BIE to be used should be determined by whether the innovation is mainly technological or organizational, and they suggest a generic BIE scheme for each case (Sein et al., 2011). In the case of concept generation and selection methods, this was re-interpreted to mean that the form of BIE to be used should depend on whether the focus is mainly on the method design or organizational intervention. The form of BIE was customized to suit the specific case and the given circumstances, and the BIE was adapted according to the possibilities appearing in the course of the
project. As an example, the BIE scheme adopted in the development of the concept generation method is given in Figure 3.2.

![Diagram of BIE scheme](image)

**Figure 3.2** The form of BIE adopted in the development of the concept generation method. Adapted from Sein et al. (2011).

In the case of the development of the concept generation method, the iteration cycles started in the creative team (the “alpha phase”). Connecting relevant findings from the literature review and the evidence derived from tests of established concept generation methods in the creative team with the identified requirements, tentative design principles were formulated in the scientific team in parallel with setting up an initial concept generation method. The suggested method was discussed in the group and a few changes were introduced before the performance of interventions. The creative team was quite satisfied with the method already after the first iteration. After the second iteration, the participants were satisfied with the method and did not want to make any further modifications. The successfully introduced changes were conceptualized as design principles and added to the tentative design principles.

At this stage, the question arose as to whether the act of taking part in the development of the method had biased the creative team’s satisfaction with it, and it was desirable to try it in a wider context and preferably with several groups. This opportunity was provided in two field tests, one in the EU project In2Rail and one in a Swedish project called UPPSAMT (the field tests are reported in Chapter 7). These field tests allowed the performance of two comprehensive interventions, corresponding to the “beta phase”, which involved judging the value of the method in typical use settings. The tests in the beta phase resulted in unexpected events that gave new insights into the method.

In the case of the concept selection method, the starting point was the modified Pahl and Beitz (MPB) method (see Section 2.3.1.3). The MPB method was developed for the same context as that described in the present thesis, but had never been implemented. It was suspected that the method had become too
cumbersome and that an “encounter with reality” would reveal why it had not been implemented and, together with the findings from the literature review, lead to an improved method. After the first cycle, the scientific team carefully considered all the data from the intervention and all the results from the evaluation and proposed a revised method. After the second cycle, the scientific team proposed some minor changes to the revised method, which the participants in the creative team agreed upon, and after that they agreed that they did not need to try the method one more time.

3.3.3 Reflection and learning

The reflection and learning stage, a continuous stage that parallels the first two stages, moves conceptually from building a solution for a certain case to applying the learning derived from that process to a broader class of problems (Sein et al., 2011). Continuous reflection also allows adjustment of the research process according to increased understanding of the ensemble constituted by the artefact (Sein et al., 2011). This stage is guided by the principle of guided emergence, which emphasizes that the ensemble will reflect not only the preliminary design created, in this case, by the scientific team, but also its ongoing shaping by organizational use, different perspectives and the participants, and by the outcomes of authentic and concurrent evaluation (Sein et al., 2011). In other words, the reflection and learning stage emphasizes the importance of incorporating the outcome of addressing Principle 1-5 (see Figure 3.1) in the final artefact.

The first task of the reflection and learning stage is to reflect on the design and redesign of the artefact during the project. The concept generation and selection methods developed in the present research have been adapted continuously according to the evaluations that have taken place, to reflect the increasing understanding of both the organizational context and the emerging methods. The increased understanding of the method characteristics clarified what broader class of problems the methods may be applied to. Task 2 involves evaluating the adherence to the principles and task 3 is to analyse the intervention results according to the stated goals. These two tasks were performed throughout the course of the project and at the end of the project.

3.3.4 Formalization of learning

The fourth ADR stage involves the formalization of learning and draws on the principle of generalized outcomes, which emphasizes that generalization is a challenge since the artefact has been developed to address a specific situation (Sein et al., 2011). Generalization is accomplished by viewing the developed artefact as a solution that addresses a problem, and through making a conceptual move from “the specific and unique to the generic and abstract” on three levels (Sein et al., 2011): 1) generalization of the problem instance, 2) generalization
of the solution instance, and 3) derivation of design principles from the design research outcome, i.e. recommendations on how methods addressing the same type of problems should be designed. The first and third task of this stage involve abstracting the learning into concepts for a class of field problems and articulating the outcomes as design principles.

The developed concept generation method has been viewed as belonging to the class of concept generation methods for cross-functional inter-organizational groups. Reconceptualizing the learning from this specific instance into design principles for the class of solutions to which it belongs was performed tentatively in the first proposal for the new method. The tentative design principles were confirmed during the alpha and beta cycles, and three additional design principles emerged as a result of the interventions.

The class of field problems for the concept selection method had to be re-framed. It was judged that the concept selection method should belong to the class of methods that facilitate the selection of one or more large-investment products that are to contribute to the provision of a public utility in the railway sector. The design principles were formulated for this class of solutions.

The second task involves sharing the outcomes and assessment with practitioners. The practitioners involved in the project were continuously updated on its progress and emerging outcomes. A step that belongs to this task, but has not yet been executed, is the implementation of the methods in the routines of STA. Furthermore, the outcome was shared through presentations, workshops and publications.

Task 4 involves articulating the learning in the light of the theories selected in Stage 1 and task 5 involves formalizing the results for dissemination. It is an important step to relate the findings to previous research and theories, and this step is not restricted to ADR. The developed artefacts and the learning from the development process have been viewed in the light of relevant literature.

3.4 The OptiKrea Group

The OptiKrea Group had six participants, four railway professionals and two academics. The participants were carefully chosen based upon their experience from the railway sector and knowledge of their respective organizations. Hence, they were able to give representative views on what would be useful for the actors which they represented. Each participant had deep knowledge of the organization which they represented, which made it possible for them to relate the application of the methods to a real-life context and judge how useful it would be in other typical use contexts. Each of the railway professionals had 20-30 years of experience from the railway sector. Two professionals worked with turnout-related issues at the infrastructure manager (STA), one professional worked at an international company manufacturing turnouts and other railway assets, and one at a contractor performing maintenance. The academics worked with railway-related projects in close cooperation with STA and industry. De-
tails of the participants are given in Table 3.1 and a schematic representation of the group is shown in Figure 3.3.

Table 3.1 Details of participants, valid for the period during which the OptiKrea project took place.

<table>
<thead>
<tr>
<th>No.</th>
<th>Organization</th>
<th>Position</th>
<th>Field of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>Maintenance methods and product development</td>
</tr>
<tr>
<td>B</td>
<td>Supplier</td>
<td>Development manager</td>
<td>Product development of turnouts, turnout engineering</td>
</tr>
<tr>
<td>C</td>
<td>STA</td>
<td>Specialist</td>
<td>Maintenance management of turnouts</td>
</tr>
<tr>
<td>D</td>
<td>STA</td>
<td>Specialist</td>
<td>Track and turnouts</td>
</tr>
<tr>
<td>E</td>
<td>Academia</td>
<td>Professor</td>
<td>Mechanical engineering design, maintenance, product development</td>
</tr>
<tr>
<td>F</td>
<td>Academia</td>
<td>Postdoc. research fellow</td>
<td>Applied acoustics and signal processing, maintenance</td>
</tr>
</tbody>
</table>

STA is the only national buyer of services and products from the contractor and supplier, resulting in an asymmetric power relation between the organizations which the participants represented. Similarly, academia is dependent on funding from STA in many projects, and several different universities and university divisions compete for such funding. None of the participants had tried any of the ideation methods before, except for participant E, who had previous experience of Method 635 and the gallery method, from both student and industrial contexts for a period of almost 30 years. None of the participants had experience of concept selection methods similar to the one developed in Chapter 8, except for participant E, who had extensive experience of such methods. All of the participants were males.

Figure 3.3 Schematic representation of the OptiKrea Group.
4. Assessment of the Swedish deregulated railway market

4.1 Introduction

This chapter presents Stage 1 of the research conducted for this thesis (see Table 1.1). This stage concerned a qualitative assessment of the prerequisites for collaborative product development on the Swedish deregulated railway market. This assessment was carried out partly before the second stage started and partly in parallel with the other research stages. The assessment was performed in three different studies, each with a different objective.

The objective of Study 1 was to investigate the impact of the current procedure for procuring maintenance contracts, and especially its effects on the development of maintenance methods by maintenance contractors.

The objective of Study 2 was to explore the status of the collaboration and coordination between the different railway actors on the Swedish deregulated railway market.

The objective of Study 3 was to investigate systematically the current process for product development on the Swedish deregulated railway market, from the perspective of the Swedish Transport Administration (STA), and especially to understand what distinguishes this product development context from a typical such context.

The following research questions were posed to address the objectives.

- What is the impact of the current procedure for procuring maintenance contracts and, especially, what is its effect on the development of maintenance methods by maintenance contractors?
- What is the status of the collaboration and coordination between different railway actors on the Swedish deregulated railway market?
- What is the process for product development on the Swedish deregulated railway market, from STA’s perspective, and is there something that distinguishes this product development context from a typical such context?

The method for and results from each study are presented in Section 4.1-4.3, respectively. Discussions of and conclusions from the results from the three studies are given in Section 4.4 and 4.5, respectively.
4.2 Study 1: The impact of procuring maintenance contracts

4.2.1 Method

4.2.1.1 Participants

The first round of data collection consisted of interviews held with one person from STA, one from a maintenance contractor company and one from a supplier company. The respondents are detailed in Table 4.1 (Respondent 1-3). These respondents were deliberately chosen because they had worked at the Swedish Rail Administration (SRA) before the deregulation started, had an engineering background and had been working within their respective segments of the railway industry for 20-30 years. Moreover, all of them were members of the OptiKrea Group. The respondents had worked in different positions and possessed a broad knowledge of the overall situation within their respective organizations. Therefore, the validity of their answers is high regarding the situation which they can describe. The first round of interviews took place at the end of 2012 and the beginning of 2013.

In the second round of data collection, another three respondents were interviewed. They were chosen because of their long experience from the Swedish railway sector and because they had other perspectives than the first three respondents. These respondents are detailed in Table 4.1 (respondents 4-6). Respondent 4 had previously worked at SRA and STA. Respondent 5 had been a field worker before becoming a union representative. Respondent 6 had previously worked at STA. The second round of interviews took place at the end of 2013 and the beginning of 2014.

All the respondents were males.

<table>
<thead>
<tr>
<th>No.</th>
<th>Organization</th>
<th>Current position</th>
<th>Railway experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>STA</td>
<td>Specialist</td>
<td>24 years</td>
</tr>
<tr>
<td>2</td>
<td>Supplier</td>
<td>Development manager</td>
<td>31 years</td>
</tr>
<tr>
<td>3</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>27 years</td>
</tr>
<tr>
<td>4</td>
<td>Consulting firm</td>
<td>Consultant</td>
<td>&gt;20 years</td>
</tr>
<tr>
<td>5</td>
<td>Contractor</td>
<td>Union representative</td>
<td>34 years</td>
</tr>
<tr>
<td>6</td>
<td>Consulting firm</td>
<td>Consultant</td>
<td>&gt;15 years</td>
</tr>
</tbody>
</table>
4.2.1.2 Data collection

The data collection was performed using qualitative semi-structured interviews. The first round of interviews concerned a broader range of topics than what is presented in this study, and the overall goals of the interviews were the following:

- to get acquainted with the situation of each actor (infrastructure manager, supplier, and contractor),
- to understand how the development, purchase and maintenance of turnouts work today,
- to understand what works well and what does not work well,
- to understand where there is room for improvement.

Each interview was guided by a set of pre-formulated questions that were adapted to each respondent depending on the organization that he represented. All the interviews in the study were performed by the author. The interview questions (translated from Swedish into English) are to be found in Appendix 4.1 at the end of this chapter. The pre-formulated questions were not followed strictly, but were rather used as starting points and support during the interviews. The respondents were encouraged to develop further their replies to the questions as well as their ideas on other topics which the questions led them to think of. Possible follow-up questions were included in the guide and used when appropriate.

The second round of interviews took its starting point in the preliminary analysis of the first round of interviews. The first part of the interviews in the second round was unstructured, and the respondent and the interviewer discussed deregulation and maintenance from the respondent’s point of view. During the second part, the interviewer brought up the themes that had emerged in the first round of interviews and asked the respondent to give his view on them. Many of the themes had already to some extent been discussed in the unstructured part of the interview.

4.2.1.3 Data analysis

The data analysis in this study took its starting point in the basic strategy for analysing qualitative data presented by Merriam (2007). According to Merriam (2007), all qualitative data analysis is primarily inductive and comparative, but moves from being totally inductive, when trying to derive tentative categories from pieces of data, to being primarily deductive as one is looking for more evidence in support of the final set of categories that have emerged during the comparison and adaption of the tentative categories to subsequent data. For the purpose of the present study, the data from the first round of interviews were thematically analysed according to the procedure described in Table 4.2. The
results of the analysis were sent to and confirmed by all the respondents in the first round before the second round of interviews started.

Table 4.2 The procedure used to analyse the interviews.

1. The recordings of all the interviews were transcribed by the author.
2. The recordings were listened to while the transcriptions were simultaneously perused. The quotes that were interesting with respect to the research questions were marked. Notes were made about the author’s reflections concerning these quotes, for the purpose of assisting in defining categories (themes).
3. Step 2 was repeated for each interview or part of an interview one or several times, depending on the complexity of the text and how many research questions were involved in the text. (Sometimes this involved either listening to the recordings or reading the transcript.)
4. All the marked quotes were collected in a single word processing document. For each quote, information was provided as to which participant had contributed the quote.
5. Quotes that concerned similar matters were listed together. Tentative categories were formed. The data were perused as many times as required to find exhaustive and mutually exclusive categories (i.e. so that all the quotes could fit into a category and no quotes could fit into more than one category). Quotes that concerned more than one category were split if possible, or otherwise the category was changed to reflect the content of the quotes. The categories were given descriptive labels.
6. The quotes in each category were perused and sorted into subcategories as necessary (using the same procedure as for the main categories, see point 5). The author examined how many participants had agreeing views in each category/subcategory. The content of the quotes in each category/subcategory was descriptively summarized.
7. Point 6 was performed for each category.

The data analysis of the second round of interviews was performed similarly to that of the first round, with the difference that data relating to any of the tentative categories were directly sorted under these categories. The author examined whether any data could be found which contradicted the findings from the first round of interviews.

4.2.2 Results

Sections 4.1.2.1-4.1.2.6 summarize the primary data gathered from the respondents. Quotes (translated from Swedish into English) are included to exemplify the qualitative data that contributed to the results. The number after each quote shows which respondent has been quoted according to the numbering in Table 4.1.
4.2.2.1 Forcing the tender bids down to win contracts

The purchase of operation and maintenance for the Swedish railways is governed by the Swedish Act on Procurement within the Water, Energy, Transport and Postal Services Sectors (APWETPSS) (Lag om Upphandling inom Områdena Vatten, Energi, Transporter och Posttjänster, 2007). STA acts as the purchaser of services with the required performance through public procurement. STA specifies the criteria for a procurement in a tender request and contractors show their interest in the contract through tenders that specify the cost at which they will perform the work. The tenders are then evaluated by STA based upon the criteria in the tender request.

According to the respondents, the competition for maintenance contracts is very tough. The basic services, typically including operation and daily maintenance, safety inspection and emergency duty, are included in the basic agreement. The contractors bid a fixed price for the basic services during a contract period of five years, whereas several of the emergency operations and special operations are charged for per hour or operation. Charging for services in addition to the basic services is a way for the contractor to increase their income, as exemplified by the following quote:

“You try, of course, to bid as low as possible, otherwise you don’t get the contract, and then you need to chase those additional costs all the time […], and everyone does that of course.” (3)

Obtaining a contract for provision of the basic services is a way to be in the right position when other jobs appear or are announced. In this way money can be earned. One consequence is that it is advantageous for the contractors if the railway infrastructure is in a bad condition, since this means more emergency work that can be charged for outside the basic service agreement, as exemplified by the following quote:

“So you bid extremely low and then you hope there will be as much additional [work] as possible. And the worse the infrastructure is, the more money the maintenance contractor will earn.” (3)

The basic provisions of the contract are complied with using as little resources as possible, just to ensure that the requirements are fulfilled, which is exemplified by the following quote:

“[The contractors] will break their backs to get the work, I mean underbid each other, to get a contract and then they don’t give an extra inch without getting paid more. And if the contract is not good enough, and they have undercut their bid, then they earn money by using as little resources as possible to complete [the work]. Then the maintenance will be bad.” (5)
One difficulty in procuring a maintenance contract is the lack of knowledge of the events taking place and the actions required in the future, which is exemplified by the following quote:

“Purchasing a normal contract, e.g. [for] replacing a turnout, that’s a completely different matter. [The contract says,] ‘Do this and do that,’ and you can easily do an inspection. Procuring a maintenance contract is something completely different.” (3)

The contractors have difficulties interpreting the tender requests and foreseeing what will have to be done to fulfil the performance requirements. Examples of typical considerations are how much snow removal will be necessary (depending on the weather conditions), the length of time during which the maintenance staff will be able to access the track to carry out their work (influencing how much time the railway workers will have to spend waiting to complete a job), and the condition of the track section in question (determining, for example, how many times the track will have to be tampered to sustain the agreed quality). The contractor is not completely responsible for the track section and there is a possibility that STA, during the time of validity of the contract, will decide to take actions that might benefit the contractor (e.g. upgrade the railway foundation) or might not benefit the contractor (e.g. increase the traffic).

The tender writers try to find the loop-holes in the tender requests that open up the possibility of cutting costs in the tender and having STA paying for additional work outside the basic part of the contract later on. This is exemplified by the following quote:

“If that’s due to STA issuing tender requests that are open to interpretation, so that we can then be cunning and say, well, they have not expressly asked for this, so it’s probably not included and the [APWETPSS] and the general provisions [for contracts by tender] say that we should always calculate the lowest price. And one result of that can be, for example, that, if [STA] hasn’t expressly asked for the heaps [of material resulting from] an excavation or a ballast cleaning to be taken care of, and that can of course be a question of contaminated heaps that must be taken care of for environmental reasons, […] then we reckon we can put the heaps to the one side [at the work site]. And then [STA] comes and says, ‘Oh no, you were supposed to take care of that’. [But then we say,] ‘Well no, that's not written anywhere [in the contract]. Then you have to pay extra.’” (5)

Deviations from the contract specifications, as interpreted by the infrastructure owner or the contractor, have occurred and have led to legal cases, because the specification of what should happen in the event of deviations has not been detailed enough in the contract. This is exemplified by the following quote:
And when [the traffic] increased, all of [the contractor’s] time on the track disappeared. So now they have had terribly high costs, and they want to be compensated for that.” (4)

Eventually it comes down to winning the contract by bidding lower than one’s competitors, even though it is clear from the start that one’s bid is unrealistic, as exemplified by the following quote:

“This money doesn’t at all cover the work we must do.” (3)

If the contract is won, the quest to cover the deficit in other ways starts. How large the profit or loss will be will not be known until the contract period has ended.

4.2.2.2 Optimization of one’s own company’s profit

4.2.2.2.1 Short-term profits

According to the respondents, maintenance was centrally controlled before the deregulation. Maintenance was planned and performed with the aim of profiting in the long-term future and with the spirit of contributing to an overarching societal goal. This is exemplified by the following quote:

“We made sure we prevented faults and the like, to avoid rushing out in the middle of the night and remedying faults and the like. […] [We] used to regard ourselves as a kind of builder of society more than contracted manpower. So it used to be our railway, we used to look after it, because if we didn’t and a fault arose, well, guess who would have to go out and fix it […], and it’s better to do it properly from the start.” (5)

When visiting a site to perform maintenance, a railway worker would simultaneously look for other faults and attend to them. Replaced parts were installed at a site where they would have the right performance or were taken care of and used as spare parts later. This is exemplified by the following quote:

“When I started […] everyone’s reflex action, when you were out doing a job, was to throw [the replaced parts] onto vehicles and then you drove somewhere and took care of everything that was worth taking care of.” (1)

Today the focus is on short-term profits, as exemplified by the following quote:

“When you have to go out and repair and fix things that you shouldn’t have to go out and [fix] at all, [this] also disturbs the train traffic, which in turn costs
money. Our experience is that this is also much worse after this deregulation. Extremely short-term economic profit is governing a lot of things.” (2)

According to the respondents, an important factor for understanding the behaviour of the contractors is the fact that they are companies, and as companies, their goal is to make profits. This is exemplified by the following quote:

“Companies are supposed to generate profit. And that is the only thing that is important for the shareholders. They’re also supposed to conduct themselves properly […], but that is still the bottom line, profit is crucial. That’s what governs them.” (6)

Each company will act in such a way that its own profit is optimized on the basis of the given conditions. It is the view of the respondents that neither the given conditions nor the follow-up evaluation of maintenance performance gives the contractors the possibility of or the incentives for working with the aim of creating a railway that is sustainable in the long term.

4.2.2.2.2 Questionable follow-up of maintenance contracts

There are three different sets of rules and regulations describing how the operation and maintenance of the railway should be executed, governing safety inspection, maintenance inspection and periodical replacement. While safety inspection is included in the basic service agreement of the maintenance contracts, maintenance inspection and periodical replacement are optional. This is connected to the fact that the contracts are performance-based and, therefore, STA is not supposed to direct how the contractors maintain the railway infrastructure, as long as it meets the agreed performance criteria. The recommended rules and regulations are not followed, which according to the respondents affects the results in a negative way. This is exemplified by the following quote:

“Previously there used to be a set of rules and regulations for how the [infrastructure] should be looked after out along the track. And it’s our experience that one is losing competence or isn’t bothering about [applying] them fully. One isn’t lifting things in the proper way, one isn’t doing the jobs one should be doing and this of course affects the result negatively.” (2)

As the contractors carry out the inspections themselves, they might adjust the inspection comments to suit their current work load, as exemplified by the following quote:

“Is it right to assign the inspection to a contractor which itself is then going to perform the job; what possibilities does that lead to?” (6)
Each track section has a track manager who follows up the condition of the infrastructure. Meetings are held regularly where all the parties concerned discuss different issues concerning the section. The following quote exemplifies a problem in this connection:

“It turns into an investigation of what’s been done and what hasn’t been done, more than focusing on the actual point at issue, the problem out on the track.” (3)

The contractors, trying not to exceed their budgets, try to find someone else who could be responsible or to delay work, as exemplified by the following quote:

“So the more brazen you can be, the more you can charge. And the more you can bear your client telling you off, the cheaper it will be for you, the more you will earn from it […]. You earn money by delaying things.” (3)

At the end of a contract period, an inspector is to examine the railway infrastructure to assure that the condition has been maintained at the agreed level. The inspector examines the signal, track and electricity systems. It is difficult to be an expert in all these areas and, consequently, the contractor has an advantage, knowing each area in more detail and being familiar with the railway section concerned. This is exemplified by the following quote:

“And then, of course, I know as a contractor, or because I have technical competence with regard to railway track, what I can cut down on towards the end [of the contract period]; I reduce the rail grinding, I do emergency tamping, I don’t give a damn about packing the macadam; well, in other words, I skip the expensive maintenance measures […] that are a prerequisite for the other things [on the track] deteriorating slowly or being protected, as it were […]. And the [inspector] cannot see that.” (4)

Moreover, according to the respondents, there have been instances when the final check at the end of the contract period has not been performed. Even if the quality of the infrastructure is found to be unsatisfactory by the inspector at the end of the contract period, this will not result in any consequences for the next procurement, since the contractor with the lowest bid will still win regardless of their previous performance. The final check only determines whether or not the track is being taken over in the condition agreed upon. This is exemplified by the following quote:

“There is a takeover inspection and then you get a number of complaints that you have to address before the next contractor takes over.” (5)
One means of control that STA has is the option of rewarding a contractor by extending their contract by two years. It is up to STA to decide if it wants to exploit this option. One respondent thought that it had been common for STA to exploit that option previously, but that it was not so common for them to do so any more. Other respondents, however, were of the opinion that it was customary for STA to make use of that option, except when they were very dissatisfied with the contractor. This is exemplified by the following quote:

“Often [STA has given them] these two extra years, and it has been the case that [STA] hasn’t done so because they are dissatisfied, but then [STA] is manifestly dissatisfied. [Throughout the contract period] there have been incessant and bitter disputes about invoices and such things, which is also a very […] common phenomenon.” (5)

No respondent was able to mention any specific requirement that was related to the exploitation of the two-year contract-extension option.

4.2.2.3 Risk of losing or not utilizing competence and knowledge

The respondents expressed a concern regarding the competence and knowledge possessed by both STA and the maintenance contractors.

4.2.2.3.1 Competence and knowledge at STA

According to the respondents, STA requires the right competence to carry out its mission as an infrastructure manager and as a purchaser of services and products. Firstly, STA requires information and knowledge about the usage, maintenance and degradation rates of every part of the infrastructure, and the right competence to judge the effects of these parameters, in order to take the right decisions about the maintenance of the infrastructure. Resources can then be guided to where they have the greatest effect. This is exemplified by the following quote:

“A lot of information is collected concerning the infrastructure and the rolling stock. And that information should somehow be used for different types of decisions. […] An enormous range of decisions [to be taken] […]. There is a varied need for competence to evaluate the information in the data, and a very high level of competence can be required in complicated contexts, and calculation methods and the handling of statistics, etc. […] Is the overall analysis organized with the competence required to deliver supporting data for that type of decision?” (6)

Secondly, STA needs to have the competence required to transform its knowledge about the infrastructure into tender requests and maintenance con-
tracts that give the maintenance contractors incentives to work towards the overall goal of STA. STA must not only know what to ask for in a tender request, but also be able to judge if the contracts are fulfilled in an honest way.

According to the respondents, the railway-specific knowledge possessed by STA is decreasing and/or is not being utilized. This is exemplified by the following quote:

“There are a huge number of people at STA who have railway knowhow, [but] are pinioned, who are actually not allowed to do what they are capable of doing.” (5)

4.2.2.3.2 Competence and knowledge at the maintenance contractor

A natural part of procuring maintenance contracts is changing the contractor due to a more competitive bid. According to the respondents, the turnover of contractors is high. Once a contractor has gained experience of a certain railway section, they become familiar with the track and understand what price tag is relevant. This is exemplified by the following quote:

“You get to know a railway infrastructure, you get to know a track section, you know what it costs and then you win [the contract once] and then the next [time] you’re kicked out, because then you know how to set the price and then there’s always someone who [bids] lower.” (3)

However, knowledge and experience of a track section are essential in maintenance work. Knowledge about where on the track problems normally arise makes both preventive and corrective maintenance more efficient. This is exemplified by the following quote:

“In the long run you develop such things as good competence, good knowledge about your infrastructure and thereby fast repair work.” (3)

Changing the contractor maintaining a railway section has several consequences. The organization of the contractor leaving the area in question is dismantled. Typically, the personnel living in the area are hired by the new contractor. This gives the possibility of some transfer of track knowledge to the new organization that has to be established when a new contractor wins a bid. However, the interviewed contractor representatives are concerned that knowledge is lost in every change of contractor.

4.2.2.4 Specialized machinery is replaced by standard machines

According to the respondents, high railway quality is dependent on specialized machines that are typically expensive to acquire. When SRA was in charge of
the operation and maintenance nationally, the machines could be used within
the entire railway infrastructure. Today there are no requirements concerning
machines in the contracts since the specifications are performance-based. As the
contractors are smaller organizations, the utilization of the machines in each
company is not high enough to motivate the purchase of specialized machinery,
due both to the reduced length of the infrastructure to maintain and to the fact
that the contract period is not long enough to make it possible to write off the
depreciation of the machines. This is exemplified by the following quote:

“You know, the trouble here is that [...] the new machines that were bought
some time in the 2000s, they have high productivity and large depreciation
costs. This means that these machines must be used very extensively to be prof-
itable and then you get a very high work volume [from them], or maintenance
performance. [...] But [...] if you don’t use them, then they become dreadfully
expensive, of course, and then the contractor thinks, ‘But I don’t want such an
expensive machine because I can do this with somewhat lower performance and
somewhat lower quality; it will be the same, approximately the same result.’”
(4)

In addition to the problem of the length of the contract period, one has the typi-
cally long delivery time for advanced machines, which can be several years.
Certain machines would give the possibility of performing maintenance work
more cheaply, if the machines could be utilized enough. This is exemplified by
the following quote:

“And then you can understand this phenomenon, that if every contractor is to
own their machine and run it 200-300 hours a year, then it will be expensive.
But if one owns the machine and succeeds in [hiring it out], then it will be very
cheap instead.” (3)

The contractors do not take the risk of purchasing new specialized machinery in
case they might not win another contract. Instead, excavators with different
accessories are typically used. Although excavators can be used for several dif-
ferent purposes, they cannot give the same quality as specialized machines.
However, multi-purpose excavators are a cheap solution, at least in the short
term. This is exemplified by the following quote:

“Good machines give good track quality, there is no doubt about that. But what
you find now is all types of cheapest possible solutions.” (3)

While contractors are avoiding buying new machines, the machinery dating
from the time before the deregulation and still in operation is growing older and
less reliable. Machinery is also discarded because it is used to a small extent.
This is exemplified by the following quote:
“[…][All] the contractors have more machines than are required to keep the track in shape. But no work orders are coming in, of course. So eventually what happens is that you decommission the machine; in the end we have cut the number of machines we own by half and when we actually are in need of a good track standard then we can’t fix that. Because then we don’t have any gear. It takes four years to get a machine after ordering it and they cost 100 million [Swedish crowns] and now we are scrapping machines […].” (5)

According to the respondents, many operations could be performed in a shorter time by using specialized machines, meaning that large operations that are currently being performed during weekends could be carried out during night slots for maintenance. In this way, the utilization of machines and personnel would be more efficient.

4.2.2.5 Insecure working conditions for railway workers

The maintenance organization as a whole must be continuously adapted when contracts are lost and won and the work load subsequently changes. Hence, railway workers experience an insecure work situation, as illustrated by the following quote:

“Every fifth year you may risk changing employer.” (5)

If a maintenance contractor is replaced, the workers typically need to be rehired by the new contractor if they want to continue working in the same area, as illustrated by the following quote:

“If you get a job at [another contractor], it’s not a question of a business takeover, so the new company chooses who they want to employ. […] For a few people this has meant a positive wage spiral, but that’s not the case for most people. For most people it has meant that [the contractor] wants to have temporary appointments.” (5)

The respondents representing contractors expressed a concern that the safety of the railway workers in the field is questionable for the following reasons, among others: a larger amount of emergency work, with no time to arrange adequate protection and perform safety planning beforehand, an increasing number of temporarily employed and hired personnel who are not familiar with the line, difficulty in coordinating the actions performed by workers from different organizations on the track, and compromising safety to save money. This is exemplified by the following quote:

“In my opinion, this is because of the procurement system; it is expensive to have people at the sites. If you are going to have one person assigned to attend
to safety work, then that’s one more salary. ‘Since you are already at the site, you can take care of that.’” (5)

4.2.2.6 No incentives for maintenance contractors to develop maintenance methods

According to the STA representative, STA is not supposed to carry out development itself, and is only supposed to direct and finance development and innovation if there is an obvious need. As far as STA is concerned, development is, like all other things purchased, governed by the APWETPSS. Collaboration concerning development with maintenance contractors is mainly carried out within the framework of a contract. Depending on the content of a contract, development may be within the scope of the contract, or even the main target of the contract. However, the key aspects of a contract cannot be changed after it has been awarded to the winner of the procurement, since that would be unfair to the other competitors. The purchase of a developed product or process must be procured in a new tender request. It is difficult to specify development in a contract, since research and development is inherently a process of trial and error. This is exemplified by the following quote:

“We are allowed to procure development. The difficulty lies in describing the product so you get the right thing [in the end].” (1)

Other research and development frameworks like collaboration projects within the Nordic countries or the European Union offer opportunities for development outside the framework of a contract, in some cases together with contractors. For the contractors, collaborating on a development project ultimately comes down to whether or not there is something to gain with respect to winning future procurements. Otherwise STA has to pay for the development. This is exemplified by the following quote:

“I think the innovations are limited to what each separate company can find possible to develop and do, and use for [their] own competitiveness and profitability. As soon as I think of an innovation, for example you can do such and such a thing and be able to halve the time for repairing a torn-down overhead line. Yes, but then my competitors can just take this method as soon as it becomes known, because STA demands that, that everyone should do the job in half the time as soon as there’s someone doing that. So [the news] will spread; they will employ someone from us and thus find out about it. So I can never protect [an innovation]. So why should I invest that money. Better to wait until my competitor does it. Or STA pays.” (6)

According to the respondents, contractors lack an incentive for innovation in the maintenance contracts. It is expensive to innovate and development costs
must be included in the bidding price. As maintenance contractors currently make little profit, there is simply not any money to put into development. This is exemplified by the following quote:

“As long as the companies are not making any money it’s not possible to spend money on research and development.” (5)

According to the respondents, research and development does not always result in a usable process or product. For small companies, the economic risk of engaging in development projects is too great, even though such projects may have a high potential for the future. Further, if one company has developed a new and improved method, STA can just demand in procurements that all the companies should use this method and request tenders. The innovating contractor will not be rewarded and the other companies will not have to contribute to the development cost. This is exemplified by the following quote:

“It costs [the contractors] time and resources to develop that optimal welding method. How easy it is for [STA] to take the method and send it to [all the contractors’ competitors]!” (1)

Research and development performed by contractors is to a great extent being replaced by monitoring what others are developing, by reading project reports and articles in journals, and by visiting exhibitions, for example, and by trying to utilize findings made by others when possible. The contractors are taking a passive role instead of engaging in in-house research and development. This is exemplified by the following quote:

“I rather experience that the contractors [are] passive when it comes to process development. They haven’t been participating so much in product development, but they have processes that could be developed. And there I think one could have come much further.” (2)

According to the respondents, there is room for the development of maintenance methods. For example, as the traffic increases on the railway network, time slots for maintenance work are limited. Therefore, it would be desirable to develop methods that would decrease the time required for maintenance on the track. There are examples of new procedures that have significantly cut the time required to perform large operations. However, in one such case, the contractor did not follow up on their innovation after demonstrating it, arguing that STA would not value such an innovation during procurement. Hence, there was no incentive for the contractor to acquire the expensive machines that would be necessary. This is illustrated by the following quote:
“[The contractor] took that idea and adjusted one of their machines and rented a [special crane] from [another country]. And then they performed [the operation] in a very short time. It was very successful. […] Then [the contractor] said, ‘No, we will not continue. [The infrastructure manager’s] tendering procedure doesn’t value such things.’ So [the contractor] completely dropped it.” (1)

As far as the contractor is concerned, the focus is fixed on making everything as cheap as possible to win procurements by forcing down the price and receiving contracts. Therefore, the aim of the development of maintenance methods performed by contractors is to do everything as cheaply as possible, as exemplified by the following quote:

“The focus is on how to solve it in the cheapest way. Not the best way. Not preventive.” (3)

4.3 Study 2: Collaboration on the Swedish deregulated railway market

4.3.1 Method

4.3.1.1 Participants

Twenty participants from the UPPSAMT workshop, described in more detail in Chapter 7, took part in the study on collaboration on the Swedish deregulated railway market. The participants were assigned to four different groups by the UPPSAMT project leader. The aim of this assignment was to obtain a mix of competences in each group, and to ensure that the participants were from different organizations and that there was at least one expert on overhead contact wires in each group, as this would facilitate the ideation part of the workshop. The details of the respondents (as reported by the respondents themselves) and the groups are shown in Table 4.3. Eighteen participants were males and two participants were females. For integrity reasons, the author has chosen not to detail which respondents were males and which were females in Table 4.3. For the same reason, the author has excluded the area of expertise of each respondent. The age of the participants was not requested. Two of the twenty participants worked for a Norwegian train operator in Norway. The written data from these (male) participants were removed, since the focus of this study was on the situation on the Swedish railway market. The participants did, however, take part in the discussions that took place in their respective groups. The analysis of these discussions, however, has focused on the Swedish situation and has not considered any contributions concerning the Norwegian situation.
Table 4.3 Details of the respondents.

<table>
<thead>
<tr>
<th>No.</th>
<th>Organization</th>
<th>Railway experience (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>STA</td>
<td>10-15</td>
</tr>
<tr>
<td>2</td>
<td>Contractor: maintenance of rolling stock</td>
<td>10-15</td>
</tr>
<tr>
<td>3</td>
<td>STA</td>
<td>10-15</td>
</tr>
<tr>
<td>4*</td>
<td>Train operator: passengers</td>
<td>5-10</td>
</tr>
<tr>
<td>5**</td>
<td>Contractor: maintenance of infrastructure</td>
<td>5-10</td>
</tr>
<tr>
<td>6</td>
<td>STA</td>
<td>15-20</td>
</tr>
<tr>
<td>7</td>
<td>Train operator: passengers</td>
<td>&gt;25</td>
</tr>
<tr>
<td>8</td>
<td>STA</td>
<td>5-10</td>
</tr>
<tr>
<td>9</td>
<td>Contractor: maintenance of infrastructure</td>
<td>&gt;25</td>
</tr>
<tr>
<td>10</td>
<td>Other organization</td>
<td>20-25</td>
</tr>
<tr>
<td>11</td>
<td>STA</td>
<td>&gt;25</td>
</tr>
<tr>
<td>12</td>
<td>Academia</td>
<td>15-20</td>
</tr>
<tr>
<td>13</td>
<td>STA</td>
<td>0-5</td>
</tr>
<tr>
<td>14</td>
<td>Academia</td>
<td>0-5</td>
</tr>
<tr>
<td>15</td>
<td>Other organization</td>
<td>10-15</td>
</tr>
<tr>
<td>16</td>
<td>STA</td>
<td>&gt;25</td>
</tr>
<tr>
<td>17</td>
<td>Other organization</td>
<td>&gt;25</td>
</tr>
<tr>
<td>18</td>
<td>Train operator: passengers</td>
<td>5-10</td>
</tr>
<tr>
<td>19</td>
<td>Contractor: maintenance of infrastructure</td>
<td>15-20</td>
</tr>
<tr>
<td>20*</td>
<td>Train operator: passengers</td>
<td>20-25</td>
</tr>
</tbody>
</table>

*Norwegian participants whose data were removed from the written part of the study.

**This participant did not submit replies to the questions in Table 4.4, but participated in the discussions in the group.

4.3.1.2 Data collection

The data collection in this study took place during a workshop arranged within the scope of a collaboration project called UPPSAMT, which is described in more detail in Chapter 7. The overall aim of the workshop was to apply the ideation method developed within the OptiKrea project to topics concerning the continuation of the UPPSAMT project. However, time was also dedicated to working on the topic of collaboration and the implementation of results from pilot projects. More specifically, the task given to the participants was to pro-
provide answers to the following two questions (translated from Swedish into English).

“1) What follows after a demo project like UPPSAMT, and how can the results [from a pilot project] be implemented in regular activities?
2) How should the deregulated railway sector organize itself to be able to match and coordinate data? (In answering this question, one can, for example, touch upon issues such as the ownership of data and business models.)”

The participants were informed that the purpose of this task was to highlight current problems and to identify critical factors for success with respect to the two questions posed above.

The respondents were given a subset of the topic clarification questions presented in Section 8.3.2.4, which had been adapted somewhat to the task of the UPPSAMT workshop. These questions are presented in Table 4.4. The participants were instructed to answer each question individually by writing down their replies on a sheet of paper. Thereafter, the members in each group took turns to present their replies to each question and discussed the group members’ replies to each question before moving on to the next. Each group was audio-recorded. One participant in Group 1 did not hand in their written replies to the questions. Consequently, the written replies to the questions from 17 respondents were available.

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What is the problem with the current solution/situation? Why does it need to change?</td>
</tr>
<tr>
<td>2</td>
<td>What is the core of the problem and wherein lies the greatest need?</td>
</tr>
<tr>
<td>3</td>
<td>Who wants the problem to be solved and why?</td>
</tr>
<tr>
<td>4</td>
<td>What are the (root) causes of the problem?</td>
</tr>
<tr>
<td>5</td>
<td>What function should a solution have? What should it bring about?</td>
</tr>
<tr>
<td>6</td>
<td>Which possibilities are open and which possibilities are closed for the achievement of an altered solution/situation?</td>
</tr>
<tr>
<td>7</td>
<td>What alternative solutions exist today? Make a survey of how the situation has been addressed in other contexts (e.g. in other countries or other branches).</td>
</tr>
<tr>
<td>8</td>
<td>What legislation, standards and the like influence what a solution can look like?</td>
</tr>
<tr>
<td>9</td>
<td>What are the wishes and demands concerning flexibility and the possibility of changing the solution in the future?</td>
</tr>
<tr>
<td>10</td>
<td>Are there other aspects that should be taken into account?</td>
</tr>
</tbody>
</table>
4.3.1.3 Data analysis

The replies to the questions that each participant had written down on a sheet of paper were categorized into a) replies that were technical and specific to the topic at hand and b) replies that concerned the overall situation regarding collaboration as well as other issues relevant to the Swedish railway market. The replies of the second type were interesting for the scope of Stage 1 of the present research and were analysed further using a thematic analysis; the analysis procedure is described in Table 4.5.

Table 4.5 The procedure for analysing the replies to the topic clarification questions used during the UPPSAMT workshop.

1. The replies from the participants were collected in a single word processing document according to the question which they concerned.

2. The document was copied and in the copy the headings in the form of the questions which the replies concerned were removed. The replies were examined and replies concerning the same matter were listed together and tentative categories were formed.

3. A check was performed to ensure that all the replies fitted into a category, otherwise the replies which did not fit were split, if possible, or the categories were changed. When all the replies had been placed in a category, the categories were given descriptive labels and a description. For each matter described in a category, a count was made of the number of participants who had mentioned that matter.

4. In the original word processing document (see point 1), the replies under each question were sorted into the categories. A check was then performed to determine which questions had generated the replies in each category.

5. For each theme, a count was made of the number of respondents who had mentioned that theme in their replies.

The audio recordings were examined to find out what issues each group had discussed concerning the overall situation regarding collaboration on the railway market in connection with the presentation of the respondents’ replies to the topic clarification questions. The recording from each group was listened to and all the comments that concerned the overall situation regarding collaboration and other issues relevant to the Swedish deregulated railway market (as opposed to technical comments and comments specific to the topic at hand) were written down in a document. Comments concerning the same matter which were made in each group were listed together. The comments concerning each matter which the groups had discussed were then described. It was found that all the matters did not contribute specifically to the objectives of Study 2, and that some contributed to the objectives of Study 1. If any matter was found to contribute to any of the objectives of Stage 1 of the present research, it was included in the analysis.
4.3.2 Results

4.3.2.1 Individual replies to the topic clarification questions

The individual replies to the topic clarification questions were classified into six different categories and, in addition, an “Other data” category for data which did not fit into the six different categories and which concerned matters only mentioned by one respondent. The categories are presented in Table 4.6.

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ownership of the issue</td>
<td>Nobody taking responsibility for the issue</td>
</tr>
<tr>
<td>Hesitation in data sharing, different incentives and goals</td>
<td>Concerns regarding the actors being hesitant to share data with others and actors having different incentives and goals for their activities</td>
</tr>
<tr>
<td>(Maintenance) contracts</td>
<td>Concerns regarding the design and implementation of (maintenance) contracts</td>
</tr>
<tr>
<td>Individual solutions, lack of coordination and collaboration</td>
<td>Concerns regarding each actor having an individual solution and/or working without taking the complete picture into account (suboptimization), and concerns regarding a lack of coordination and collaboration between actors</td>
</tr>
<tr>
<td>Competition and APWETPSS</td>
<td>Concerns regarding the competition on the market or the APWETPSS</td>
</tr>
<tr>
<td>Resources</td>
<td>Costs for a solution or implementing an activity in terms of both money and time</td>
</tr>
<tr>
<td>Other data</td>
<td>Data that do not fit into the other categories (mentioned by a maximum of one respondent)</td>
</tr>
</tbody>
</table>

Table 4.7 shows, for each of the above-mentioned categories, how many respondents mentioned a theme belonging to the category in question in any of their replies to the questions in Table 4.4. The respondents who did not mention these themes did not mention any alternative concerns or explanations for the current overall situation, but rather focused on technical or topic-specific issues. From Table 4.7, it is found that a lack of coordination and collaboration, with each actor optimizing their own solution, is a concern mentioned by a majority of the respondents.

Four questions generated the majority of the relevant replies, namely questions 1, 2, 4 and 6 in Table 4.4. Examples of replies given by the respondents to these questions are listed in Table 4.8.
Table 4.7 For each category, the number of respondents mentioning a theme belonging to that category in their replies to the topic clarification questions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ownership of the issue</td>
<td>7</td>
</tr>
<tr>
<td>Hesitation in data sharing, different incentives and goals</td>
<td>6</td>
</tr>
<tr>
<td>(Maintenance) contracts</td>
<td>5</td>
</tr>
<tr>
<td>Individual solutions, lack of coordination and collaboration</td>
<td>12</td>
</tr>
<tr>
<td>Competition and APWETPSS</td>
<td>8</td>
</tr>
<tr>
<td>Resources</td>
<td>6</td>
</tr>
</tbody>
</table>

4.3.2.2 Discussions within the groups

4.3.2.2.1 Task question 1: Implementation of results from pilot projects

Group 2 discussed the continuation of collaborative research projects. The experience in this group from other collaborative research projects was that typically not much happens after such a project has ended. According to the discussion in the group, the industrial partners are not interested enough in implementing the results. Three reasons for this lack of interest were given in the group: a) the projects are conducted at quite a distance from the daily operations; b) implementation of the results costs money; and c) there are barriers within an organization preventing the implementation of something new, for example employees within an organization who might have personal reasons for countering an implementation that might remove the need for their knowledge and skills.

4.3.2.2.2 Task question 2: Organization of the deregulated railway sector to enable the coordination of data

Two of the groups discussed the sharing of data, or rather the lack of data sharing. A topic discussed in both these groups was the fear of actors of sharing any data in case this would give advantages to companies which they might meet in a competitive situation. According to the experience of one of the members of Group 4, if representatives of competing actors were to gather for a meeting, there would be complete silence. In Group 3, it was mentioned that one challenge with regard to data sharing is to get all the relevant actors to join in. The experience of Group 3 led them to predict that there would always be actors that would find data sharing to be a great idea, but then in a few months would say that they would not take part because it would disturb their business. In Group 4, it was mentioned that this worry about sharing data appeared to be unfounded to a great extent, because the actors concerned had usually not evaluated what data were actually sensitive and what data could be shared and provide value to all the actors involved.
<table>
<thead>
<tr>
<th><strong>Table 4.8</strong> Examples of replies (quotes) given by the respondents to the questions that generated most relevant replies with respect to the research question.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Examples of replies from the respondents</strong></td>
</tr>
<tr>
<td><strong>What is the problem with the current solution/situation? Why does it need to change?</strong></td>
</tr>
<tr>
<td>- Many actors have their own solutions, which are not coordinated and do not generate experience/learning.</td>
</tr>
<tr>
<td>- Information is collected in different sources and important information that would help/solve problems is not accessible for making decisions.</td>
</tr>
<tr>
<td>- There is an abundance of different information-owners with widely differing incentives.</td>
</tr>
<tr>
<td>- Insufficient/non-existent collaboration between the parties involved.</td>
</tr>
<tr>
<td>- The problem of today’s railway is NOT the fragmentation of the system in itself. It is rather that no party has the task of keeping the system together in the prevailing conditions. STA must be the party that leads the work and creates incentives/sanctions [to encourage/force] the other actors to participate.</td>
</tr>
<tr>
<td><strong>What is the core of the problem and wherein lies the greatest need?</strong></td>
</tr>
<tr>
<td>- A lack of concrete practical collaboration within the branch.</td>
</tr>
<tr>
<td>- The parties have widely differing company goals, are competitors or have a customer-supplier relationship.</td>
</tr>
<tr>
<td>- The need for analysis is today not coupled to the primary interests of the information owners.</td>
</tr>
<tr>
<td>- Short contracts.</td>
</tr>
<tr>
<td>- No one wants to own the problem, everybody thinks [a common solution] is a good idea.</td>
</tr>
<tr>
<td><strong>What are the (root) causes of the problem?</strong></td>
</tr>
<tr>
<td>- Increasing efficiency within “watertight compartments” (deregulation without consideration of the need for coordination).</td>
</tr>
<tr>
<td>- Absence of any common goal (providing guidance in each business) for the functioning of the whole.</td>
</tr>
<tr>
<td>- Competition – fear of sharing something which can favour someone else.</td>
</tr>
<tr>
<td>- Difficult to find time/money to invest (one has perhaps not worked out the advantages).</td>
</tr>
<tr>
<td>- Short-term economic perspectives/conditions of competition.</td>
</tr>
<tr>
<td><strong>Which possibilities are open and which possibilities are closed for the achievement of an altered solution/situation?</strong></td>
</tr>
<tr>
<td>- Disjointed and conservative mode of working.</td>
</tr>
<tr>
<td>- An independent party needs to take responsibility for the coordination. There is an opening [here] in that all the parties seem to realize that openness is required to improve [the situation].</td>
</tr>
<tr>
<td>- Closed possibilities due to the competition [between stakeholders].</td>
</tr>
<tr>
<td>- Uncertainty regarding data (integrity issues) shuts down [possibilities].</td>
</tr>
<tr>
<td>- Positioning regarding costs is applying the brake.</td>
</tr>
</tbody>
</table>
Group 2 brought up the fact that actors focus on the competition within the market and forget that the market as a whole has competitors, for example in the form of road and air transportation. Group 2 discussed the fact that the Swedish deregulated railway market is an immature and introvert sector that is focused on itself and believes that the railway is very special. The group was of the opinion that other markets had similar challenges, such as many actors, short contracts and many rules and regulations, and that the railway sector should study how other businesses have solved similar issues to those experienced in the railway sector. Group 2 also brought up the fact that the contractors are unaccustomed to daring to invest and taking risks, and instead expect other actors to take the initiative or pay the cost.

4.3.2.2.3 Maintenance

Group 1 discussed the fact that the contracts on offer today do not give the maintenance contractors any incentives for proactive maintenance and development work. Further, the maintenance contractors do not have any resources for development work since they do not make any profit, despite the fact that the goal of a company is to make a profit. The contract time span of five years (with an option of a two-year extension), during which the maintenance contractors try to optimize their activities to make a profit, was found by Group 1 to be counterproductive and to prevent a long-term approach to infrastructure maintenance, since that time span is much shorter than the life span of the infrastructure. Group 2 brought up the fact that the actors making use of the loopholes are the winners and that those which adopt a long-term perspective are the losers. Examples of alternative types of contracts (with alternative conditions) in use in Sweden were mentioned in Group 1, and these might create incentives for working proactively. Group 1 found that the option of a two-year extension is usually exploited since STA is usually late in initiating a new procurement.

In Group 4, it was mentioned that it is mostly faults associated with a risk of derailment which are remedied, and not faults affecting the comfort of the passengers. It was also mentioned that some failures are reported several times without any measure being taken, and as a result people stop bothering about such failures and stop reporting them.

4.3.2.2.4 Being in control of the infrastructure

Three of the groups discussed issues connected with STA not being in full control of the infrastructure. Group 1 discussed the fact that STA does not have control of the condition of the infrastructure and that inspections of the line when a change of contractor has been taking place have not been good enough. There are examples of procurements where STA has stated that the condition of the line was on a certain level, and where the winning bidder has examined the actual condition of the line in the field and has found that the line was in a
worse condition (leading to the contractor demanding extra money to restore the line to the stated condition). Group 1 brought up the fact that contractors are supposed to hand back the line section that the expiring contract concerns in the same condition as it was in when given to them, but there is no good way to measure this condition. In Group 3, worries were raised about the fact that STA does not own the measurement wagons that are used to inspect the infrastructure, while Group 4 brought up the fact that inspectors typically find the same failure, but make different judgments about how acute the failure is (e.g. for the same failure, sometimes it is deemed necessary to perform a certain maintenance action within a week and sometimes within a month).

4.3.2.2.5 Competence and knowledge

Group 4 mentioned that there are many specialists at STA, but there might be a lack of generalists with knowledge about the whole transport system. The group also discussed the fact that the specialists seem to have problems explaining to the management why measures are required. Further, they brought up the fact that, to be able to develop the right things, the development department must know the target context extremely well.

Two groups brought up the fact that politicians and officials do not seem to understand the prerequisites of the railway. Group 3 brought up the fact that STA receives additional funding very late and is unable to use it efficiently since the train schedules are then already in place and the contractors are already booked. For example, such additional funding was given in the spring two years ago, and then a large amount of that money had to be used to buy back track time from the operators, since the money had to be spent before the end of the year. Group 4 brought up examples of politicians and officials who in a regional context have taken decisions about providing train services with rolling stock that does not have the reserve capacity required to reach the regularity that they demand.

4.3.2.2.6 Ambitions of STA

Three of the groups brought up intentions on the part of STA to make changes. Group 2 mentioned that there is a desire in the management of STA to find solutions that create positive incentives, and Group 3 mentioned that STA’s views on the value of information have changed and that they have demonstrated a will to adopt a more proactive maintenance regime. Group 4 mentioned that STA has an ambition to improve and to remedy problems and Group 3 mentioned that STA is close to adopting an LCC philosophy with regard to the replacement of parts of the infrastructure.
4.3.2.2.7 Barriers preventing change at STA

Three groups expressed the criticism that there are barriers preventing change at STA. Group 3 and 4 expressed the criticism that STA is a slow and conservative organization where it takes time for changes to happen. Group 4 mentioned that many good ideas and suggestions on working methods emerged at STA, but that the organization was slow in implementing them. Group 3 brought up the fact that there have been many employees at STA who have become tired of the railway and changed their employer because they have hatched ideas which have been ahead of their time. Changes are now being introduced at STA which were suggested several years ago by these said employees.

4.4 Study 3: Conditions for product development

4.4.1 Method

4.4.1.1 Participants

The respondents of this study were the professional participants of the OptiKrea Group, i.e. participants A-D in Table 3.1.

4.4.1.2 Data collection

The data for this part of the study were continuously collected during different activities taking place as part of the OptiKrea project, i.e. during Stage 1-5 in Table 1.1. All the recorded individual interviews, group interviews and workshops were included in the data used. In addition, the following documents, which were mentioned by the participants, were used for the data collection:

- **TDOK 2014:0306 Utveckla och utforma system- och teknikkrav** [Develop and formulate system and technical requirements] (Trafikverket, 2014c),
- **TDOK 2014:0307 Införande av ny eller modifierad komponent i anläggningen** [Introduction of a new or modified component in the infrastructure] (Trafikverket, 2014b),
- **Samhällsekonomiska principer och kalkylvärden för transportsektorn: ASEK 5.2** [Socio-economic principles and calculation values for the transport sector] (Version 2015-04-01) (Trafikverket, 2015):
  - Kapitel 2: Grundläggande kalkylteknik [Chapter 2: Basic calculation technique]
  - Kapitel 3: Kalkylprinciper och generella kalkylvärden [Chapter 3: Calculation principles and general calculation values].
4.4.1.3 Data analysis

The transcribed recordings from the interviews and workshops were examined and all the excerpts that related to the procedure for product development when STA is involved were marked. These comments were collected in a single word processing document. The comments were arranged in a logical order (where comments concerning the same matter were grouped together) and then summarized in a descriptive text. The documents were read through and the parts relevant for the study were used to triangulate the data that had been collected from the professional participants in the OptiKrea Group. The descriptive text was adjusted accordingly.

4.4.2 Results

It is not within the scope of STA’s mission that they should develop railway products themselves and STA has no in-house production of railway products. The role of STA is to specify the scope of the product required, whereas the role of companies in the railway sector is to develop and provide the actual product. At the national level, the infrastructure manager is the sole purchaser of railway assets. Railway products are purchased through procurement and, as a public utility, STA must follow the APWETPSS (Lag om Upphandling inom Områdena Vatten, Energi, Transporter och Postjänster, 2007). The delivery of a required product is procured for a time span of several years through a framework agreement. To procure a product, STA compiles a tender request with requirements for the product. Interested suppliers reply by sending a tender to STA and the tenders are evaluated according to the procedure specified in the tender request to find the winner of the procurement. According to the respondents, STA aims to use functional requirements in their tender requests and the tender award criteria and evaluation outcome should be expressed in a monetary value. Naturally, the development of railway products must be performed in accordance with the applicable legislation.

Often the procured product, as specified by the requirements given by STA in the tender request, does not exist in reality during the procurement. Hence, the information in the tender offers is often marred by uncertainty, which, according to the respondents, entails a risk that the product when eventually delivered will not meet the requirements as specified in the tender request. According to the respondents, the only figure with no uncertainty in the tender request is typically the purchase price, and this is one of the reasons why the purchase price has been given most weight during procurement. Once a winner of the procurement has been found, STA can work together with this supplier in developing a product that meets the requirements stated in the tender request. According to the respondents, it is difficult to formalize the trustworthiness of tenderers in a fair way that can be used during the procurement to sort out unrealistic offers and determine which company has the best necessary qualities to
succeed in delivering a product that fulfils the requirements, while simultane-
ously taking the price into account. It is also the experience of the respondents
that the suppliers, on the other hand, take a risk in developing a product that
might be sold to only one buyer and, in their eagerness to win the procurement,
in offering a price which is too low and will end up not covering the actual de-
velopment costs.

There are situations when the APWETPSS does not apply, for example in
pre-commercial procurement (where development work is conducted under
competition) and in research and development projects. The purchase of the
delivery of a product from such a project must be procured, but the learning
from such projects can, according to the respondents, be transferred to regular
procurements and act as input to the requirements in the tender request. The
requirements and stated evaluation process cannot be changed after the tender
request has been distributed. According to the respondents, a change will result
in major consequences such as a repeated procurement. While socio-economic
calculations are taken into account during the analysis of large infrastructure
investments at STA, for example, they are not included in the procurement of
railway products through a framework agreement. STA has rules and regula-
tions regarding how socio-economic calculations are to be performed.

When STA intends to introduce a new or modified product, the standards
TDOK 2014:0306 (Trafikverket, 2014c) and TDOK 2014:0307 (Trafikverket,
2014b) are followed. The processes and routines described in these documents
follow the European Standard EN 50126 (CENELEC, 1999), which is also the
Swedish standard. EN 50126 (CENELEC, 1999) is a framework that describes
the process to be followed to assure that all the aspects of an asset from “the
cradle to the grave” are considered. According to the respondents, different
methods can be used to provide the information required by EN 50126
(CENELEC, 1999).

According to the respondents, there is always an alternative when STA eval-
uates a situation. The alternative does not have to be a similar product, and can
also be, for example, the option of driving a car instead of taking the train.

4.5 Discussion

4.5.1 Forcing the tender bids down to win contracts

There exists a common understanding that replacing a public monopoly or a
publicly guaranteed private monopoly with competition between companies
results in improved efficiency (Alexandersson and Hultén, 2008). However, a
certain level of competence is required to design tender requests and evaluate
bids to avoid the negative consequences of obtaining unrealistic or strategic bids
(Alexandersson and Rigas, 2013); otherwise the possible gains may be lost. Study 1 has found that maintenance contractors make unrealistically low bids to
win contracts and then chase additional work which must cover the loss generated from the basic part of the contracts. In a review of STA’s procurement of roads and railroads in general, the Swedish National Audit Office (Riksrevisionen, 2012) pointed out that contractors often have the possibility of compensating for a low bid with additional reimbursements. Some mechanism for time and material reimbursement is difficult to avoid in contracts, since there will always be conditions which the contractor and procurer cannot foresee.

The survival of maintenance companies is uncertain due to repeated tendering, and a high turnover may be unavoidable when the market is learning and adapting to new conditions. One possible reason for a company placing a low bid is that the company actually is operating in a cost-efficient manner. However, Study 1 points to other explanations in the case of maintenance contracting. It is difficult to estimate the actual cost of uncertain actions taking place in the future, and the true cost becomes obvious to the contractor once they have been maintaining the line for a period of time. Nash and Smith (2006) noted that it was expected in the British market that the incumbent contractor would start with an advantage in terms of knowledge of costs and markets. However, they found instead that thus far only a small fraction of the franchises obtained through refranchising had been won by the main incumbent. Yet many incumbents went on to win other franchises in different parts of the country. This rather suggests that overoptimistic bids might be, at least partly, the result of poor information, leading to the “winner’s curse” (Nash and Smith, 2006), if not merely the result of a strategic decision taken by the contractor. On the other hand, Lingegård (2012) found, when interviewing contractors and STA staff, that a contractor in charge of maintaining a track section was likely to win the next procurement, since their organization was already established in the area and could therefore offer a lower bid than their competitors. The respondents in Study 1 and 2 have expressed mixed views as to whether there is typically a change of maintenance contractor on line sections after a new procurement, and whether two-year extension options are typically exploited. As the procurement of maintenance started in 2001 and was subsequently introduced at a slow pace, not that many reprocurements have actually taken place. Further, the turnover of contractors may be different depending on the geographical area and respondents may therefore have different impressions.

4.5.2 Optimization of one’s own company’s profit

The Swedish maintenance contractors are companies, and their goal is to make a profit from their business. The current design of the maintenance contracts gives the contractors an incentive to work in an unsustainable manner to maximize their short-term profit, with a short-term operative maintenance perspective. As STA has an explicit policy of proactive maintenance (Stenström, 2012), it is alarming that the results suggest that the contractors have an incentive for corrective maintenance. Yvrande-Billon and Ménard (2005) suggest that merely
changing the ownership does not provide the incentives and outcomes typically desired when performing deregulation. In reforming public utilities, regulators must take into account how the conditions which they impose influence choices (Yvrande-Billon and Ménard, 2005).

The quality of maintenance performance seems also to correlate to the fact that STA cannot take earlier contractor performance into account during tender evaluation. This reduces the contractors’ incentive to aim for a good reputation through actions that might increase their costs or decrease the amount of additional work that STA will pay for. Companies with higher standards will either be outperformed by bolder companies with lower standards or have to adopt a similar behaviour to survive on the market. According to Stenbeck (2007), so-called soft parameters (i.e. other parameters than price) were initially given a 30% weight during the bid selection at the Swedish Road Administration, which was the former counterpart of the SRA in the area of road transport. Due to complaints from the contractors that the bid selection had become too arbitrary and subjective, the weight was gradually reduced to 3.5% in 2004. The former SRA assigned a 70% weight to price in their bid selection in 2004 and experienced similar complaints to those received by the Swedish Road Administration. The soft parameters in both agencies were, however, based on documentation submitted with the offer and not actual performance in the field (Stenbeck, 2007). It was reported by respondents in Study 1 and 2 that alternative types of maintenance contracts existed which were expected to give incentives to the contractor to work proactively. In the literature there are examples of successful bonus and fee systems leading to a decreased number of faults and improved quality, see Stenbeck (2008).

It has been found from Study 1 that one difficulty in procuring performance-based maintenance contracts is the lack of knowledge of what will happen in the future and, consequently, what maintenance measures will be required to fulfil the contract. Stenbeck (2007) described three success factors for performance-based specifications: how well the infrastructure owner can describe and define the contracts, how compliance is measured and how deviations are handled, i.e. how the contractor is penalized for non-fulfilment or rewarded for excess delivery. According to Study 1, the maintenance contracts that have been awarded by STA have not fulfilled these success criteria and this has led to legal cases. As noted by Nash (2008), in a contract involving parties whose knowledge of future circumstances cannot be perfect, situations are likely to arise (e.g. increased traffic) which require details to be renegotiated, entailing the risk of opportunistic behaviour on the part of either party. Stenbeck (2007) found that technical acceptance limits were not written explicitly into contracts. The contract owners hinted, during interviews, that technical acceptance limits were part of the contract code, but were embedded in the code and only understandable by experts. This definitely creates a risk of diverging interpretations, opportunistic behaviour from the contractor and legal cases. Legal cases that could have been...
avoided by a better formulated contract are a waste of both the contractor’s and the tax payer’s money.

The findings suggest that the follow-up assessment of the contractors’ maintenance work is unsatisfactory, entailing the risk of unfair competition. The Swedish National Audit Office (Riksrevisionen, 2012) noted in its review of the procurement of roads and railways that the contractors mentioned experiencing insufficient monitoring of the fulfilment of contract conditions. This issue was also raised by STA’s internal audit. According to Stenbeck (2007), inspections are even more important in performance-based contracts than in ordinary contracts, because, without inspections, the contractors have an incentive not to meet demands and thereby save money in both the short and the long term. It seems to be a rather naïve approach to allow Swedish contractors to operate without regular inspections performed by either the infrastructure owner or an independent third party. The contractors perform the asset inspections themselves, giving them the possibility of adjusting the results to suit their current work load or to fulfil the contract requirements. At the end of a maintenance contract, an independent inspection of the line’s condition is supposed to take place, but the results from Study 1 and 2 indicate that these inspections are not performed well enough, or are not carried out at all, and do not provide STA with enough data to give them a correct picture of the condition of the infrastructure. Famurewa et al. (2013) suggest a framework for the successful implementation of performance-based railway infrastructure maintenance which includes a performance monitoring tool that allows assessment of the deviations of maintenance performance from targets and consequently quick interventions.

4.5.3 Competence and knowledge

The results from Study 1 suggest that there is a risk that STA will have difficulties formulating a successful contract if they lose technical knowledge. An issue that is perhaps even more crucial is that the information asymmetry between the infrastructure manager and the contractor can be expected to increase if the technical knowledge of the STA staff, especially the inspectors, decreases. This will give the contractor the possibility of concealing technical errors and omitting to fulfil contract requirements. Further, technical competence is crucial for making the right decisions regarding the railway infrastructure. Therefore, there are several incentives to define and maintain the technical competence at STA that is required for efficient management of the infrastructure, which includes awarding contracts for maintenance and following them up. On the other hand, the results from Study 1 and 2 also point to the fact that the problem might not be a lack of specialists at STA, but rather a lack of generalists with knowledge about the whole transport system and that specialists might not have enough knowledge of the field to develop good solutions. A third possibility that has emerged from Study 1 and 2 is that the staff at STA with technical competence experience difficulties in gaining a hearing for their opinions before decisions
are being made. It is possible that all three phenomena exist in parallel, and possibly also to different extents in different units at STA.

The risk of the contractors losing competence was also brought up in Study 1. In this case, it was the line-specific knowledge which was a concern. Line-specific knowledge involves knowing where on the line problems typically emerge, for example under certain circumstances, and therefore being able either to prevent them or to attend to them quickly. As such competence stems from experience of the line, it is difficult to uphold if the railway workers are replaced, for example because another contractor has won a procurement or because of a high level of temporarily hired staff. There is also a possibility that experienced railway workers are less motivated to use their competence effectively, due to their working situation. Not only the maintenance companies, but also their employees experience an insecure situation because of repeated tendering. When a new company wins a contract, it typically rehires the staff of the previous company, creating a situation where the new employer can dictate the conditions. There is a risk that railway workers will leave the business for other opportunities, impairing the organization’s competence and local track knowledge.

Study 2 found that the competence on the level of politicians and decision-makers regarding funding might be lacking when it comes to understanding the prerequisites of the railway well enough to make the right decisions regarding funding. Simply adding more funding to STA might not be a very efficient way of achieving a more reliable infrastructure if STA does not receive them in time to make long-term plans concerning their use.

4.5.4 Maintenance machinery

In Sweden, the length and volume of the maintenance contracts have led to contractors using modified excavators and aging machines dating from the time before the deregulation, instead of specialized modern machines that could enhance the quality of the infrastructure. The contractors are not confident enough about winning another contract to dare to purchase high-cost machines. Similarly to the present findings, Stenbeck (2007) found that cost reductions had primarily been achieved through staff cuts and using less expensive and less advanced standardized machinery, rather than process and product innovations, as a result of maintenance contracting for the Swedish road and railway infrastructure. Further, STA does not own the measurement wagons that measure the condition of the infrastructure, and this can impair their ability to control the infrastructure.
4.5.5 Maintenance method development

The maintenance method development and innovation in Sweden can be divided into development that is directed by STA and development that is carried out within the contractor organizations at their own initiative.

It has been found that maintenance contractors have no incentive for maintenance method development at their own initiative. Firstly, the maintenance contractors have little money to invest in research and development. Secondly, they do not gain any benefits from developing new methods. This constitutes an extra cost and other contractors can copy the developed methods. There is one exception, and that is the development of procedures that can help them save money. There is nothing wrong in trying to cut costs. However, if this is done at the expense of quality and possibly safety, as suggested by the findings in Study 1, it is quite alarming. This approach is cheaper in the short term, but might be very expensive in the long term if the railway network deteriorates unnecessarily fast and the life cycle costs increase. Additional drawbacks are the societal costs caused by traffic disturbances, incidents and, in the worst case scenario, fatal accidents. However, these costs are not paid by the contractor with company funds and consequently they do not take them into account when trying to optimize their own profits.

As the conditions for the railway are changing, with higher volumes of traffic and greater demands for high punctuality, maintenance method development is necessary to address the challenges. For example, the increasing traffic volume means that the contractors will have less time on the track to perform maintenance and therefore procedures are needed which will involve preparation outside of the track and quick actions on the track, while achieving a high level of performance.

4.5.6 Collaboration among actors and data sharing

The lack of collaboration and coordination between the individual actors while they optimize their own solutions was a concern expressed by the majority of the respondents in Study 2. This appears to be closely related to the organizations’ different incentives and goals and the competitive situation on the railway market. The railway actors appear to see the advantages of collaboration, at least for the railway market as a whole. The fact that the respondents of Study 2 decided to dedicate an entire day to collaborating and discussing collaboration is a further sign that there is an interest in collaboration (although actors with no such interest would probably not have shown up). However, railway actors do not dare to share information with each other because they are afraid that their competitors would gain an advantage through shared information. It appears that two issues need to be addressed to enhance collaboration: 1) someone should take responsibility for providing the platform required to share data/information, and 2) a business model should be developed which gives all the
actors involved an incentive to work towards a common goal. The ePilot119 project (Junntti et al., 2014) is an example of a collaborative project involving data sharing between actors by means of a neutral platform. Jägare et al. (2015) investigated factors that had an impact on the collaboration and the implementation of results in the project. They found several legal factors (e.g. intellectual property and data ownership rights), organizational factors (e.g. commitment and communication) and financing factors (e.g. funding and return-on-investment strategies) that influenced the collaboration and needed to be resolved for a successful collaboration (Jägare et al., 2015). It is obvious that developing a business model that would encompass all the Swedish railway actors is quite a delicate process. Further, by viewing the Swedish railway market as a business ecosystem, Ingwald and Kans (2015) show that the Swedish railway market is a very complex system in which STA has a very dominating role and is directly or indirectly responsible for the actors taking part in different activities connected to the infrastructure. Therefore, at present it is difficult to see any other actor than STA taking the initiative to create a sector-wide platform for information sharing and a business model directing the actors towards a common goal, irrespective of whether it is administrated by a third party or not.

4.5.7 Product development

During product development there is in general a risk that unsuitable requirements will be specified at the start of the process, and consequently the most suitable concept will not be chosen. In the typical conceptual design process, iterations between different stages of the process are necessary and adjustments to the requirement list are made as new circumstances are discovered along the way (Pahl et al., 2007; Ulrich and Eppinger, 2014). However, in the case of a public procurement, according to APWETPSS (Lag om Upphandling inom Områdena Vatten, Energi, Transporter och Posttjänster, 2007), one is not allowed to change the requirement specification once a tender request has been issued. If such a change is made, this might lead to major consequences that are costly in terms of both time and money. This means that STA must treat the formulation of requirements as a crucial step that requires special attention during a procurement situation. Further, the fact that many products, as specified by the requirements stipulated in the tender request, do not exist in reality at the time of procurement means that STA must rely on the figures provided by the tenderers during the evaluation, although these figures are unavoidably marred by uncertainty (the extent of this uncertainty will depend on several factors, e.g. how different the product specified by the tender request is from existing products). It is difficult to judge and take into account the trustworthiness of the tenderers. This means that there is a risk that the selected tender will not be able to deliver the stated product at the promised time and to the stated cost. Despite this fact, the current tender process appears to presuppose that nothing can go
wrong during the product development process taking place at the contract winner. There is no redundancy in the process if something goes wrong.

As STA has outsourced the production and maintenance of railway assets, there is a concern that the knowledge required to write the requirement specification has to some extent been transferred to the suppliers and contractors, and this might to some extent be compensated for by involving suppliers and contractors during the topic clarification and requirement specification stages. STA aims to write functional requirements, but the fact that railway products must fit into the existing infrastructure and existing maintenance contracts means that there are several interfaces which the product must be compatible with and that it will be difficult to write merely functional requirements. Further, STA needs to specify requirements that assure that the product will comply with the applicable legislation and has a responsibility to provide a public utility while being economical with the tax payers’ money. The conditions specified in the tender request will direct the development of the product and the eventual design of the product has an impact on the direct costs (e.g. investment and maintenance costs), which are financed by taxes, and the indirect societal costs (e.g. costs originating from a failure to provide public transport for citizens). Despite this fact, the major weight in procurements of railway assets has been given to the purchase price, and socio-economic calculations have not been considered. However, it is not difficult to understand why this situation has arisen, since the purchase price has been the most certain figure in the evaluation of the tenders. On the other hand, the final outcome of the procurement in terms of the delivered product is much more difficult to judge from the purchase price and calls for a more comprehensive procedure that also takes uncertainty and risk into account.

In conclusion, drawing up a requirement list and composing a tender request constitute a delicate process during which many aspects need to be considered. The situation is complicated by the fact that uncertainties and risks, which are an inherent part of product development processes, are not addressed in the procurement process that has been used so far.

4.6 Conclusions

The current procedure for procuring maintenance contracts has in Study 1 been found to have the following impacts:

- tender bids forced down to an unrealistically low price level by the maintenance contractors,
- the optimization of one’s own company’s profit on the basis of the given conditions,
- the risk of losing or not utilizing competence and knowledge,
- specialized machinery being replaced by standard machines,
- insecure working conditions for railway workers,
the absence of incentives for maintenance contractors to develop maintenance methods, with the exception of cheap methods devised at the expense of quality.

Study 2 has shown that there is a lack of collaboration between the different actors on the Swedish railway market and each actor suboptimizes their specific sphere of interest. No one has taken responsibility to provide a platform for sharing data/information, and the actors are hesitant to share data/information since they are afraid that shared information might be used to their own disadvantage.

In Study 3, it has been found that, because STA must follow the APWETPSS and because their mission is to provide a public utility while efficiently spending public funds, the prerequisites for product development that apply for STA are different from those which typically apply for a profit-driven company. Further, in the case of procurement, the requirement specification and stated procedure for evaluating concepts in the form of tenders cannot be changed after the tender request has been issued.
Appendix 4.1: Interview questions

STA

*Purchase process*

- What is the procedure today when a turnout is to be purchased?
- How has the process changed since the market was deregulated?
  - What are the advantages/disadvantages compared to the previous situation?
  - How has the role of STA changed?
- What is the procedure for procurement?
  - Who draws up the specifications?
    - What are they based upon?
  - Time perspective?
  - Who takes the decision?
  - Are there any problems regarding the time length of the contracts (considering that the railway is a long-term product)?
- What are the consequences of today’s process for
  - maintenance?
  - technical development?
  - traffic punctuality?
- How does STA work with
  - [turnout supplier]?
  - [contractor]?
  - other actors?
    - In a formal or an informal way?
    - What is their relationship like?
- How does STA behave regarding directives “from above”?
- In what way would STA like to buy turnouts if it was possible to choose?
  - Are there different opinions in this regard between different units/professions at STA?

*Technical development*

- How does STA obtain information about new turnout concepts?
- How can STA exert an influence on how [turnout manufacturers] design future turnouts?
How does it work today?
How does STA want it to work?

- Who decides whether new concepts should be purchased by STA?
- How should new technology be demonstrated/tested before STA can start using it?
  o Is STA doing its own tests?
- Are there examples of new technology being installed too early and causing problems?
- What are the necessary prerequisites for being able to direct product development?
- What are the greatest difficulties in developing the technology required?

Maintenance

- How is maintenance purchased?
  o What is the procedure for the procurement?
  o What do the specifications look like?
- How is maintenance planned?
- What are the greatest difficulties regarding the maintenance of turnouts?
- How does one determine whether or not a turnout should be replaced?
  o Who makes that decision?
- Do [maintenance contractors] make comments regarding how easy/difficult turnouts are to maintain?
  o If so, how is that information used?
- What advantages (utility) must be obtained before one can accept an increase in the maintenance costs?
- Do [maintenance contractors] earn more money on bad turnouts that must be maintained more?
- How is the work performed by [maintenance contractors] inspected?

Other questions

- Is there anything else which would be important for us to know?
- What is the single most important measure that one should take to avoid turnouts being the cause of train delays?
Maintenance contractor

- Is [the present contractor] responsible for all the maintenance on the Swedish railway?
- What is included in maintenance?
- What is the procedure during a procurement?
  - What does [the present contractor] do?
- What do the specifications from STA look like?
  - Clear?
  - Function?
  - Do they specify what maintenance should be performed?
- How does [the present contractor] regard the uncertainty associated with the short contract lengths?
  - What does one do to make sure that one wins the next contract too?
- How is the maintenance planned?
  - Who plans the maintenance?
  - Long-term/short-term?
  - How long are the line sections that have to be maintained?
- How much is planned maintenance and how much is corrective maintenance?
  - Would it be possible to avoid corrective maintenance?
    - If yes, how?
- How are new turnouts installed?
- What is maintained on a turnout?
- What is the greatest difficulty concerning the maintenance of turnouts on the Swedish railway?
- What happens if [the present contractor] has an idea for a technical development?
- What happens if [the present contractor] comes up with a better way of performing the maintenance?
- Does [the present contractor] earn more money if there are bad turnouts (that must be maintained more) in the infrastructure?
- How does STA inspect the work that [the present contractor] performs?
- Are there parts of the maintenance [work] that are dangerous to health/involve great risks?
  - How can that be avoided?
  - Does that affect the maintenance?
  - Does that affect train delays?
- What does the relationship with STA look like?
- What does the relationship with [suppliers of turnouts] look like?
- Is there anything else which would be important for us to know?
- What is the single most important measure which one should take to avoid turnouts being the cause of train delays?

**Turnout supplier**

**Purchase/procurement**

- What is the procedure today when a turnout is to be purchased (STA)?
  - What do the specifications look like?
- How has the process changed since the market was deregulated?
  - What are the advantages/disadvantages compared to the previous situation?
  - How has the role of [the present turnout supplier] changed?
- What different procedures do other purchasers have? For example Jernbaneverket.
  - What do the specifications look like?
- What possibilities does one have of influencing the purchaser’s decision regarding, for example, different parts and details, as well as new technology?
- In what way would [the present turnout supplier] prefer the purchases to be made, if it were possible to change the procurement procedure?
- What difficulties exist regarding the purchases today?
- How is [the present turnout supplier] affected by the time lengths of the contracts?
  - How does one act to assure that one wins the next procurement?
- What is the relationship with STA like?
  - Formal/informal?
  - How does one obtain information?

**Technology development**

- How does technology development take place at [the present turnout supplier]?
- Business intelligence?
- To what extent is the technology development based on feedback from customers?
  - Does [the present turnout supplier] receive feedback from STA regarding what technology developments are desired?
- How should new technology be demonstrated/tested before it is sold, for example to STA?
  - How are the tests performed?
  - Are there examples of new technology being installed too early and causing problems?
- What are the necessary prerequisites for being able to direct product development?
- What are the greatest difficulties in developing the technology required?

**Maintenance**

- What is the relationship with [maintenance contractors] like?
- Do [maintenance contractors] share their views on how easy/difficult turnouts are to maintain?
  - If yes, how is that information used?
- Is [the present turnout supplier] affected by the way in which the procurement of the maintenance and installation of turnouts is performed?
  - If yes, how?

**Other questions**

- How large a part of [the present turnout supplier’s] market is in Sweden/the Nordic countries?
- Is there anything else which would be important for us to know?
- What is the single most important measure which one should take to avoid turnouts being the cause of train delays?
5. Test of established ideation methods

5.1 Introduction

This chapter presents Stage 2 of the research performed for this thesis (see Table 1.1). The main objective of the work carried out during this stage was to apply and evaluate three established ideation methods, which were chosen from a larger set of methods by the members of the OptiKrea Group. By combining an account of the ideation outcomes with a presentation of the participants’ views on the methods and an analysis of their behaviour during ideation, the study performed in Stage 2 aimed to provide a rich picture of how the methods worked in the given context. A secondary objective was to learn about other effects that arose due to the composition of the group.

To meet the objectives, the following research questions were posed.
- What did the participants think about working in the group?
- What did the participants think about the methods?
- Does the outcome of an ideation session depend on what ideation method is applied? If this is so, in what way does the outcome depend on the method?
- How did the participants behave during the application of the methods?
- How did the diversity of the group manifest itself?

5.2 Method

5.2.1 Procedure

The group met four times over a four-month period. Between the meetings, the participants worked with their normal tasks and did not use the ideation methods. At the first meeting, the group was introduced to ideation methods and had 16 established ideation methods presented to them, see Table 5.1. After the presentation, the participants were asked which ideation methods they would like to try during the course of the OptiKrea project. In an informal discussion among the group participants and the researchers (note that one of the OptiKrea Group participants was also a member of the research team) the group chose the methods to be tested. If a participant suggested a method and explained why he thought it would be interesting to test, and if no other participant disagreed, it
was put on the list of methods for testing. One change of method (from the quick-Delphi method to Method 635) was made after the meeting, after the project group had had some time to reflect upon the choice. The methods to be tested were Method 635 (Rohrbach, 1969; Pahl et al., 2007), the gallery method (Pahl et al., 2007), and the SIL method (VanGundy, 1984). The basic criterion for the choice was that the method should be rather simple to perform. The gallery method was chosen because the group believed that it would give well-devised concepts, the SIL method because it was believed that it could help the group to overcome mental barriers, and Method 635 because it was believed to be capable of providing many ideas and because it was different from the others. It should be noted that the participants wanted to be able to use both sketches and words when describing their ideas, and therefore the version of Method 635 used in this study is similar to the modified version found by Linsey et al. (2011) to be the most effective in terms of the number of ideas generated. The reason for involving the participants in the choice of the ideation methods to be evaluated was that their choice would indicate what type of method their organizations would be prepared to use in the future, which would increase the probability of such methods being implemented during future product development in collaboration between different actors. Another reason was to get the participants engaged and motivated to remain involved in the study, which naturally took time from their normal work load.

Table 5.1 Ideation methods presented to the participants and from which they chose the methods to be tested.

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brainstorming</td>
<td>Osborn (1957)</td>
</tr>
<tr>
<td>Electronic brainstorming</td>
<td>Gallupe (1991)</td>
</tr>
<tr>
<td>Individual brainstorming</td>
<td>Finke et al. (1992)</td>
</tr>
<tr>
<td>Method 635</td>
<td>Rohrbach (1969); Pahl et al. (2007)</td>
</tr>
<tr>
<td>Brainwriting pool</td>
<td>Geschka et al. (1984)</td>
</tr>
<tr>
<td>ICR grid</td>
<td>Wodehouse (2012)</td>
</tr>
<tr>
<td>Gallery method</td>
<td>Pahl et al. (2007)</td>
</tr>
<tr>
<td>Collective notebook</td>
<td>VanGundy (1984)</td>
</tr>
<tr>
<td>SIL method</td>
<td>VanGundy (1984)</td>
</tr>
<tr>
<td>Lead user method</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Delphi method</td>
<td>Dalkey and Helmer (1963)</td>
</tr>
<tr>
<td>Quick-Delphi method</td>
<td>Developed by Prof. Jan Lundberg, LTU</td>
</tr>
<tr>
<td>TRIZ</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Synectics</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Bionics</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
</tbody>
</table>
During each ideation session, one person was attending via video conferencing due to travel restrictions imposed by his organization. It is a typical desire from employers and employees that telework should be possible. Due to delays caused by technical issues concerning this person’s participation via video conferencing, all the methods could not be carried out exactly as planned. This had an impact on the ideation times. This was the only undesired outcome of the technical issues; in all other respects the participant in question contributed on the same level as the other participants. Table 5.2 shows how much time the participants spent working individually (in the case of Method 635 with inspiration from the other participants’ ideas) and as a verbally interactive group, respectively, during the application of each method. The execution of each method is presented in Table 5.3. Regarding the participant joining through video conferencing, the circulation of the sheets of paper containing suggestions (step 2 in Method 635, Table 5.3) was solved through scanning the sheets and sending them electronically to him. The sheets which he had worked on were sent electronically by him, printed out at the location of the group and handed out to the “neighbour on his left” (the “neighbours on his left and right” were determined before the ideation session started). During the gallery method, all the sheets of paper were scanned and sent to the person participating through video conference after step 1, and this person’s suggestions were sent by him and printed out at the location of the group after step 1 (before being put on a wall). During the application of the SIL method, the camera was directed towards the whiteboard so that the person participating through video conference could see what was sketched and written on it.

Table 5.2 The time (min.) spent on individual and group activities, respectively, and the total time (min.) for the execution of each method.

<table>
<thead>
<tr>
<th></th>
<th>Individual</th>
<th>Group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>30</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>Gallery</td>
<td>20</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>SIL</td>
<td>10</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

Different ideation topics were chosen for each ideation method, since using the same topic for each method may have resulted in the solutions from the previous ideation session(s) being reused in the subsequent session(s) and comparison of the ideation outcome between the methods would not have been meaningful. Instead, the aim was to find equivalent problems that were based on actual needs, were domain-specific and open-ended, and had a large amount of possible solutions. Firstly, the participants were asked to come up with topics which they wanted to work on within the OptiKrea project and which, based on their experience, they thought were equivalent. Thereafter, the scientific team made a review of the suggestions to pick out the topics that best fulfilled the equivalence criterion. The participants were informed of this selection and agreed upon the ideation topic about one week before each ideation session. The
Table 5.3 The execution of each ideation method selected.

**Method**

1. Each participant works individually during five minutes and comes up with at least three suggestions on how to address the ideation topic. The ideas are sketched and/or written down on a sheet of paper.
2. When the time is up, each participant gives the sheet containing the suggestions to the neighbour on their left.
3. Each participant reads through the suggestions on the sheet of paper which they have received from the neighbour on their right and adds improvements/comments to suggestions, combines suggestions to form new suggestions, and/or uses suggestions as a source of inspiration to come up with new ideas, during a total time of five minutes. One is allowed to ask the neighbour on one’s right what is meant by a suggestion that has been received.
4. When the time is up, each participant gives the sheet of paper which they have most recently been working on and which now contains the work of two participants to the neighbour on their left, and step 3 is repeated.
5. The process continues until each sheet of paper has been passed between all the participants, i.e. when it has been returned to the person who started working on it as a blank sheet of paper.

**Gallery method**

1. The participants individually sketch and/or write down suggestions on how to address the ideation topic on a sheet of paper for 15 minutes.
2. The sheets of paper are attached to a wall. The group gathers around one of the sheets hanging on the wall and the creator of the suggestions explains them to the other participants, who give constructive feedback. The group then moves on to the next sheet, which is explained, after which feedback is given, and this process is continued until all the participants have received feedback on their suggestions. Each participant can use approximately five minutes for presenting their ideas and receiving feedback.
3. Each participant takes down their sheet of paper and works individually on that sheet for five minutes to develop their suggestions or come up with new ideas using the feedback which they have received from the other participants and using the other participants’ suggestions as a source of inspiration.

**SIL method**

1. The participants individually sketch and/or write down their suggestions on how to address the ideation topic during ten minutes.
2. Two participants are randomly selected and each presents one of their suggestions to the rest of the group.
3. All the participants try to combine these suggestions into one concept by interacting verbally and sketching/writing on a whiteboard.
4. When the group is finished with the first two suggestions, a third group member presents another suggestion. The group then tries to combine this suggestion with the suggestion which resulted from step 3. Alternatively, a fourth suggestion is presented by any participant and an attempt is made to combine this fourth suggestion with the third suggestion.
5. The process of presenting suggestions and trying to combine them with each other continues until all the ideas have been presented or the time is up.
chosen ideation topics are presented in Table 5.4. The issues on which the ideation was performed required at least a basic technical understanding in general and railway-specific knowledge in particular.

<table>
<thead>
<tr>
<th>Method</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>How can deterioration of the track geometry in turnouts be prevented?</td>
</tr>
<tr>
<td>Gallery</td>
<td>How can the transition zones between rail sections of different rigidity be designed to ensure a smooth transition?</td>
</tr>
<tr>
<td>SIL</td>
<td>How can turnouts be protected from snow and ice?</td>
</tr>
</tbody>
</table>

At the start of each session, one of the participants presented the ideation topic, after which a common group-analysis of the topic took place for 30-60 minutes in order to identify causes of problems and known solutions to problems. Thereafter, one of the methods was tried. The author of this thesis presented the instructions for the method concerned to the group by means of a projector, ran a timer to keep track of the time, announced when the group should move on to the next stage of the method, and answered questions about the instructions during the session. During the test of the SIL method the participants became so confused about what they were supposed to do or reluctant to do it that the author had to take an unplanned active role in leading the session.

5.2.2 Views of the participants

After each ideation session, a short group interview regarding the participants’ experience of the applied method took place to capture the participants’ immediate reactions to the ideation method which they had tried; individual interviews about the ideation methods took place after all the methods had been tried. Moreover, after all the methods had been tried, each participant answered a questionnaire presenting different statements about each method. The participants answered on a continuous scale from “Do not agree at all” to “Agree completely”, by making a mark on a line which was about 100 millimetres in length and where “0” represented “Do not agree at all”. The position of the mark was measured in millimetres from “0” with a ruler and divided by the total length of the line. The average value for each statement was calculated and, despite the small number of participants, the average values for the subgroups of academics and railway professionals. The questionnaire which was filled in for each method is shown in Appendix 5.1 (translated from Swedish into English). The questionnaire was given to the participants in the form of an A4 hard copy. The participants were also asked to rank the methods from “one” to “three” based on their total impression of the methods.

The questionnaire formed the basis of an individual semi-structured interview. Questions about the group and the ideation methods were prepared in
advance and compiled in an interview guide, shown in Appendix 5.2 (translated from Swedish into English). The participants were encouraged to develop their replies. Thematic analysis was used to assess the content of the interviews. The aim was to use an inductive approach to make descriptive summaries of the participants’ views on working in the group and on the different methods which they had tried. The procedure used for the analysis is found in Table 5.5.

Table 5.5 The procedure used to analyse the interviews.

<table>
<thead>
<tr>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The recordings of all the interviews were transcribed by the author.</td>
</tr>
<tr>
<td>2. The recordings were listened to while the transcriptions were simultaneously perused. The quotes that were interesting with respect to the research questions were marked. Notes were made about the author’s reflections concerning these quotes, for the purpose of assisting in defining categories (themes).</td>
</tr>
<tr>
<td>3. Step 2 was repeated for each interview or part of an interview one or several times depending on the complexity of the text and how many research questions were involved in the text. (Sometimes this involved either listening to the recordings or reading the transcript.)</td>
</tr>
<tr>
<td>4. All the marked quotes were collected in a single word processing document. For each quote, information was provided as to which participant had contributed the quote.</td>
</tr>
<tr>
<td>5. Quotes that concerned similar matters were listed together. Tentative categories were formed. The data were perused as many times as required to find exhaustive and mutually exclusive categories (i.e. so that all the quotes could fit into a category and no quotes could fit into more than one category). Quotes that concerned more than one category were split if possible, or otherwise the category was changed to reflect the content of the quotes. The categories were given descriptive labels.</td>
</tr>
<tr>
<td>6. The quotes in each category were perused and sorted into subcategories as necessary (using the same procedure as for the main categories, see point 5). A count was made of the number of participants who had agreeing views in each category/subcategory. The content of the quotes in each category/subcategory was descriptively summarized and, if it was found that all the participants did not agree with the summary of the category/subcategory, information was provided in the summary as to how many did agree. As necessary, the author returned to the recordings or transcriptions to clarify if a certain participant had agreed, had not agreed, or had not expressed any opinion about the category/subcategory.</td>
</tr>
<tr>
<td>7. Point 6 was performed for each category.</td>
</tr>
</tbody>
</table>

5.2.3 Ideation outcome

5.2.3.1 Number of concepts and ideas

The number of concepts suggested by the group as possible solutions to each ideation topic was counted, as well as the number of concepts that were sketched. For the purpose of the present research, a concept has been regarded
as one idea or a combination of two or more ideas which serves as a solution to the problem on its own merits. A concept may address only one subpart of the problem and it may be possible to combine a certain concept with other concepts. As a starting point, concepts were defined as sketches and text that clearly belonged together and were marked off from other sketches and text on the sheets of paper. It was a straightforward task to identify the concepts on the sheets of paper. The concepts generated using the SIL method were distinguished in a different way from the concepts generated with Method 635 and the gallery method, because there were no sheets of paper with solutions after the ideation session using the SIL method. The concepts had instead to be identified by means of photographs of the whiteboard where the participants had made sketches of and written their ideas and on the basis of the transcript records. The concepts were constructed based on how the participants talked about them. Typically, a certain time interval would concern the same concept and then the group would switch over to a discussion of something new that was clearly separate from the preceding concept. The concepts turned out to be a very broad collection of solutions, ranging from a single abstract phrase to technical solutions which several participants had worked on. A popular-scientific overview of most of the generated concepts is to be found in Part II of this thesis. In all the examples of concepts and ideas presented throughout this chapter, the text or parts of text have been translated from Swedish into English. Examples of abstract concepts generated by a single person are shown in Table 5.6. These concepts are abstract since they describe something that should be carried out, but do not state how it should be accomplished.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>635</td>
<td>Recurrent tamping with stabilization until the track position is stable (3-4 times).</td>
<td>Recurrent tamping Stabilization 3-4 times</td>
</tr>
<tr>
<td>2</td>
<td>635</td>
<td>Decrease axle load.</td>
<td>Decrease axle load</td>
</tr>
<tr>
<td>3</td>
<td>635</td>
<td>Clean the ballast under the turnouts and drain.</td>
<td>Clean the ballast Drain</td>
</tr>
</tbody>
</table>

An example of a concept which several participants worked on is Concept 4 in Table 5.7. Four different people added ideas to the concept, consisting mainly of words and two sketches. Concept 5 in Table 5.7 is an example from the gallery method of a sketch with annotations which one person had worked on. The concepts contained not only physical suggestions on how to solve the problem, but also organizational, operational and maintenance-based suggestions which also clearly represented possible ways of addressing the ideation topics.
<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Sketch and text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>635</td>
<td><img src="image" alt="Slab cast in one piece under turnout's vulnerable parts" /></td>
<td>Slab cast in one piece under turnout’s vulnerable parts Crossing Switch system Injection casting Lifted to right position Jack Cranes Under sleeper pads softer than those used today</td>
</tr>
<tr>
<td>5</td>
<td>Gallery</td>
<td><img src="image" alt="Steel slab Angle Bolt fixation" /></td>
<td>Steel slab Angle Bolt fixation</td>
</tr>
<tr>
<td>6</td>
<td>SIL</td>
<td><img src="image" alt="Increase air current Fence Low, 5 dm Funnel-shaped Narrowest at the switch blades Shutter One joint Open when train passes Arm fixed to ground Hydraulic cylinder Telescopic Part of fence Increase air supply Plates on locomotive Turn up at turnout GPS control" /></td>
<td>Increase air current Fence Low, 5 dm Funnel-shaped Narrowest at the switch blades Shutter One joint Open when train passes Arm fixed to ground Hydraulic cylinder Telescopic Part of fence Increase air supply Plates on locomotive Turn up at turnout GPS control</td>
</tr>
</tbody>
</table>
Several concepts which several participants had worked on included suggestions on how to accomplish some function or detail that were incompatible with another suggestion concerning the same function or detail, either physically, because the suggestions could not be contained in the same concept, or logically, because it made no sense to use both alternatives at the same time or to solve the same function. When a larger part of the concept (more than half of its inherent ideas) remained the same upon changing between two different variants, the concept would be counted as one concept and the variants would be captured in the number of ideas, as is the case with Concept 4 in Table 5.7. This concept consists of ten different ideas. Of these ideas only two would not be compatible, i.e. using jacks and cranes to lift the slab, which is not meaningful to do simultaneously. Since this incompatibility involved less than half of the ideas, this concept was not split into two. If more than half of the ideas of a concept were incompatible, it was split into two or more different concepts.

When applying the SIL method, several times the group discussed a new concept and then made a suggestion on how this concept could be combined with a previously discussed concept. In these cases, the new concept was incorporated in the previously suggested concept if, as in Concept 6 in Table 5.7, less than half of the ideas in the combination of concepts were incompatible. In Concept 6 in Table 5.7, the first suggestion was to protect the turnout by means of a shutter. Another participant then saw the possibility of using the shutter as a part of a fence which he had been thinking of. In connection with a discussion about how to use locomotives to remove snow, e.g. with a plough or brush, it was suggested that locomotives could increase the air current through the fence by putting a big plate on their front. A summary of the rules for counting concepts is to be found in Table 5.8.

### Table 5.8 Summary of the rules for counting concepts.

<table>
<thead>
<tr>
<th>No.</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A concept consists of one idea or a combination of ideas, and stands as a solution to the problem on its own merit. The concept in question may address only one subpart of the problem and it may be possible to combine the concept with other concepts.</td>
</tr>
<tr>
<td>2.</td>
<td>If a concept contains incompatible ideas and half or more of the ideas are incompatible, the concept is split and counted as two or more different concepts that share some ideas.</td>
</tr>
<tr>
<td>3.</td>
<td>If more than half of the ideas of each of two concepts are shared by those concepts, the two concepts are counted as one concept containing variants.</td>
</tr>
<tr>
<td>4.</td>
<td>If a concept is a subpart of another concept, these two concepts are counted as one concept.</td>
</tr>
<tr>
<td>5.</td>
<td>Concepts that reframe the problem and do not specifically address the problem as described, but meet higher-level needs are counted.</td>
</tr>
</tbody>
</table>
The total number of non-redundant ideas was counted for each method. Because the concepts were of different types, it was challenging to define an “idea” as a unit for counting purposes. The procedure developed by Linsey et al. (2005, 2011) was used as a starting point. In this procedure, the basic definition of an idea is that it is something that solves one or more functions in the functional basis (Linsey et al., 2011), with the functional basis defined as a standardized set of function-related terminology that allows repeatable and meaningful function representation (Hirtz et al., 2002). However, as a result of the ideation topics, it was not meaningful to try to fit several of the generated concepts into a functional basis, and other generated concepts required completely different functional representations. For the purpose of the present research, ideas were instead defined as the units which each concept could be systematically decomposed into, and which could be expressed as a key phrase consisting of a verb phrase containing a maximum of one verb, or a noun phrase. In this way it was possible to capture ideas from concepts of different types that were expressed by different means. The counting rules are shown in Table 5.9.

<table>
<thead>
<tr>
<th>No.</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Ideas are units which a concept is systematically decomposed into and an idea can be expressed as a key phrase consisting of a verb phrase containing a maximum of one verb, or a noun phrase.</td>
</tr>
<tr>
<td>2.</td>
<td>If a certain idea is used multiple times in a concept, this idea is counted once for that concept.</td>
</tr>
<tr>
<td>3.</td>
<td>Ideas in different concepts are counted as the same idea if they imply physically doing the same thing, although different words are used to describe that idea.</td>
</tr>
<tr>
<td>4.</td>
<td>New combinations of already-counted ideas are counted as separate ideas.</td>
</tr>
<tr>
<td>5.</td>
<td>Categories of ideas only count as ideas when no subordinates are given.</td>
</tr>
<tr>
<td>6.</td>
<td>Ideas count even if they are unnecessary or deteriorate the concept.</td>
</tr>
<tr>
<td>7.</td>
<td>Ideas must be shown (words and/or sketches) and not just implied.</td>
</tr>
</tbody>
</table>

Examples of how ideas were counted are shown in Table 5.6 and Table 5.7. In Concept 4 in Table 5.7, the sketches illustrate what is mentioned in the text and do not add any additional key phrases. Concept 5 in Table 5.7 involves attaching a steel slab onto the concrete slab at an angle with bolts to soften the transition between track areas of different rigidity. The transition from ordinary rail to slab track was derived from the ideation topic and does not count as an idea. This way of defining ideas reflected the scope of a concept (how much work and thought had been put into it) and made it easy to find redundant ideas that had been used in several concepts. However, it did not cover all the information contained in sketches, e.g. the positioning of items in relation to each other. Judging the redundancy turned out to be a straightforward task, as ideas were
either identical (e.g. the use of sleeper pads was a common suggestion when testing Method 635) or clearly different.

A valid question is whether information contained in the audio recordings but not written down or sketched by any participant should be included in the analysis of concepts and ideas, since, without the recordings, this information would either have been lost or possibly stored only in the mind of the participants. It was decided that the recorded information would be used to show the potential of the ideation methods. Otherwise a concept that was only described by spoken words would not have been captured through the inherent documentation of the methods. The extent to which concepts were actually documented during the ideation sessions is captured through the elaboration of the concept descriptions, which is presented in the next section.

5.2.3.2 Elaboration of concept description

The elaboration of each concept description was defined as the number of sketched details, the number of written words, and the number of spoken words used during verbal interaction. Sketched details were defined as follows: a) physical objects, b) intentional holes through a physical object, c) indications of dimensions, d) indications of quantity and e) arrows indicating movement. The counting rules are shown in Table 5.10.

Table 5.11 shows examples of how sketched details and written words were counted. Concept 7 in Table 5.11 shows a straightforward example containing two rails and six sleepers, counted as eight sketched details. Concept 8 in Table 5.11 shows another example with more and different sketched details. A hole created because it was surrounded by different other objects, like the hole in the area between the rail, two sleepers and the ballast body in Concept 8, was not included. In contrast, the holes in the sleepers in Concept 9 in Table 5.11 were counted. Each length or angle indicated is counted as a separate detail. An example of an indication of dimension is shown in Concept 9 in Table 5.11.

The elaboration of a concept measures the amount of details given about the concept. It reflects how much work the participant(s) put into communicating the concept and how much discussion took place regarding the concept during the ideation session. A high degree of elaboration does not necessarily mean that the concept in question is of high quality or involves a large number of ideas.
Table 5.10 Summary of the rules for quantifying the elaboration of a concept description.

<table>
<thead>
<tr>
<th>No.</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The elaboration of a concept description is defined as the number of sketched details, written words, and spoken words related to a concept in its description.</td>
</tr>
</tbody>
</table>
| 2.  | The number of written or spoken words is counted, without distinguishing between the different parts of speech, on the sheets containing solutions and the transcript records, respectively.  
   a) An abbreviation of one word is counted as one word.  
   b) An abbreviation of several words is counted as the number of words included in the abbreviation.  
   c) Sketched arrows indicating “gives”/ “results in” are counted as one word.  
   d) A number is counted as one word. |
| 3.  | Sketched details are defined as:  
   a) physical objects,  
   b) intentional holes through a physical object,  
   c) indications of dimension (only lengths and angles were given in the present set of concepts),  
   d) indications of quantity,  
   e) arrows indicating movement. |
| 4.  | A physical object is a discrete object that must be possible to distinguish in the sketch to be counted as a detail. The shape of the object is not considered. |
| 5.  | Holes are counted as objects if they are added intentionally and go through a physical object. |
| 6.  | Each length or angle indicated is counted as a separate indication of dimension. |
| 7.  | Each quantity indicated is counted as a separate indication of quantity. In a diagram, different values of dependent quantity are counted as separate indications of quantity. A changing quantity of constant slope is only counted once. |
Table 5.11 Examples of concepts and the number of sketched details and written words in the concepts.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Sketch and text</th>
<th>Sketched details</th>
<th>Written words</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>635</td>
<td></td>
<td>8 physical objects</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2 rails</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-6 sleepers</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Gallery</td>
<td></td>
<td>17 physical objects</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 rail</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-4 sleepers</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-1 ballast body</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-11 glued lumps</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Gallery</td>
<td></td>
<td>9 physical objects</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-2 rails</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-7 sleepers</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 holes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 indications of dimensions</td>
<td></td>
</tr>
</tbody>
</table>

5.2.3.3 Viability and ability

The quality of an idea was described in terms of its viability and ability, which reflect its usefulness. The viability of an idea describes the possibility of practically implementing the solution with respect to technology, cost, and safety. The current rules and restrictions were not considered, since these can be changed. The ability of an idea describes its capacity to solve the problems of the ideation topic in question. While the other parameters were analysed by the author of this thesis, the viability and ability were judged by the participants individually about three months after the last ideation session by means of a questionnaire. The participants received textual instructions concerning what to consider when judging the viability and ability, and a description of each concept (copies of sketches and written words from the ideation sessions and a summary of spoken key phrases, when relevant). They were asked to make a mark on a line between “0” and “1”, indicating how viable or able, respectively, they thought each concept was. “Zero” corresponded to “not at all viable/able” and “one” corresponded to “completely viable/able”. The value was calculated using the same procedure as was used for the questionnaire described in Section 5.2.2. Some of the concepts were combinations of suggestions that could have
been a concept on their own merits. The participants were asked to judge the
different inherent concepts in a sample of such concepts. The participants could
choose themselves how much time they put into the judgment process and the
participant who did it most carefully used about three hours. They did not use
external information, but relied on their knowledge and prior experience. A
check was performed to determine whether there were any obvious differences
between how the academics and the railway professionals, respectively, judged
the ideas. The instructions for the questionnaire (including examples and trans-
lated from Swedish into English) are given in Appendix 5.3 at the end of this
chapter. In addition to the instructions in Appendix 5.3, the participants were
informed (in the e-mail sent with the questionnaires) a) that they should ignore
existing regulations during their judgment of the viability, b) that they should
evaluate “the concept’s ability to solve the problem” of the actual ideation topic,
and not whether the concept was a good concept in general or from some other
point-of-view, c) about the other parameters to be used in the evaluation of the
ideas (some of these parameters were, however, changed later), and d) that they
did not have to think too much about each concept and could use their “gut feel-
ing” (which was stated partly not to scare them away from making a judgment,
which involved quite a substantial amount of work).

The parameters used to measure the ideation outcome are summarized in
Table 5.12.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Measurement method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>The number of: non-redundant ideas, non-redundant concepts, sketched concepts, non-sketched concepts</td>
<td>Counted by the author of the thesis according to the rules in Table 5.8 and 5.9.</td>
</tr>
<tr>
<td>Elaboration</td>
<td>The number of: details in sketches, written words, verbal words used in the statement and elaboration of the concept</td>
<td>Counted by the author of the thesis for each concept according to the rules in Table 5.10.</td>
</tr>
<tr>
<td>Viability</td>
<td>The possibility of practically implementing an idea with respect to technology, cost, and safety.</td>
<td>Judged by the participants individually on a continuous scale between 0 and 1 (0=not at all possible, 1=completely possible).</td>
</tr>
<tr>
<td>Ability</td>
<td>The ability of the idea to solve problems connected with the ideation topic in question.</td>
<td>Judged by the participants on a continuous scale between 0 and 1 (0=not at all able, 1=completely able).</td>
</tr>
</tbody>
</table>
5.2.4 Behaviour of the participants

To analyse the content of the ideation sessions, a variant of the coding system developed by Jackson and Poole (2003) was used. Instead of the time used for the different activities, the number of words used in the different activities was counted. The main activities specified by Jackson and Poole (2003) were used, i.e. idea statement, elaboration, criticism, direction, and going off at a tangent. A description of each activity is to be found in Table 5.13.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea statement</td>
<td>Spoken contribution to be recorded as a possible solution to the problem.</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Non-critical clarification (explaining), discussion or rephrasing of ideas.</td>
</tr>
<tr>
<td>Criticism</td>
<td>Negative statements or judgments about proffered ideas.</td>
</tr>
<tr>
<td>Direction</td>
<td>Guiding or structuring the idea generation activity.</td>
</tr>
<tr>
<td>Going off at a tangent</td>
<td>Interaction that is off-topic and breaks the “singlemindedness” of the idea generation, whether or not it relates to other group work tasks.</td>
</tr>
</tbody>
</table>

An analysis was performed of the distribution between the participants of the words spoken by them during the ideation sessions. During this analysis, agreeing expressions consisting of a maximum of three words that were not an answer to a direct question were neglected (e.g. “yes”, “h’m”). This was partly because such expressions did not add anything to the discussion and partly because they were difficult to hear on the recordings (especially when uttered by those participants sitting furthest away from the microphone) and counting them would therefore not give a fair picture of the speaking distribution. However, when a speaker was interrupted by such a comment and then had to “speak up” again to continue, this was counted as speaking twice, since speaking up again after being interrupted takes an effort and is important in understanding the speaking distribution. The same rule applied when a speaker said something that made the others laugh and then had to start speaking again. It was not possible to interpret all the verbal communication of the recordings of the workshops, and the uninterpretable part was not included in the analysis. However, this part is so small (estimated to be roughly 1% of the total number of words, and impossible to calculate exactly, since one cannot judge exactly how many words are not heard in each instance) that it would only have had minor effects on the results of the analysis and would not have changed the overall conclusions.
Recordings from and observations of the ideation sessions were continuously analysed during Stage 2 of the research for this thesis, with respect to the behaviour of the participants. The purpose of this analysis was a) to follow up on other results from the study (e.g. from the interviews and content analysis) to determine whether these results were reflected in the recordings and observations, b) to follow up on earlier findings in the literature that might be possible to compare to the findings of Stage 2, and c) to find data that were relevant to the research questions, but had not been specifically investigated in any other way (e.g. because one had not expected that this type of data would turn up). This analysis had a broad approach and the author tried to keep an open mind and follow up on “intuitive guesses” during the study. Once interesting data were found, they were analysed similarly to the procedure described in Table 5.5 or that described in Table 5.14, depending on the nature of the data.

The recordings were specifically used to analyse if and how the diversity of the group had manifested itself during the interaction between the participants during the ideation sessions. The procedure for the analysis is shown in Table 5.14.

**Table 5.14** The procedure used to analyse the audio recordings from the ideation sessions with respect to the diversity of the group.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The recordings of all the ideation sessions were transcribed by the author.</td>
</tr>
<tr>
<td>2</td>
<td>The recordings were listened to while the transcriptions were simultaneously perused. The excerpts that were interesting with respect to the diversity of the group were marked. Notes were made about the author’s reflections concerning these data, for the purpose of assisting in defining categories (themes).</td>
</tr>
<tr>
<td>3</td>
<td>Step 2 was repeated for each ideation session or part of an ideation session one or several times depending on the complexity of the text. (Sometimes this involved either listening to the recordings or reading the transcript.)</td>
</tr>
<tr>
<td>4</td>
<td>All the marked excerpts were collected in a single word processing document. For each quote contained in an excerpt, information was provided as to which participant had contributed the quote.</td>
</tr>
<tr>
<td>5</td>
<td>Excerpts were listed together which were found to represent similar ways in which the diversity expressed itself in the group during the group members’ interaction. Tentative categories were formed. The data were perused as many times as required to find exhaustive and mutually exclusive categories (i.e. so that all the quotes could fit into a category and no quotes could fit into more than one category). The categories were given descriptive labels.</td>
</tr>
<tr>
<td>6</td>
<td>For each category, 2-3 excerpts that illustrated the descriptive labels were chosen and were presented together with the labels as the results from this analysis.</td>
</tr>
</tbody>
</table>
5.3 Results and discussion

5.3.1 Views of the participants

Sections 5.3.1.1-5.3.1.4 present the results regarding the participants’ experience of the group composition and the ideation methods, and rely on data from the interviews, questionnaires, and observations. Quotes (translated from Swedish into English) are included to exemplify qualitative data that contributed to the results. The letter in brackets after each quote indicates who is quoted, according to Table 3.1.

5.3.1.1 Group composition

All the participants were convinced that the composition of the group, with members expressing different views on the subjects, was a winning concept. They all thought that the group possessed a large amount of competence and experience, and that this competence and experience facilitated the finding of good solutions and helped them avoid continuing to work on solutions that had previously proven to be bad. This is exemplified by the following quotes:

“[Bringing up the same problem at your own place of work] would have been something completely different. [In this group we] come from the different parts [of the railway sector], a turnout manufacturer, the infrastructure manager, STA, and […] a contractor; that’s as clear as a bell. It’s a strength.” (A)

“People are sometimes working on things that are not especially successful. Abortive. But if we had been able to sit together [and work] in teams like this, we could have avoided that. Look out, be careful, this is going to cause you trouble.” (A)

All the participants thought that the presentation and discussion of the ideation topic were important due to the diversity of the group. It was important to achieve a common understanding of the issue and survey the known causes before starting ideation, otherwise one had different starting points and understanding of the topic. This is exemplified by the following quote:

“I think it has something to do with the participants too, […] quite a motley crew here, all of them with different angles of approach. If [X] and I had been sitting with our colleagues, sitting in one of those more homogeneous groups, probably with roughly the same picture of [the current situation], then [a group-analysis of the problem] would have perhaps felt a bit like a creativity killer. Then we could have started [ideating] directly.” [In this quote, “X” stands for one of the members of the OptiKrea Group.] (C)
In addition, the participants thought that it was beneficial to use the stated causes of problems as inspiration when generating ideas. They mentioned that they had reviewed the causes to find new angles of approach, as is exemplified by the following quote:

“I noticed myself that, when I tried to come up with an idea, [I thought about] what the origin of [this problem] was. You see, so to speak, for example [the stated cause] that water is accumulated, well, then you think, we must have some solution, better drainage. It was helpful to have a list of possible causes.” (E)

All the participants thought that working in the group was really enjoyable, inspiring, stimulating and motivating, and that its activities could improve turnouts, as is exemplified by the following quote:

“Yes, it’s stimulating [to be in the group], I think it’s great fun.” (A)

Generating ideas was, however, not the only objective of participation in the group. Using it as a means of interfacing with the other group members was also important. Naturally, the industrial participants wanted to promote their company and show the infrastructure manager why they should be their choice among competing companies. This is exemplified by the following quote:

“The group is interesting; it gives contacts that have been cut off thanks to the [rules concerning] restrictive practices.” (A)

Both the industrial participants mentioned that if the competitors of their companies had participated, they would not have been able to speak freely and would have kept good ideas to themselves. This effect has been observed in another related project run by Prof. Jan Lundberg of LTU, where several actors from the same actor segment participate.

5.3.1.2 Method 635

Method 635 was perceived as stressful by the majority of the participants. They thought that it was tough to describe one’s suggestions and ideas in five minutes without talking, especially if one is not skilled in drawing. This is exemplified by the following quote:

“I thought five minutes was dreadfully stressful.” (B)

On the other hand, half of the participants stated that they had experienced the stress factor and the regular input of ideas from the other participants as exciting and enjoyable, as is exemplified by the following quote:
“I thought it was a bit more thrilling. […] You sort of felt tossed around. It was kind of fun, I think.” (F)

They considered the method to be very dynamic because new ideas were continuously generated by the input of ideas from the other participants. According to the participants, it was inspiring to think about other participants’ suggestions and, after exhausting one’s own ideas, the input from others helped in generating new ideas. The participants thought that the time restriction prevented the apprehension of their own ideas. Some participants described how they were pushed out of their normal thinking paths. This is exemplified by the following quote:

“You are probably most […] creative during the first minutes and then not very much happens. I can sit for three hours and nothing happens. I think in the same old rut. It’s better then to throw in something new, […] new input and then you empty yourself [of the new ideas] and then the next [input] comes.” (B)

As the sheets of paper with ideas were circulated, the participants thought that the suggestions were elaborated more and more, making it harder to add something. This is exemplified by the following quote:

“So you could also see that the ideas arriving were more and more developed; there was less and less to give, so to speak. As a result, you didn’t contribute anything, of course.” (C)

The participants thought that the high productivity might mean that the quality of the ideas was worse and the applicability lower compared to the other methods. An element of scrutiny was desired by the participants, as some suggestions were perceived as unrealistic. This is exemplified by the following quote:

“Very productive method, it may be the case that the quality and applicability are lower.” (B)

5.3.1.3 Gallery method

The participants appreciated the fact that, when applying the gallery method, they had had more time available than when they had worked with Method 635. There had been time for describing and understanding the suggestions. This is exemplified by the following quote:

“There was more room for thought. Time is a fairly important parameter when you are sitting thinking.” (C)
The method was described as being a comfortable method to use. One’s own suggestions could be developed in peace and quiet, and then the benefits and drawbacks could be figured out. The participants’ presentations and the feedback from the other participants were highly appreciated since they gave insights that the receiver of the feedback had not thought of himself, as is exemplified by the following quote:

“I think it’s very important [to get feedback], that’s what gives additional [insights].” (A)

The participants valued the fact that, afterwards, time was available for revising the suggestions based on the feedback and by using ideas from other participants.

Half of the participants, all of them railway professionals, selected the gallery method as their preferred method, as can be observed in Table 5.15, which shows how each participant ranked the three methods.

<table>
<thead>
<tr>
<th></th>
<th>Railway professionals</th>
<th>Academics</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>A 1 B 2 C 2 D 3 E 3 F 1</td>
<td>13 2.2</td>
</tr>
<tr>
<td>Gallery</td>
<td>A 3 B 1 C 3 D 3 E 2 F 1</td>
<td>13 2.2</td>
</tr>
<tr>
<td>SIL</td>
<td>A 2 B 3 C 1 D 1 E 2 F 2</td>
<td>10 1.7</td>
</tr>
</tbody>
</table>

One railway professional, who in the end chose the SIL method, said that he too would have chosen the gallery method as his preferred method if the choice had been based on what method he felt was the nicest. Nevertheless, he finally chose the SIL method as his number one method because he thought that it pushed him out of his comfort zone and forced him to think in new ways. Both academics chose Method 635 as their preferred method.

The gallery method was described by one participant as being rather ordinary, resembling a normal meeting:

“[The gallery method] was actually rather ordinary. What ideas do you have? Write them down on a sheet of paper and present them.” (F)

The participants thought that the productivity was lower and that the concepts were better devised and had higher applicability compared to those generated during Method 635, as exemplified by the following quote:
“The advantage, of course, is that you get a little more time to look into a small number of suggestions; [...] one gets quite a lot of details, but maybe [the gallery method] is somewhat limited in the idea generation phase, because it doesn’t give so many ideas.” (E)

5.3.1.4 SIL method

The participants thought that combining ideas was a way to come up with more suggestions and that by combining ideas they were forced to think differently, as exemplified by the following quote:

“Normally I would never [combine] certain things; I would never even dream of combining two [of those ideas], but here you are forced to try and concoct something, and at that point, I think, then you have really left your comfort zone and [started thinking in a new way].” (B)

However, not all the participants were impressed by the method. One participant found it abstract and forced, and could not see the point of it. The majority of the participants were reluctant to try to combine ideas that initially appeared impossible to combine (see further in Section 5.3.3 and Table 5.22). During the ideation session they had to be convinced by the author that they should actually continue trying, and then they discovered that they had been able to create new suggestions.

During the general discussions on how to combine ideas, there was naturally an element of scrutiny present, since the participants had to consider if and how ideas could be combined and what needed to be added in order to make a suggestion work. This element of scrutiny was appreciated by all the participants, since they were anxious not to work on suggestions that were completely unrealistic and eager to find ways of improving a good suggestion. This is exemplified by the following quote:

“To elaborate on almost unrealistic things feels a bit wrong somehow, but to inject some quick criticism: stop, there it’s getting unrealistic.” (A)

5.3.1.5 Questionnaire

Figure 5.1 presents how the participants on average rated different aspects of each method in the questionnaire. The statements are given in Table 5.16. The questionnaire results were also analysed separately for the academics and railway professionals. It turned out that the variation between participants was greater than that between the two subgroups, and therefore no conclusions about possible differences between the two subgroups could be established. The only tendency that could be determined was that the railway professionals found the gallery method more useful than the academics, who preferred Method 635 in
this respect, similarly to what was found based on the ranking of methods in Table 5.15.

![Figure 5.1](image.png)

**Figure 5.1** The participants’ average level of agreement with different statements about the method.

**Table 5.16** The statements in the questionnaire.

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am engaged in the topic.</td>
</tr>
<tr>
<td>2</td>
<td>If I had had more time, I would have come up with more ideas.</td>
</tr>
<tr>
<td>3</td>
<td>I had more ideas than I presented.</td>
</tr>
<tr>
<td>4</td>
<td>Many interesting ideas were presented during the workshop.</td>
</tr>
<tr>
<td>5</td>
<td>Many ideas were new to me.</td>
</tr>
<tr>
<td>6</td>
<td>I think ideas that were presented can be used in practice.</td>
</tr>
<tr>
<td>7</td>
<td>I felt engaged during the workshop.</td>
</tr>
<tr>
<td>8</td>
<td>I am satisfied with my own contribution during the workshop.</td>
</tr>
<tr>
<td>9</td>
<td>I think the method is useful.</td>
</tr>
</tbody>
</table>

The gallery method was popular and received the highest average scores from the participants for several parameters, such as “I think the method is useful” and “I felt engaged during the workshop”, as shown in Figure 5.1. When Shah et al. (2001) compared C-sketch to the gallery method and Method 635 (using only textual description), the gallery method was also found to be popular and scored highest with regard to the participants’ views on the creative outcome and the promotion of creative cognitive processes (Shah et al., 2001). The gallery method allowed the participants to talk and discuss, which they
expressed was very exciting and important to them. However, the SIL method allowed even more time for group work, but was still not as popular. Most of the participants actually preferred the silent Method 635 to the SIL method when ranking the methods, see Table 5.15. This implies that working in a verbally interactive group is not sufficient to create an enjoyable session. Possible reasons for the SIL method being the least popular method may be that the participants felt uncomfortable trying to combine ideas which they did not consider to fit together, and that they perceived the method as less structured than Method 635 and the gallery method. It was expected that the SIL method would challenge the participants, but not to such a degree that they would become reluctant to use it.

It is interesting to note that all the railway professionals preferred the gallery method (except for one, who said that he would have chosen it if he had only based his choice on what method felt most comfortable), whereas the academics preferred Method 635. This might reflect the way in which the participants are used to working in their own organization. An interesting finding was that the professionals working at STA consistently awarded the lowest marks or marks that were among the lower marks for “Many ideas were new to me” compared to the other participants. An explanation could be that, as employees of the infrastructure manager, they are exposed to more ideas from others compared to the other members of the team.

5.3.2 Ideation outcome

The concept and idea generation rates are presented in Figure 5.2. Figure 5.3 shows the number of ideas for each combination of viability and ability for each method. Since the methods had different total times for ideation, the histograms are normalized to 60 minutes. The viability/ability of each idea is the average of the participants’ ratings.
Figure 5.3 Histograms, normalized to 60 minutes, of concepts as a function of their average ability and average viability as judged by the group members.

In terms of concept and idea generation rates, Method 635 outperformed the other methods. Since verbal interaction has been found to contribute to process losses (Mullen et al., 1991), this result is not surprising as Method 635 provided more time for individual ideation than the gallery and SIL methods. However, other factors might also have played a role in this case. The results might be partly explained by expectations. The participants knew that the “3” in Method 635 stands for coming up with three ideas in every round. Although the participants were informed that fewer or more ideas would also be acceptable, they probably felt that they were expected to generate several ideas before giving the sheet of paper to the next person. During the other workshops no such specific goal was specified. Specifying goals has been found to increase performance during group work (Wegge and Haslam, 2005) and even eliminate the productivity gap between nominal and interactive groups (Paulus et al., 1993). The stress factor probably also exerted an influence, since the time constraint prevented the participants from evaluating their own ideas before writing them down. In addition, previous research has shown that individuals under time pressure work at a faster rate and that time limits are inversely related to the amount of task focus shown by groups (Karau and Kelly, 1992).

The gallery method invited the participants to sketch their suggestions and this method had the highest fraction of sketched concepts, see Figure 5.2, as expected. Method 635, however, had almost as high a number of concepts including a sketch, generated over a shorter period of time. Table 5.17 shows the
distribution of the number of sketched details, written words and verbal words describing each concept. Not all the concepts were described by both sketches and written words. Therefore, the average number of sketched details was calculated with respect to both the total number of concepts and the number of concepts with at least one sketched detail; and the average number of written words was calculated with respect to both the total number of concepts and the number of concepts with at least one written word. The methods exhibited different characteristics with regard to the distribution of sketched details, written words and verbal words in the concept descriptions. This is interesting, because only sketches and written words are saved for post-ideation processing if no additional documentation is performed, and sketches typically contain information that is difficult to capture by textual description. Not surprisingly, the number of spoken words per concept increased and the number of sketched and written details decreased as the time for verbal group ideation increased, see Table 5.2 and Table 5.17.

Table 5.17 Distribution of the sketched details and written and verbal words in the descriptions of concepts. The figure in parenthesis shows the value in front of it divided by the time in minutes that was used and the number of concepts generated during the ideation session.

<table>
<thead>
<tr>
<th></th>
<th>635</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of sketched details</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for all the concepts</td>
<td>5.9 (0.0042)</td>
<td>18.2 (0.014)</td>
<td>3.2 (0.0021)</td>
</tr>
<tr>
<td>Average for concepts with at least one sketched detail</td>
<td>12.8 (0.0091)</td>
<td>20.4 (0.016)</td>
<td>9.1 (0.0061)</td>
</tr>
<tr>
<td><strong>Number of written words</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for all the concepts</td>
<td>15.4 (0.011)</td>
<td>11.2 (0.0086)</td>
<td>4.9 (0.0033)</td>
</tr>
<tr>
<td>Average for concepts with at least one written word</td>
<td>15.4 (0.011)</td>
<td>12.0 (0.0092)</td>
<td>6.2 (0.0041)</td>
</tr>
<tr>
<td><strong>Number of verbal words</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for all the concepts</td>
<td>0 (0)</td>
<td>105 (0.080)</td>
<td>274 (0.18)</td>
</tr>
</tbody>
</table>

The gallery method had a considerably lower concept and idea generation rate than Method 635 (see Figure 5.2). This is in accordance with the participants’ perception, as found from the interviews, that the number of suggestions was smaller with the gallery method than with Method 635, although this was not reflected in the questionnaire. The participants suggested that the concepts would be better devised, and that the quality, as well as the applicability, might be higher with the gallery method. Table 5.17 shows that the number of sketched details in the concepts from the gallery method was higher compared to the corresponding number for the other methods, but on average the gallery method did not result in more ideas per concept than Method 635, despite the
fact that the average available time for working on each concept was considerably higher with the gallery method (see Table 5.18).

### Table 5.18 Average number of ideas and average available time for each concept.

<table>
<thead>
<tr>
<th></th>
<th>635</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of ideas per concept</td>
<td>3.4</td>
<td>3.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Average time available for each concept (min.)</td>
<td>3.9</td>
<td>9.3</td>
<td>15.7</td>
</tr>
</tbody>
</table>

This indicates, although some information in the sketches might have been lost in the idea counting, that the participants sketched their concepts more carefully when working with the gallery method, and the resulting sketches might have been easier to comprehend, but the participants did not use the time to add ideas as to how to solve problems connected with different features of their concepts or to create new concepts. For example, adding more sleepers and holes, as in Concept 9 in Table 5.11, adds more sketched details that can clarify the concept to others, but does not add any new ideas as to how to address the ideation topic. According to Table 5.19, neither the average viability nor the average ability differs significantly between the methods. However, Figure 5.3 shows that the average ratings of the ability and viability of the ideas generated when applying the gallery method are less scattered than the corresponding ratings of the ideas generated with the other methods, especially Method 635, which might explain the participants’ perception that concepts generated using the gallery method would be of higher quality.

### Table 5.19 Viability with respect to technology, cost and safety, and the ability of the idea to solve the issue.

<table>
<thead>
<tr>
<th></th>
<th>635</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.62</td>
<td>0.56</td>
<td>0.54</td>
</tr>
<tr>
<td>Average range</td>
<td>0.78</td>
<td>0.61</td>
<td>0.67</td>
</tr>
<tr>
<td>Average standard deviation</td>
<td>0.29</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td>Ability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.54</td>
<td>0.66</td>
<td>0.64</td>
</tr>
<tr>
<td>Average range</td>
<td>0.59</td>
<td>0.51</td>
<td>0.57</td>
</tr>
<tr>
<td>Average standard deviation</td>
<td>0.21</td>
<td>0.18</td>
<td>0.21</td>
</tr>
</tbody>
</table>

As can be observed in Figure 5.1, Method 635 scored lower than the other methods with respect to the participants’ perception of the possibility of using the ideas presented in practice. One explanation might be the very high number of generated ideas, including more low-rated ideas, compared to the other methods. These low-rated ideas might have dominated the participants’ perception of the method, although it actually generated a higher number of ideas with
high ability than the other methods, as shown in Figure 5.3. In the study by Shah et al. (2001), C-sketch and the gallery method had higher mean quality scores than Method 635 (using only textual description). However, since they did not mention how many ideas from each method had a certain quality (and variety and novelty) score (only average values were given), it is still possible that Method 635 had a similar or higher absolute number of ideas with a high-quality score, if many more ideas were generated. This would be interesting to know, as this was found to be the case in the present study.

Linsey et al. (2011) found, contrary to this study, that when participants were allowed to use both sketches and words to describe solutions, gallery viewing generated more product solutions (which are analogous to concepts in the present study) than rotational viewing (during a 40 minute session). Rotational viewing, however, produced an overall greater number of ideas per function than gallery viewing. A major difference between the execution of the gallery method in the present study and how the method was executed in Linsey et al.’s study (2011) was that no presentation and discussion step was included in the gallery viewing condition in the latter study. Therefore, one possible explanation of the different findings may be that the time used for presentation and discussion in the present study resulted in significantly fewer suggestions being generated in this step than in the ideation steps, and that this was not compensated for by an increased number of ideas being generated in the following short ideation step. The role of presentation and discussion steps is investigated in Chapter 6.

In conclusion, the participants had on average a realistic view of the concept quantity, but misjudged the concept quality after the ideation sessions. This is intriguing, since they themselves judged the ability and viability of the concepts, which means that, when they themselves had to think specifically about the usefulness of each concept, they reached conclusions that were different from what they spontaneously felt about the collection of concepts after the ideation sessions.

The participants were also asked to judge a sample of concepts at different stages of the ideation session, to find out if they thought that the concepts had been improved by the elaboration of others (in the case of Method 635) or by the performed combinations (with the SIL method). It turned out that these samples included concepts that had improved and concepts that scored approximately the same or worse after receiving additions or being combined with another concept or idea. An example of a concept that became worse after it had received an addition from another participant is Concept 4 in Table 5.7. In this example, the initial suggestion was to have a slab cast in one piece under the turnout’s most vulnerable parts and to post-adjust it by means of injection casting while lifting the slab to the right position with a jack or a crane. This suggestion received an average viability score of 0.74 and an average ability score of 0.60. When one participant added the idea of a having a separate slab for the point driver, the average viability score did not change considerably, but the
average ability score decreased to 0.40. When a third participant added the idea of under-sleeper pads to attain an exact rigidity, the concept received an average viability score of 0.92 and an average ability score of 0.77, which constitutes a considerable increase. However, when a fourth participant suggested making the under-sleeper pads softer than those used today, the scores decreased to 0.75 and 0.45, respectively. Another example is the suggestion of letting an engine remove snow by means of a brush or some other special mechanical device, which received an average viability score of 0.63 and an average ability score of 0.67. However, when another participant added the idea of adapting the turnout to the engine’s special devices, the average viability score stayed the same, but the ability score decreased to 0.33. The overall score given to these concepts concerned the entire concept, including the “bad” modification. It is not within the scope of this study to investigate this in detail, but rather to conclude that modifications of concepts can deteriorate the original concept. Linsey et al. (2011) noted that the quality of a product solution frequently changes as team members add their ideas. They found that embellished product solutions tend to be higher-quality product solutions and gave examples of how an overall product solution can drastically improve as individuals add ideas (Linsey et al., 2011). They did not, however, investigate how product solutions changed for the worse as they received additions and how common this was. Shah et al. (2001) found that when the gallery method was used, the first ideation step produced suggestions with high scores for variety, but low to medium scores for quality. After the second ideation step, the scores for variety decreased, but the quality of the ideas improved. Shah et al. (2001) suggested that the participants picked up ideas from others that improved their own concepts, but also made them less varied. It appears that there is a knowledge gap concerning how often individual concepts deteriorate or improve as they are modified. This is important to know when the group moves on to concept selection. If a concept receives an overall low score, there might still be high-quality features embedded in that concept that can be useful when developing the end product.

From Table 5.19 it can be observed that the average range (the average difference between the minimum and the maximum score) and the average standard deviation of the viability and ability are quite high for all the methods. This reflects the fact that the participants in general answered differently as to how viable and able the ideas were. In some cases, the range for a specific idea was 1, meaning that at least two participants totally disagreed on its viability or ability. No difference could be found between the judgements of ideas made by the subgroups of railway professionals and academics (the variation between the participants was similar to or greater than that between the average values of the two subgroups). This is in contrast to previous studies using raters, where the inter-rater agreement concerning quality-related parameters is typically quite high. This is believed to be a result of the domain-specific topics and the ability of the highly experienced expert participants to generate advanced solutions. These factors made it much more difficult in the present study to judge the qual-
ity of ideas in comparison with previous studies without domain-specific topics and with participants who did not possess expert knowledge or experience of those topics. It is an important finding that, although all the participants had substantial knowledge of and experience from the railway sector in general and turnouts in particular, at this stage they had different opinions on what ideas were useful. An additional reason for the participants’ different opinions might be their different perspectives on the ideas and what they actually knew about the possibility of implementing them. It is interesting to note that Chulvi et al. (2012) used expert raters with at least eight years of professional experience in the domain-specific area concerned to rate solutions generated by multidisciplinary teams of PhD students or professional designers working on design problems; they found, similarly to this study, that the experts’ responses had a significant dispersion, especially concerning the degree of usefulness. Chulvi et al. (2012) suggested that this might “indicate that in the absence of available data, even experts find it difficult to assess and compare the potential usefulness of a product”. This finding implies that it is inadvisable to discard ideas at this stage, especially based on the opinion of only one person, since all the possibilities of the solution might not be understood. On the other hand, some participants might not have understood certain drawbacks of certain ideas, and these drawbacks might change their judgment. In conclusion, this underlines the importance of a concept selection method to guide the choice of ideas to be developed further.

Two of the participants expressed concern that not many ideas were novel, although all the methods were rated fairly highly concerning the participants’ apprehension of the number of ideas that were new to them, see Figure 5.1. An idea may be new to the individual who conceived it, to humanity, or to some subset of humanity greater than one (Nickerson, 1999), e.g. the ideating group. Some of the suggested ideas had been thought of before the ideation session by the participants or had been noticed in another context. Such ideas had been encouraged by the fact that the participants had got to know the ideation topic beforehand. The participants had already been working with some of the ideas that were presented. The proportion of these categories of ideas is probably very different depending on whether the ideation topic is a new or old issue. Old issues have already been the subject of much deliberation, and ideas have been conceived as to how one can solve the problems which they consist of. However, if an old issue has still not been solved, an ideation session is a good opportunity to share ideas and discover a new angle on one’s own ideas. If this is the case, it might be especially appropriate to use methods that push the participants outside of their normal thinking paths and involve people with different perspectives in order to find new ideas. The remarks stating that not many new ideas had been presented might also highlight the fact that the “old” ideas had not yet been discussed and scrutinized thoroughly, and therefore would reappear until they had finally been discarded or accepted for further development. The author suggests that this has been a neglected topic in the research on ideation.
methods. Sarkar and Chakrabarti (2009, 2014) offer a framework for exploring this topic through their definition of twelve different types of searches for ideas. According to their definition, a new search in the solution phase is activated when a designer comes up with an idea that was not previously known to the designer, i.e. was not derived from any knowledge base available or known to the designer. In an experiment letting novice and experienced designers work on two different problems individually, without the use of any design methods, it was found that only 15 out of 814 searches in the solution phase were categorized as new (Sarkar and Chakrabarti, 2014). In another experiment (Sarkar and Chakrabarti, 2009), groups of three members worked on two different problems using brainstorming, functional analysis, ideal design or an innovation situation questionnaire. In this experiment, it was found that with brainstorming, about 10% of the total number of search solutions were new solutions, whereas, when the other methods were used, on average 1% or less of the total number of solutions were new. According to these findings, the generation of new ideas is very rare, but can be enhanced by an appropriate ideation method. The distribution of known and novel ideas has typically not been taken into account when evaluating ideation outcomes, possibly due to the complexity of the issue. There is quite a difference between listing already known ideas and aiming to encourage creativity in such a way that novel and potentially radical ideas arise.

5.3.3 Behaviour of participants

Table 5.20 shows the distribution of spoken words between the participants during the verbally interactive group time of the gallery and SIL workshops. The total number of interpretable spoken words during the application of the gallery method and the SIL method was 3,781 and 7,392, respectively.

Table 5.20 Distribution of spoken words between the participants (%) during the verbal interaction steps of the gallery and SIL workshops.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>E</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>F</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Leader</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

The speaking time was unevenly distributed, since during both sessions almost half of the spoken words were uttered by two participants. According to Gibson (2010), researchers within the field of small-group dynamics assume that some
individuals speak infrequently on the grounds that others are thought to have more to contribute to a task that everyone wants to see performed well. Other reasons might be that participants might choose not to share their ideas for different reasons, or they might not get anything said because of group communication structures and personal low “verbal latency” (i.e. the ability to “jump in” as an opportunity appears in a discussion, see e.g. Burke (1974)). During interviews the participants mentioned the problem of low verbal latency as the reason for the unequal distribution of words. Group members have different personalities, and while some persons are more outgoing and talkative (more extrovert), others tend to be more reserved (more introvert), and therefore some group members are more likely to dominate the interactions in a group. Other possible reasons for the uneven speaking distribution could be the asymmetric dependence of different actors on each other and that an informal hierarchy exists among the participants. One negative consequence of such group behaviour is that it can hinder access to the full range of skills and resources, and the content of a meeting may be governed by those who speak the most. Unequal participation might also have consequences for post-encounter behaviour, e.g. compliance with decisions (Gibson, 2010). When the members of a team possess diverse sets of knowledge, less information is shared within it, and therefore it is particularly important to implement procedures that structure discussion, to enhance team sharing and consequently performance (Mesmer-Magnus and DeChurch, 2009). It is therefore believed that an unequal distribution of spoken words is a drawback and should be mitigated to release the full potential of the different knowledge and experience represented in a cross-functional group.

The distribution of spoken words between the activities during the verbally interactive time of the gallery method and the SIL method is shown in Table 5.21.

Table 5.21 Distribution of spoken words between activities (%) during the verbal interaction steps of the gallery and SIL workshops.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea statement</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Elaboration</td>
<td>80</td>
<td>75</td>
</tr>
<tr>
<td>Criticism</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Direction</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Going off at a tangent</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

It turns out that most of the spoken words were dedicated to the elaboration of ideas, whereas only a minor part of them was dedicated to stating an idea. In the gallery method, it was in the verbally interactive stage that the participants, according to instructions, presented their suggestions from the individual ideation and gave feedback concerning other participants’ suggestions, and therefore
it is not surprising that not very many of the words were used for idea statement. However, during the SIL method, the suggestions were generated during the verbal interaction time and still only 5% of the words spoken were used to state an idea. Jackson and Poole (2003) suggested that group members are not creating a list of ideas as much as they are constructing shared meaning about each idea and recording it publicly. As shown in Table 5.21, criticism was rare. Accordingly, the participants reported that they thought that the group was not criticizing ideas and that they could express whatever came into their mind. One reason for the fact that little criticism occurred might be that the group had been informed during the presentation of ideation methods at the start of the project that it is typically recommended to defer criticism to later stages of the conceptual design phase.

The participants of the group represented different organizations and it could be observed, in agreement with the participants’ own views, that they sometimes took the opportunity – consciously or unconsciously – to promote things that their organization provided or considered developing. This concerned not only the industrial actors, but also the academics, e.g. in talking about suggestions for work that academia could perform or prototypes that could be built and would benefit their projects. These attempts made by participants to promote their own organizations did not dominate the ideation sessions and often there was the potential of a win-win situation, but this is something that must be taken into consideration when ideating in real-life settings.

One of the reasons for choosing the three methods tested in Stage 2 of the present research was that they were expected to be rather simple to perform. In this respect, it turned out that Method 635 and the gallery method were easy for the participants to understand and they found the instructions for these methods easy to follow. In contrast, the participants had problems grasping the SIL method. They did not understand the instructions, had many questions about what to do, and seemed quite confused, as is exemplified in the excerpts from the transcript records in Table 5.22. Moreover, the original instructions by VanGundy (1984) about integrating all the generated ideas into the same solution had to be abandoned, and instead the participants tried to combine two suggestions at a time; and when a new suggestion was presented, the participants reviewed previous suggestions to check if they could find something appropriate to combine it with. The participants also reviewed the remaining suggestions that they had come up with individually to see if something could fit, rather than trying to combine any two suggestions. In the end this worked quite well, but one cannot maintain that the method was easy to perform at the first attempt.

The version of Method 635 used in the present study allowed the participants to ask the neighbour sending them his sheet of paper to explain anything they did not understand. This is not allowed in C-sketch, for example, in order to encourage misunderstandings that can lead to novel ideas (Shah et al., 2001). However, in the present study the participants never asked their neighbour for an explanation when performing Method 635.
Table 5.22 Excerpts from the transcript records showing how participants were confused about using or reluctant to use the SIL method. L denotes the author of this thesis.

<table>
<thead>
<tr>
<th>No.</th>
<th>Excerpt</th>
</tr>
</thead>
</table>
| 1   | L: Then I would like you first of all to present an idea.  
E: Okay.  
L: And then F presents an idea. Then together [you] try to combine these two ideas. You are welcome to use the whiteboard.  
E: Shall I start?  
F: Aha, okay, so then one should see if it can be combined with…  
L: [E’s] idea.  
E: How long shall I talk about my idea? Or my two [ideas].  
L: It’s mostly, it’s more on a conceptual level.  
E: Shall I present one or two, I had two.  
L: You should start with one […]  
E: [Presents his first idea.]  
L: Okay, that can be added to.  
E: Then I only had one [more] idea.  
L: Okay, there will be another round.  
E: Aha, only one idea at a time, okay. |
| 2   | L: Do we see any contradiction? Or is it possible to combine those two?  
E: No, everyone can discuss. Everyone’s free to speak.  
L: But they’re like two separate ideas in their own right.  
A: They’re two different things.  
L: Yes, so they can actually be combined, but the point is that they are to solve the same thing.  
E: But maybe they can’t be put together. |
| 3   | F: May I just ask, this method, is it a question of two [participants] presenting their ideas and then the others giving feedback, or […]?  
A: It’s supposed to be based on […]  
L: It’s a question of jointly, as it were, it’s not so much a question of feedback and criticism, but rather of trying jointly to combine [the ideas] to something that can be useful. |
The qualitative analysis of the transcript records showed that the elaboration of ideas was performed partly to make sure that all the participants understood them, which is a finding that might be typical of cross-functional inter-organizational groups dealing with technical and other complex problems and their solutions. Through relevant discussions, the participants were directed towards thinking of variations of their ideas and in this way concepts evolved. Four different ways in which the diversity of the group manifested itself were found in the qualitative analysis of transcript records: a) objections to other participants’ ideas, b) associations inspired by other participants’ ideas, c) former experience of similar solutions and d) immediate feedback to questions.

Table 5.23 shows excerpts from the transcript records where participants made objections to other participants’ suggestions. These objections in several cases led the group to explore how objections could be addressed by altering suggestions.
Table 5.23 Excerpts from the transcript records of verbal interaction showing examples of participants raising objections to other participants’ suggestions.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SIL</td>
<td>F: […] those solutions, what is attractive about them is that they don’t cost anything other than a little programming work. D: […] the latter yes, the first is about safety. F: What did you say? A: Then you have to remake all the point machines.</td>
</tr>
<tr>
<td>2</td>
<td>SIL</td>
<td>B: And I imagine a few wings and things within the turnout precisely to make that interplay with the vehicle […] and try to arrange for there not to be any lee-sides inside the turnout, because the trains actually come from both directions, so that you in some way can lead the snow out through wind power and let it carry the snow out. D: Now, there is a difference between single track and double track, if you think about double track, then you have trains coming from the same direction, is that really an advantage? F: It would be good if it could be constructed in such a way that it worked in all directions, so that it got cheaper in some way maybe.</td>
</tr>
<tr>
<td>3</td>
<td>SIL</td>
<td>F: Yes. Or that [the turnout] is high enough like that, then you can manage to plough properly with a deeper plough or something that is put down or, or something. […] D: I don’t know if I should be critical but […], what you should think about is that we pass both a crossing and a check rail, so that the height of your fixed installation on the train must be about a decimetre up, it is only those trains where you can lower and raise the plough with a driver that you dare to drive through a turnout at full speed. If you are to trust some automatic function that raises it before the check rail, then I think we will smash things to pieces. F: Yes, but it’s a bit the same with that shutter there, you have to be able to rely on it opening. D: Yes, exactly. I believe more in compressed air, in thinking along such lines rather than [a] pure mechanical [solution]. F: Or, [a] mechanical [solution], then one could have, so to speak, things that are elastic, like these rubber flaps that we talked about, things that endure contact with the turnout.</td>
</tr>
</tbody>
</table>

Elaborating ideas can lead to new ideas that would otherwise not have been obtained, in accordance with the research by Pelled et al. (1999) and Seidel and Fixson (2013), suggesting that debating ideas results in a more thorough exploration of the solution space. The elaboration of ideas may also save the group time at a later stage of the project. Table 5.24 shows examples of how participants were inspired by other participants’ suggestions and either elaborate them or create other suggestions.
Table 5.24 Excerpts from the transcript records of verbal interaction showing examples of participants using other participants’ suggestions as a source of inspiration to create an improvement of a suggestion or another suggestion.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gallery</td>
<td>A: [This is something that] nails my attention again and again, this variant of an air cushion, I think the automotive industry has really developed in that regard. If you could imagine a split sleeper […], like an air cushion part there that could be split in the middle and each side can be adjusted from side to side using air pressure. F: [OK, so it] can do the track adjustment with air. A: Yes. […] F: But let us say that all the sleepers in the entire turnout had such cushions and then you had a measurement system that measured all the time […]. A: Self-adjusting. F: Self-adjusting. So you always have a perfect track position.</td>
</tr>
<tr>
<td>2</td>
<td>SIL</td>
<td>F: So then you have a small buffer for 40 cm of snow lumps or newly-falling snow, but you still need some kind of heating system that removes that so that the buffer will be cleared, so that new snow can fall into that hole, or under […], but maybe it’s expensive to build it like that too. D: A thought comes to me when you say that, h’m, our point machines have a stroke length from 170, 120 and 90 mm, I mean, why do we allow a stroke length of 170 mm if we really want to avoid snow lumps that are too big in the turnout, I mean, maybe one should say that 120 is enough, we have, now please correct me, about 60 mm, 59 mm as the minimum requirement. B: Yes. C: And there are turnouts with 94-94, that is to say the first and the second point machine have the same stroke length, isn’t that so? B: [Model] 1 to 15. C (addressing D): But that’s quite right […], I mean the stroke length is not chosen based on the snow lumps, how big they are allowed to be or not to be […] D: But if you say 94, that means that the snow lumps that come in can only be 94, because otherwise they won’t even get down. C: Yes. D: And as a result, an increase in height of 1 dm, I mean, would have a rather big effect. For I believe in the actual basic idea there, if we can just raise the rail, then everything lying there falls down, if you press it together and crush it there and then let it up again and then a little heat on the sides and then, well, it flows downwards.</td>
</tr>
</tbody>
</table>
The diverse experience of the participants was also evident in comments where participants talked about previously proposed similar solutions (as exemplified in Table 5.25) which the other participants did not know of. This experience was in several cases exploited for the benefit of the current group interaction, for example by adapting solutions or by discarding suggestions as they had already been proven not to work.

Table 5.25 Excerpts from the transcript records of verbal interaction showing examples of participants relying on former experience of similar solutions to comment on a suggestion.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gallery</td>
<td>E: For if you have a joint between each part, then you kind of get a bump over it. A: Yes, exactly. E: A hinge there. C: In a sense there are beams like that in the City Tunnel. The problem is that they are probably not long enough. They are too short. So you only move that really stiff transition […]</td>
</tr>
<tr>
<td>2</td>
<td>SIL</td>
<td>C: And I have seen a railway built through a desert, [and there] you have the same problem, of course, but with sand.</td>
</tr>
<tr>
<td>3</td>
<td>SIL</td>
<td>E: […] I’m a bit fascinated by this simple idea of a shed. That is to say a cold space […] with a, now I’m not taking the overhead line into account. But in principle, I mean, that you have such a cold space here and here you come in and on the outside then the turnout has done its job. Yes. And so at least there’s no snow coming in. No heating. No maintenance costs. It opens up for a few, that we can add a few air fans and whatever. That can’t be so terribly expensive. D: [A mine operator] built their own terminal and to have as little snow as possible in the ore they unload, they have built a giant tent with, it’s only made of tarpaulins. I mean, a steel frame and tarpaulins, not like we do it with wooden walls and a steel roof and so on. So there has to be a cheaper solution to building one of those and then you can take it away during the summer, you can take down the tarpaulins again. E: Exactly, and then the overhead line can run under it […] D: They have built it over the overhead line […]</td>
</tr>
</tbody>
</table>

The diverse knowledge and different backgrounds of the group members made it possible for a participant quickly to obtain an answer to anything he was wondering about in relation to a suggestion; examples of this are shown in Table 5.26. It is not a straightforward task to measure quantitatively the degree to which the different backgrounds of the participants helped the evolution of ideas, but it is obvious that the more knowledge and experience there is gathered in the group, the higher is the probability that someone will know something that can help improve another participant’s suggestion.
Table 5.26 Excerpts from the transcript records of verbal interaction showing examples of participants receiving immediate feedback concerning questions.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gallery</td>
<td>E: What I’m wondering about is exactly that, if it’s the case that, that it’s too stiff in a tunnel or if it’s too pliant in the normal track. D: In my opinion it’s too stiff in the tunnel. E: Okay, h’m. But then it’s an excellent idea. Simple.</td>
</tr>
<tr>
<td>2</td>
<td>SIL</td>
<td>F: [And …] the next thing I thought about in connection with the train was, it was […] this plough which could be adjusted and then create a gust of air, but, that it also has such a shape, now it is positioned, I don’t know how high above the rail the plough is positioned, does anyone know? […] A: The fixed ploughs on the engines? F: Yes, a bit up? A: Yes. D: Almost a decimetre on the engines, but then the engine [designed for snow removal is closer to the] rail, it’s only a few centimetres, there’s a piece cut out for the rail heads and suchlike, but it’s special for snow clearance, of course.</td>
</tr>
<tr>
<td>3</td>
<td>SIL</td>
<td>F: How long a section does one say needs to be covered, I mean, how long is it? E: Yes, it should be the entire blade. D: The length of the installation is 10-15 m in these longer turnouts. F: [There should be] some cleats further back too.</td>
</tr>
</tbody>
</table>

Other situations in which the different backgrounds of the participants played a role were the group-analysis of the topics and finding out the causes of problems. It is beyond the scope of the present thesis to study this process in detail, but it was observed that the participants contributed different perspectives on the topics, which resulted in quite a comprehensive group-analysis of the topics in a short time. The identified causes were used as inspiration during the ideation sessions and this resembles the way in which an inspirational poster was used in a study conducted by Rexfelt et al. (2011), where the poster was presented to a brainstorming group but not to a Method 635 group. Rexfelt et al. (2011) reported that the Method 635 group in that study struggled to come up with suitable ideas, and it is interesting to note that similar observations were not made during the test of Method 635 in the current study. However, there may be other factors which explain or contribute to the difference in these observations between the present study and the study by Rexfelt et al. (2011), such as the personalities of the group members. In this connection, an interesting area for future research is how the scope of the group-analysis of topics is related to the outcome of the ideation.
5.4 Conclusions

The participants enjoyed working in a cross-functional inter-organizational group. They all thought that the group possessed a large amount of competence and experience and found the different and complementary knowledge about the topics that the different participants possessed with regard to the ideation topics to be very valuable, both during the group-analysis of the topic that took place before ideation and during the generation of ideas on how to address the ideation topic.

The academics preferred using Method 635, while the gallery method on average was most popular among the railway professionals in the group. Overall, the SIL method was least popular among the participants.

Method 635 was found to have a considerably higher idea and concept generation rate than the gallery and SIL methods. The quality of the generated concepts, in terms of their ability and viability, did not vary significantly between the methods. The methods exhibited different characteristics with regard to the distribution of sketched details, written words and verbal words in the concept descriptions.

The speaking distribution between the participants was uneven during the verbally interactive steps. Most of the spoken words were dedicated to the elaboration of ideas, and only a minor part to stating an idea. Criticism was rare. Method 635 and the gallery method were easy for the participants to follow. In contrast, the participants struggled to grasp the SIL method and were averse towards the way in which it was supposed to be performed. During the meetings, the participants in the group sometimes took the opportunity to promote the organizations which they represented.

During the verbally interactive steps, four different ways were found in which the diversity of the group manifested itself: a) objections to other participants’ ideas, b) associations inspired by other participants’ ideas, c) former experience of similar solutions, and d) immediate feedback to questions.
Appendix 5.1

**Instructions**

For each statement, please indicate with a cross how well you think the statement agrees with your opinion.

**Example**

I like turnouts.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I am engaged in the topic.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

If I had had more time, I would have come up with more ideas.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I had more ideas than I presented.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

Many interesting ideas were presented during the workshop.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

Many ideas were new to me.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I think ideas that were presented can be used in practice.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I felt engaged during the workshop.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>
I am satisfied with my own contribution during the workshop.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I think the method is useful.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I kept ideas to myself during the workshop due to the presence of competitors.

<table>
<thead>
<tr>
<th>Do not agree at all</th>
<th>Agree</th>
<th>Completely agree</th>
</tr>
</thead>
</table>

I like the following things about the method:

The following are things which I do not think worked well when using the method:

Other comments:
Appendix 5.2

Interview guide

Ask the participant to give his judgement of the methods and explain the reasons for his judgement with respect to the statements in the questionnaire.

Ask the participant to explain further his free-text comments on the methods.

Ask the participant to motivate his ranking of the methods.

Additional questions:

- Are the ideas unique? Do ideas come from other forums?
- How important is it to share ideas compared to coming up with completely new ideas?
- Is it difficult to disregard existing regulations during the idea generation?
- How have the participants’ different backgrounds affected the project?
  - Resource?
  - Disadvantage?
- Can the methods be implemented in your own organization?
JUDGMENT OF CONCEPTS: PART 1

INSTRUCTIONS

Ideaion topic
How can deterioration of the track geometry be prevented?

VIABILITY (TECHNIQUE, COST, SAFETY)
The viability of a concept is judged with respect to three different categories:
- principle: technical viability,
- economy: whether or not the concept can be realized at a reasonable cost, in consideration of the circumstances,
- safety: whether or not the concept can be realized with high enough safety for the personnel and the environment.
The viability is judged on a scale from 0-1, where:
0 = not at all possible,
1 = completely possible.

ABILITY TO SOLVE THE PROBLEM
Judging a concept’s “ability to solve the problem” concerns judging the degree to which the concept solves the given ideation topic.
The “ability to solve the problem” is judged on a scale from 0-1, where:
0 = not at all possible,
1 = completely possible.

If a concept cannot be judged with respect to its viability or ability to solve the problem, tick the “Cannot judge” box (for example, if you feel that you do not have enough knowledge to make a judgement, that the criteria are not applicable to the concept, or that the concept is incomprehensible). Write comments, if any, beside the concept.
END OF INSTRUCTIONS. THE EXAMINATION STARTS ON THE NEXT PAGE.

Example

Certain concepts are used as main concepts and should therefore be identified by underlining the main concepts and the additional concepts.

CONTINUATION OF INSTRUCTIONS

Reading rules that can also be found on the exam format are noted.
<table>
<thead>
<tr>
<th>Ideation topic: How can deterioration of the track geometry be prevented?</th>
<th>Ability to solve the problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viability (technique, cost, safety)</td>
<td>0 = not at all possible</td>
</tr>
<tr>
<td>0 = not at all possible</td>
<td>1 = completely able</td>
</tr>
<tr>
<td>1 = completely possible</td>
<td>0 = not at all able</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Viability (technique, cost, safety)</th>
<th>Ability to solve the problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Decrease the axle load</td>
<td>0 = not at all possible</td>
</tr>
<tr>
<td>0 = not at all possible</td>
<td>1 = completely able</td>
</tr>
<tr>
<td>1 = completely possible</td>
<td>0 = not at all able</td>
</tr>
<tr>
<td>Cannot judge</td>
<td>Cannot judge</td>
</tr>
</tbody>
</table>

| 2. Decrease the number of passengers | 0 = not at all possible |
| 0 = not at all possible | 1 = completely able |
| 1 = completely possible | 0 = not at all able |
| Cannot judge                | Cannot judge |

| 3. Good and large contact area between ballast and sleeper. | 0 = not at all possible |
| 0 = not at all possible | 1 = completely able |
| 1 = completely possible | 0 = not at all able |
| Cannot judge                | Cannot judge |
6. Development of ideation method

6.1 Introduction

This chapter presents Stage 3 of the research performed for this thesis (see Table 1.1). The first objective of the work carried out during this stage, based on the findings from Stage 2 (Chapter 5) and relevant findings from the literature, was to develop a practical ideation method which would be tailor-made for cross-functional inter-organizational teams in the context of the Swedish deregulated railway market. The second objective was to apply and evaluate the different versions of the method in the OptiKrea Group and compare them to the methods which they were based on (which had been tested in Stage 2), with respect to the participants' views, the ideation outcome and the behaviour of the participants. The third objective was to analyse how the individual ideation steps and group review steps used in the methods differed.

The following research questions were posed to meet the objectives.

- What are the goals of and the requirements for such an ideation method?
- What did the participants think about the developed method?
- How did the application of the method affect the ideation outcome compared to the methods tested in Chapter 5?
- How did the participants behave during the application of the different versions of the method?
- How do the individual ideation and group review steps differ during the application of the methods?
- What are the general design principles for such an ideation method?
- Is the developed method better than the methods which it was based on?

6.2 Method

6.2.1 Procedure

The goals of and the requirements for the method to be developed emerged continuously during Stage 1 and Stage 2 and at the start of Stage 3 (see Table 1.1) of the OptiKrea project. Comments relevant to the goals and requirements
were noted down at meetings and workshops of the OptiKrea Group. These data were continuously considered by the “scientific team” (see Section 3.3.1) in relation to the findings that were made during Stage 1 and 2, and the goals and requirements regarding the method were shaped during this process. The shaping of the goals and requirements was also influenced by the experience which the senior researcher in the “scientific team” had reaped from using conceptual design methods in and developing conceptual design methods for different types of teams during approximately 20 years. The final goals and requirements that emerged as a result were presented to and approved by the participants of the OptiKrea Group. The goals and requirements were then used to direct the development of the method.

The two versions of the developed method were tested by the OptiKrea Group with an interval of four weeks between the tests. Between the two tests of the developed method, the participants worked with their normal tasks and did not use any ideation methods. The different ideation topic for each version of the ideation method was chosen using the same procedure as was applied during the tests of established ideation methods, as described in Section 5.2.1. The chosen ideation topics are presented in Table 6.1. The tests of the two versions of the ideation method were performed in the same way as the tests of the established ideation methods. This procedure is described in Section 5.2.1. The author of this thesis presented the instructions for each version of the developed method to the group by means of a projector, ran a timer to keep track of the time, announced when the group should move on to the next stage of the method and answered questions about the instructions during the session. In this capacity the author is referred to as the “supervisor” throughout this chapter.

<table>
<thead>
<tr>
<th>Table 6.1</th>
<th>Ideation topics used for the tests of the two versions of the developed method.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Topic</td>
</tr>
<tr>
<td>OKMv1</td>
<td>How can good track geometry of a modular turnout be achieved and retained?</td>
</tr>
<tr>
<td>OKMv2</td>
<td>How can satisfactory drainage of the track superstructure of turnouts be achieved?</td>
</tr>
</tbody>
</table>

### 6.2.2 Views of the participants

The views of the participants regarding the developed method were assessed by applying a questionnaire and through interviews. Both after the test of version 1 of the developed method and after the test of version 2, the participants answered a questionnaire and a short group interview regarding the participants’ experience of the applied method took place to capture the participants’ immediate reactions to the ideation method that they had tried. The questionnaire was the same as that used after the test of the established ideation methods and is described in Section 5.2.2 and shown in Appendix 5.1.
A linear regression was performed between each pair of statements in the questionnaires to find out if any correlation existed. The questionnaire results for all the five methods tested by the OptiKrea Group during the project, both the established ideation methods and the two versions of the developed method, were included, resulting in 30 data points.

The participants’ average levels of agreement with each statement in the questionnaire were used to find a total quantitative score for the participants’ views on the methods. The first statement, “I am engaged in the topic”, was removed from this analysis since it is not dependent on the method. Two statements, “If I had had more time, I would have come up with more ideas” and “I had more ideas than I presented”, were inverted (the average level of agreement was subtracted from 1). This was because the tested methods were considered to be better the lower this value was with respect to these statements, while the opposite was the case with respect to the other statements in the questionnaire. The statements were considered to be of equal importance (have the same weight) and their values were added together to find the overall score of each method.

Appendix 6.1 shows the questions that were used during the group interview taking place after testing OKMv1. After the test of OKMv2, the group members were asked to answer a number of questions in written form, and these questions are shown in Appendix 6.2. These questions were used during the group interview that followed. The group interviews were audio-recorded and analysed in the same way as described in Section 5.2.2 and Table 5.5.

6.2.3 Ideation outcome

The execution of the established methods which were tested in Stage 2 and to which the two versions of the developed method will be compared in terms of outcome is presented in Table 5.3.

The number of concepts suggested by the group as possible solutions to each ideation topic was counted, as well as the total number of non-redundant ideas and the number of ideas contained in each concept. The same procedure and rules for counting concepts and ideas were used as when testing the established ideation methods, see Section 5.2.3.1.

The two versions of the developed method and the established methods which they were based on were divided into individual phases (IPs), where the group members worked on their own, and group phases (GPs), where the group members interacted verbally. An analysis was made of how the number of non-redundant concepts and ideas varied over the different phases and between the methods. Further, the number of ideas reused in a later phase of the versions of the developed method was counted.
6.2.4 Behaviour of the participants

The ideation sessions were audio-recorded, and these recordings served as a basis for analysing the behaviour of the participants. To analyse the content of the ideation sessions, the same procedure was used as during the tests of the established ideation methods and is described in Section 5.2.4. Being present at the workshops allowed the author to act as an observer, although this role was combined with the role of supervising the workshops and not systematically fulfilled throughout the workshops. However, if observations were made of the participants behaving unexpectedly or in a manner that was interesting in some other respect, they were noted down. The notes were used to support the analysis of the recordings. The recordings were, however, the main source of the findings regarding the behaviour of the participants.

On the basis of the transcribed recordings, an analysis was made of the distribution, between the different activities and between the participants, of the words spoken by the participants during the ideation sessions. The analysis was performed using the same procedure as for the established ideation methods, see Section 5.2.4. If the speaking distribution had been equal between the participants, each participant (in the group of six participants) would have spoken an ideal fraction corresponding to one sixth of the number of words remaining after subtracting the number of words uttered by the supervisor from the total number of words spoken. The ideal fraction was compared to the actual fraction and the average of the absolute values for how much the participants deviated from the ideal fraction was calculated. The “equality of speech”, $ES$, during the GPs of each method was thus calculated by

$$ES = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{N_i}{N} - \frac{1}{n} \right|$$  

Equation 6.1

where $n$ is the number of participants, $N_i$ is the number of words uttered by participant $i$ and $N$ is the total number of words uttered by all of the participants (i.e. after the words uttered by the supervisor have been subtracted). The first term within the absolute value bars is the actual fraction of words spoken by participant $i$ and the second term is the ideal fraction. A corresponding calculation was performed for the distribution of the speaking frequency between the participants, i.e. how many times each participant spoke during a session.

An overview of the areas of focus during the evaluation, the sources of evidence and the methods applied to analyse the data is presented in Table 6.2.
Table 6.2 Parameters that were analysed during the evaluation of the tests of the idea-
tion methods.

<table>
<thead>
<tr>
<th>Areas of focus</th>
<th>Sources of evidence</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>The participants’ views</td>
<td>Questionnaires</td>
<td>Quantitative analysis of the average value of each questionnaire statement</td>
</tr>
<tr>
<td></td>
<td>Transcribed audio recordings of group</td>
<td>Quantitative analysis of the correlations between pairs of statements in the questionnaires</td>
</tr>
<tr>
<td></td>
<td>interviews</td>
<td>Calculation of total method score from the questionnaire results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thematic analysis</td>
</tr>
<tr>
<td>Outcome of the ideation</td>
<td>Sheets of paper from the workshop</td>
<td>Quantitative analysis (during different phases of the methods) of the following:</td>
</tr>
<tr>
<td>sessions</td>
<td>Transcribed audio recordings of sessions</td>
<td>the number of non-redundant concepts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the number of non-redundant ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the number of ideas reused in later steps</td>
</tr>
<tr>
<td>Behaviour of the group</td>
<td>Transcribed audio recordings of sessions</td>
<td>Coding of the following:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the distribution of spoken words between different activities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the distribution of spoken words between participants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Calculation of the average deviation from an equal distribution of the spoken words and speaking frequency between the participants</td>
</tr>
</tbody>
</table>

6.3 Results and discussion

Section 6.3.1 presents how the initial design of the method was reached and describes what changes were made to the method after the first test in the OptiKrea Group. The changes were introduced to meet needs that were identified during the evaluation of the initial design. The remainder of Section 6.3 is dedicated to a detailed presentation and discussion of the results from the evaluation of the tests of the different versions of the developed method, and a comparison of these results with those from the test of the established methods (see Chapter 5) which the developed method is based upon.

6.3.1 Design of the method

6.3.1.1 Goals and requirements set for the method

Two overall goals emerged for the method to be developed: a) the method should produce a great number of ideas with breadth and potential, which was broken down into the requirements “must generate a great number of ideas” and
“must generate ideas with breadth and potential” and b) the method should be implemented by STA after the project finishes. The prerequisites for an implementation were found to be that the intended users must feel that the method is easy to initiate (a low usage barrier) and that the inclusion of different participants (from different organizations) on different ideation occasions must be easy. It emerged that, to meet these prerequisites, the method must fulfil the requirements “must be easy to use and understand” and “must be attractive to the users”. Despite previous findings that facilitators (e.g. Kramer et al., 2001) and training (Parnes and Meadow, 1959) can enhance the productivity of ideation, both of these alternatives were ruled out, because they would raise the usage barrier too much. Partly for the same reason, special equipment or software was ruled out. Another reason for excluding software was that organizations typically (e.g. in the present case) do not allow external software to be downloaded onto their employees’ computers. In addition, computational tools for the conceptual stage do not yet provide the same usability, the same ease of annotation of design sketches with rough dimensions, and the same possibility of making notes as do pen and paper (Yang, 2009). Emerging technology might, however, solve these issues.

The results from the tests of the established ideation methods showed that during the application of the SIL method the participants were quite confused about what they were supposed to do and asked a large number of questions about the procedure (see Chapter 5). It was therefore concluded that the SIL method would not be feasible to explore further since it was found that it would be impossible to achieve the goal of easy implementation with that method; therefore, the SIL method was excluded from the scope of Stage 3. Neither Method 635 nor the gallery method encountered similar problems and were therefore judged, with respect to how easy it would be to use the methods, to be a better starting point for the method to be developed.

6.3.1.2 Individual and group work

Drawing upon the implications of associative memory models (AMMs – see Section 2.2.3), it was determined that a feasible start for the method would be a first step of individual ideation where the participants would document the ideas which they would first come up with as a reaction to the ideation topic. In this way, they would not be affected by the other participants’ ideas initially. After the first individual step, AMMs suggest that the participants should be exposed to other ideas than their own to activate not-yet-accessed parts of their memory network. Participants can be exposed to ideas generated by others in the group or external ideas such as pictures of objects from other fields or physical objects from nature. Previous research suggests that the effectiveness of stimuli depends on the appropriateness of the stimuli with respect to the given topic, participants and context (Perttula and Sipilä, 2007). It was decided that the group would use internally generated ideas during the ideation method because a) it
was expected that these ideas would have a higher probability of appropriateness than selected external stimuli, b) someone would have had to make the effort of selecting the external stimuli before the ideation session, which would have been a barrier to using the method, and c) the different functional backgrounds of the participants were thought to provide ideas with enough variety.

6.3.1.3 Sharing of ideas

There are several ways of exchanging ideas that can activate the memory network in different ways. The ideation method to be developed did not have strict restrictions regarding time usage. Therefore, it was decided that there would be sufficient time to allow a mix of idea-exchanging strategies as a starting point. Linsey et al. (2011) suggested that “an improved process for idea generation consists of first using a gallery communication method to generate a large number of high quality product solutions and then moving to a rotational viewing method using words and sketches to develop the details of the product solutions and a large number of functional ideas”. However, in the light of AMMs, one should first aim to achieve a large pool of ideas with great variety that can activate different parts of the memory network. Therefore, the two versions of the developed method start with an individual ideation phase with rotational viewing followed by an individual ideation phase with gallery viewing. After each of the individual ideation phases, the participants exchange ideas through presentation and discussion in verbally interactive group phases. The participants in the OptiKrea Group thought that presenting and discussing ideas were very important for reaching good solutions, and this feature of the method made it particularly attractive, since the participants really enjoyed interacting verbally (see Chapter 5). Further, the presentation steps let each participant gain an overview of the entire set of ideas generated by the group. One difference between interactive and nominal groups is the fact that, after the ideation session, the members of an interacting group will have knowledge about the entire set of generated ideas, whereas a member of a nominal group will only know about the subset of ideas that he or she has generated. In consideration of this, it is not very surprising that the participants in an interactive group typically feel more productive (Stroebe et al., 1992; Paulus et al., 1993) and enjoy the ideation session more (Furnham, 2000). It is reasonable to assume that the sharing of ideas is an important factor in motivating individuals to participate and engage in the ideation session. The procedure of the first version of the developed method (OKMv1) is shown in Table 6.3.
Table 6.3  The execution of the first version of the developed method (OKMv1), IP=individual phase, GP=group phase.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP1</td>
<td>1) Steps 1-5 of Method 635 (see Table 5.3).</td>
</tr>
</tbody>
</table>
| GP1   | 2) The participants use five minutes to read through the ideas that have been added to the sheet of paper which they started out with.  
3) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.  
4) After each presentation, the other participants give feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of ten minutes is allowed per sheet of paper (for presentation and feedback). |
| IP2   | 5) The sheets of paper from step 1 are put up on a wall or some other place where all the participants can easily view them.  
6) Each participant works individually for ten minutes to develop or combine ideas from the collection of ideas from step 1. New ideas are also welcome. New sheets of paper are used to document the ideas by means of sketches and/or text. |
| GP2   | 7) Each participant presents their ideas from step 6, both the ideas conceived by themselves alone and those inspired by others.  
8) After each presentation, the other participants give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of five minutes is allowed per participant (for presentation and feedback). |

It is acknowledged that GP2, the presentation and discussion of ideas from IP2, might not add very many new ideas, but rather is used to make the group familiar with the entire set of ideas and to conclude the session for the group. However, knowing that others will view the ideas which one has come up with might enhance the idea generation in the earlier steps through social comparison (Perttula et al., 2006). Further, it might be possible that this last step activates new thoughts among the participants, who may come up with ideas directly after the session, after a period of incubation, or later when interacting with some other stimuli.

6.3.1.4 Form of representation

According to Linsey et al. (2011), embodied cognition theories suggest that external representations such as sketches are helpful for the performance of design tasks since they reduce the cognitive load, as the amount of information which an individual has to represent internally is reduced. On the other hand, whereas information such as geometry and configuration tends to be easier to express in drawings, other information such as abstract concepts might be easier
to convey in words (Linsey et al., 2011). Additional information in words might also be helpful when other individuals interpret a sketch, which saves time during both rotational and gallery viewing. Particularly worth noting is that individuals have different preferences and some individuals simply do not like to sketch and/or are bad at sketching. Therefore, the participants were instructed to use both sketches and text, to the extent desired by them, when documenting their ideas. An advantage of the developed method is that all the ideas are documented in sketches and text by the participants continuously during the workshop, so that no special person needs to be appointed to record ideas, which facilitates implementation. Table 6.4 sums up the preliminary design principles used to set up the initial version of the new method.

6.3.1.5 Length of each round of rotational viewing – adjustment made in OKMv2

Two changes were made to the initial design of the method after the first test. The time allocated for each round of IP1 was prolonged from five to ten minutes, as the participants all agreed that five minutes was too short a time to be able to document all their own ideas during the first round, and to read through and understand the ideas on the sheet of paper which they had received and then come up with new ideas during the remaining rounds. This is in line with Nijstad et al.’s (2002) reasoning that stimulus ideas can interfere cognitively with a person’s train of thought if they belong to a different semantic area of memory. The train of thought is then cut off, leading to a loss of potential ideas. Therefore, the time allocated for each round must be long enough to allow the participants to reach the end of their train of thoughts, as activated by exposure to the ideation topic and stimulus ideas. On the other hand, the time for each round must not be too long, as is further discussed in Section 6.3.4.

6.3.1.6 Distribution of spoken words – adjustment made in OKMv2

The second change concerned the unequal distribution of spoken words among the participants which was found after the first version of the developed method had been tested and which is dealt with in greater detail in Section 6.3.4. The distribution of spoken words is important, since previous research claims that it is essential that all the participants’ perspectives and knowledge are contributed in the group (Milliken et al., 2003). New information can activate new areas in the participants’ memory network, and to access and use the information which each participant possesses, it is an advantage if all the participants speak, although participants can also contribute their ideas and viewpoints during the IPs in the developed method. Some persons are inclined to speak more and others less (e.g. Burke, 1974), and therefore it was desirable to find a simple procedure
Table 6.4 Design principles for ideation methods to be used in cross-functional inter-organizational groups, their justification and how they were put into practice in the developed method.

<table>
<thead>
<tr>
<th>Design principle</th>
<th>Justification</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used to design OKMv1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developing the method based on methods proven to be easy to understand</td>
<td>To increase the probability of implementation of the method after the development. To make it easy to involve new participants.</td>
<td>Combines elements from Method 635 and the gallery method.</td>
</tr>
<tr>
<td>The initial generation of a pool of ideas with great variety</td>
<td>Many different types of ideas to create associations with.</td>
<td>The method starts with individual ideation without interaction or other stimuli.</td>
</tr>
<tr>
<td>Exposure to stimulus ideas generated internally within the group</td>
<td>To activate new areas of each participant’s memory network. High probability of appropriateness. No additional work.</td>
<td>Participants are exposed to ideas from the other participants, silently during the IPs and verbally during the GPs.</td>
</tr>
<tr>
<td>Mixing idea-exchanging strategies</td>
<td>Exchanging ideas in different ways may activate the memory network in different ways.</td>
<td>Combines rotational viewing and gallery viewing with verbally interactive steps.</td>
</tr>
<tr>
<td>The use of external representation in the form of sketches and text</td>
<td>To reduce the cognitive load. Sketch and text appropriate for different types of information. To facilitate the understanding of concepts. Personal preferences. Inherent documentation.</td>
<td>The participants themselves choose if they want to document their ideas with sketches or words.</td>
</tr>
<tr>
<td>Including time for discussion and debate</td>
<td>To increase the quantity and variety of ideas. Attractive to participants.</td>
<td>Individual phases are followed by a group phase.</td>
</tr>
<tr>
<td>Added after testing OKMv1 in the group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimizing the cycle time</td>
<td>Enough time for participants to exhaust their own ideas in the first round and to review and react to the stimulus ideas from other participants in the remaining rounds.</td>
<td>Time increased from five to ten minutes during each round of IP1.</td>
</tr>
<tr>
<td>Developing a strategy for distributing the verbal interaction</td>
<td>All perspectives and knowledge in the group should be considered.</td>
<td>Participants take turns to comment on other participants’ suggestions.</td>
</tr>
<tr>
<td>Added after reflections on the method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paying attention to stimulus ideas</td>
<td>Participants must attend to stimulus ideas to be inspired by them.</td>
<td>Rotational viewing during IP1.</td>
</tr>
</tbody>
</table>
that could even out the distribution of spoken words. Instead of letting the dis-
cussion be completely free, the change was introduced that, after each
participant had presented his suggestions, the other participants took turns to
comment on the presented ideas (the maximum time for each presentation and
feedback step is found in Table 6.3). Whenever questions and urgent comments
arose, the participants were still allowed to speak out of turn. The steps of the
second version, OKMv2, are shown in Table 6.5.

Table 6.5  The execution of the second version of the developed method (OKMv2),
IP=individual phase, GP=group phase.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP1</td>
<td>1) Step 1-5 of Method 635 (Table 5.3), but with ten minutes in every round instead of five minutes.</td>
</tr>
</tbody>
</table>
| GP1   | 2) The participants use five minutes to read through the ideas that have been added to the sheet of paper which they started out with.  
3) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.  
4) After each presentation, the other participants take turns to give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of ten minutes is allowed per sheet of paper (for presentation and feedback). |
| IP2   | 5) The sheets of paper from step 1 are put up on a wall or some other place where all the participants can easily view them.  
6) Each participant works individually for ten minutes to develop or combine ideas from the collection of ideas from step 1. New ideas are also welcome. New sheets of paper are used to document the ideas by means of sketches and/or text. |
| GP2   | 7) Each participant presents their ideas from step 6, both the ideas conceived by themselves alone and those inspired by others.  
8) After each presentation, the other participants take turns to give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of five minutes is allowed per participant (for presentation and feedback). |

6.3.2 Views of the participants

Section 6.3.2.1 and 6.3.2.2 present the results regarding the participants’ views on the versions of the developed ideation method and rely on data from the group interviews. Quotes (translated from Swedish into English) are included to exemplify qualitative data that contributed to the results. The letter in parenthesis after each quote indicates who is quoted, according to Table 3.1.
6.3.2.1 The participants’ views after testing OKMv1

All the participants thought that OKMv1 was the best of all the methods which they had tried thus far, i.e. Method 635, the gallery method, the SIL method and OKMv1. They liked combining elements from Method 635 and the gallery method, since they thought that this gave them the opportunity to exhaust the ideation topic more thoroughly than they had been able to do with the methods tested previously. This is exemplified by the following quote:

“Directly after this [Method] 635 we got to do this [gallery method]. […] On the same ideation topic, and then we could exhaust the topic more.” (F)

All the participants felt that they had progressed further regarding the complexity and details of the concepts and ideas, and that the concepts and ideas had evolved in more directions and in greater detail than when they had tried the established ideation methods, as is exemplified by the following quote:

“I do think I got further through being able to take that final gallery [step], wrap up what one had achieved, so to speak, so I experienced that I got further in the process.” (E)

All the participants agreed that five minutes in each round of IP1 was too short a time. They had not had enough time to document their own initial ideas in the first round. Four participants mentioned specifically that, when the next round had started, they had forgotten their own remaining ideas and had concentrated on making associations with the ideas on the paper which they had received, as is exemplified by the following quote:

“When I got [my neighbour’s sheet with ideas…], I sat wondering what I should do with them, but I had my own ideas, of course, and […] his ideas] were completely different from what I was thinking about. […] But then I lost [my own ideas before I got to write them down]. […] Now I was supposed to understand what he had [suggested…].” (D)

However, reading through and understanding these ideas had consumed most of the minutes that had been available, and therefore there had not been much time left to build on these ideas. This is exemplified by the following quote:

“Once you get the sheet from your neighbour you get a bit stressed, you don’t have the time to read through and understand all […] the innovations or ideas before you have to [add your own ideas], there is so little time left to try to come up with something new yourself.” (F)
Two participants mentioned specifically that they were surprised and impressed by the wealth of ideas that had emerged during the session, as is exemplified by the following quote:

“We chose between [ideation topics] A and B, and B had […] been so broad that the number […] of possibilities of generating ideas had been greater. […] So I wrote that I am still surprised that we generated so many ideas.” (D)

When asked if they would have preferred working in groups of two, for example, during IP2, all the participants agreed that they preferred working alone during the individual steps. All the participants thought that they could create ideas with more breadth by working individually, that working on their own saved time because working in pairs or teams required more time for explanation, and that there was enough group work during the presentation steps. Also mentioned was the fact that it was good to have time to develop one’s own ideas on one’s own. This is exemplified by the following quote:

“We have better possibilities of [covering a wide range of ideas] if we work individually in this third step, in my view. The generation of ideas and then the presentation of them, that is when you have the common discussion. Then the group is working. (D)

Two of the participants mentioned specifically that they had not been in a good creative mood on the day when OKMv1 had been tested, in one case because of a lack of sleep.

6.3.2.2 The participants’ views after testing OKMv2

All of the participants agreed that OKMv2 represented another step forward and was the best method that they had tried thus far. This is exemplified by the following quotes:

“I wrote that it is a step forward, that I think that this very fact that [we have more time] means that we can process, think about and grasp what people mean instead of just spurting out ideas. So I feel, as it were, that this is a method that, as I see it, could be put into practice.” (D)

“I agree, this was the best we accomplished, I believe. It felt completely right, all the way. [Sometimes] there were even some spare minutes, that was first-rate, I thought.” (C)

The participants all thought that prolonging the time for the rounds of rotational viewing was excellent, but had somewhat different opinions about how to distribute the time. One participant thought that it would be better if the first round
was longer than the others. The other participants, however, thought that ten minutes was appropriate for every round. All the participants thought that there was sufficient time to understand, consider and make associations with other people’s ideas, as well as to document one’s own ideas. This is exemplified by the following quote:

“No, ten [minutes] was excellent, it’s quite sufficient, because somehow I felt there was peace and quiet and we had the time to read through [the ideas] before starting to comment; [during OKMv1] we started to comment directly, as it were, one didn’t have the time.” (D)

Four of the participants mentioned that, during the last rounds of rotational viewing, they had actually had a few minutes left which they had thought could perhaps have been used for IP2, the second individual step, instead. None of the participants wanted to remove IP2, despite the fact that IP1 had been prolonged. They wanted to use IP2 to summarize or to develop concepts, or to be able to work on new ideas which they had conceived during GP1, as is exemplified by the following quote:

“I didn’t [work on] the [concept] I came up with first, those ideas, anyone of those […] I came up with first, but when we arrived at the gallery [step], then I was [working on] something else. Then I had let go of the first [ideas] I believed in.” (C)

Half of the participants thought that the time allocated for IP2 had been sufficient, whereas the other half felt that they could have used a few minutes more. Three participants mentioned specifically that they were surprised by the large amount of ideas which they had conceived regardless of whether they had been very familiar or less familiar with the ideation topic. This is exemplified by the following quote:

“Just to see visually what we [created], what we got down on paper, my goodness, we were never close to that any time before. And on a topic which I think, [although there is good knowledge of it within the group,] as for me, I have only scant knowledge of it, yes, well, incredibly good.” (C)

Letting all the participants sitting around the table speak in turn was appreciated, which is exemplified by the following quote:

“Really good, in my opinion.” (E)

One participant suggested that the developed method could be improved by having each group member think of three ideas on the topic in question before the ideation session and bring it to the session as a starting point for IP1. The
other participants, however, were reluctant to accept this suggestion, because they thought that it would be difficult to arrange a presentation and clarification of the ideation topic in advance of the session, that “crazy” ideas would disappear and that it would not be possible to achieve the quick responses and communication which the developed method made possible, as is exemplified by the following quote:

“[…] I think it removes crazy ideas. Then you have to be so serious all at once. Then I think one loses good ideas.” (B)

6.3.2.3 Questionnaire

The participants’ average levels of agreement with the statements in the questionnaire are shown in Figure 6.1, while the statements are shown in Table 6.6. The questionnaire results were also analysed separately for the academics and railway professionals. It turned out that the variation between all the individual participants was greater than that between the two subgroups, and therefore no conclusions about possible differences between the two subgroups could be established. The only statement for which the variation between the individuals was less than that between the subgroups was the statement “Many ideas were new to me”. The academics exhibited a higher level of agreement with this statement during the test of OKMv2, which is perhaps not surprising, as the other participants had a longer experience from the railway sector. Moreover, there was a trend that the professionals from STA agreed to a lower extent than the others with the statement “Many interesting ideas were presented during the workshop”.

Figure 6.1 shows that the engagement in all the ideation issues was quite high, although the engagement in the issue which was ideated on during OKMv1 was somewhat lower. This may be one explanation for the lower engagement felt during the OKMv1 ideation session compared to the engagement felt during the gallery method or OKMv2. However, when examining the individual ratings, it was found that it was the very low rating given by one of the two participants who had experienced a “bad day” which had pulled down the average level of engagement felt by the participants during the OKMv1 workshop. The same explanation applies to how satisfied the participants were with their own contribution.

When testing OKMv2, it was observed during IP1 that the participants were less active during the last minutes of the last rounds. The participants reported that they had experienced difficulty coming up with new ideas at the end of these rounds and this is also reflected in the questionnaire. OKMv2 scored much lower on “If I had had more time, I would have come up with more ideas” compared to the other methods. Another way of judging whether a topic is more exhausted is to analyse the types of ideas appearing at the end of the ideation
Figure 6.1 The participants’ rankings for different statements about each method.

Table 6.6 The statements in the questionnaire.

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am engaged in the topic.</td>
</tr>
<tr>
<td>2</td>
<td>If I had had more time, I would have come up with more ideas.</td>
</tr>
<tr>
<td>3</td>
<td>I had more ideas than I presented.</td>
</tr>
<tr>
<td>4</td>
<td>Many interesting ideas were presented during the workshop.</td>
</tr>
<tr>
<td>5</td>
<td>Many ideas were new to me.</td>
</tr>
<tr>
<td>6</td>
<td>I think ideas that were presented can be used in practice.</td>
</tr>
<tr>
<td>7</td>
<td>I felt engaged during the workshop.</td>
</tr>
<tr>
<td>8</td>
<td>I am satisfied with my own contribution during the workshop.</td>
</tr>
<tr>
<td>9</td>
<td>I think the method is useful.</td>
</tr>
</tbody>
</table>

session. It was observed that the ideas that emerged at the end of IP1 of OKMv2 were obviously more crazy and unrealistic (e.g. killing plant life by using nuclear radiation or introducing animals that like water to address insufficient drainage at turnouts) than previous ideas. It has been reported in the literature that, when “the rate of suggested ideas has diminished almost to a standstill”, more intriguing and unusual ideas start to appear, but only a few people contribute then and most people would prefer to do something else (Byron, 2012). It could therefore be argued that with OKMv2 this point in time was reached, which did not happen during the other tests, and therefore the topic dealt with during OKMv2 was more exhausted. The point in time when the participants run out of ideas is probably dependent on the ideation issue and the participants them-
selves, but it can be concluded that it takes quite a long time to exhaust the

types of questions which have been dealt with in Stage 3; for example, in the
case of OKMv2, the participants used 63 minutes during IP1. Once this point is
reached, there is a choice of two alternatives: ending the ideation (or in the pre-

cent case, advancing to the next step of the method) or trying to squeeze out

unusual ideas at a lower rate while the members possibly start feeling that they

are wasting time. Nijstad et al. (1999) found that an individual working alone or

in a group stops brainstorming when they feel that it is no longer worth the ef-

tort, and this feeling is based on their subjective estimate of the probability of

them being able to generate more ideas on the ideation topic. The continuation

of ideation requires some effort to be made to overcome this stage, and perhaps

additional measures must be taken. Although many of the ideas produced in this

phase are obviously unrealistic, a few of them could lead to a breakthrough

through association.

With regard to how the participants viewed the ideas generated during the
ideation, OKMv1 scored highest on the number of interesting ideas presented

and OKMv2 on the newness of the ideas to the participants. The variations con-
cerning these parameters between the methods may, however, be a result of how

familiar the participants were with the ideation topic. When rating the extent to

which the presented ideas could be used practically, the participants awarded

higher scores to the two versions of the novel method than they awarded to

Method 635 and the gallery method. When analysing the data from the inter-
views, it was observed that this could be due to the discussions during the group

phases and the possibility of refining ideas in several steps. OKMv2 scored

highest on how useful the method was thought to be, which is reflected in the

comments by the participants that this was the best of the methods that they had

tried thus far.

The quantitative total method scores are shown in Table 6.7 (where the time
for each method application is also shown). For comparison, the total score of
the SIL method is also shown.

Table 6.7 The quantitative total score of each method.

<table>
<thead>
<tr>
<th></th>
<th>635</th>
<th>Gallery</th>
<th>SIL</th>
<th>OKMv1</th>
<th>OKMv2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>5.46</td>
<td>6.03</td>
<td>5.00</td>
<td>5.47</td>
<td>6.65</td>
</tr>
<tr>
<td>Time (min.)</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>110</td>
<td>158</td>
</tr>
</tbody>
</table>

Interestingly, this fairly rough way of calculating the total score fits fairly well
with other results. In Chapter 5, it was found that, overall, the gallery method
was the most popular method and the SIL method the least popular, which is
reflected in Table 6.7. Table 6.7 also reflects the fact that OKMv2 was found by
the participants during the interviews to be better than any of the other methods.
However, the scores do not reflect the fact that in the interviews the participants
regarded OKMv1 as being better than the methods that they had tried earlier,
since the gallery method received a higher score than OKMv1 (with Method
receiving a similar score). A possible explanation is that the questionnaires were filled in on different occasions and that the participants had different methods to compare with. Further, the method used is not the only parameter affecting the statements in the questionnaire. For example, during the test of OKMv1 two participants mentioned that they had had a “bad day” and this was reflected in one of these participants’ individual level of agreement with the statements “I felt engaged during the workshop” and “I am satisfied with my own contribution during the workshop”. If this participant’s values were to be removed from the average of these two statements, the total score of OKMv1 would change to 5.72.

Considering the correlations between different statements in the questionnaire, a few points are interesting to note. However, it should be kept in mind that the small number of participants and therefore data points makes the results uncertain and the findings should be regarded as tentative. The only correlation with any significance at all, see Figure 6.2, was between how engaged a participant felt during the workshop and how satisfied that participant was with his own contribution ($R^2=0.68$).

![Figure 6.2](image.png)

This tentatively indicates that there might be a causal relationship between the engagement felt by the participants and the effort which they make during the workshop. However, another possibility is that, if a participant feels more engaged, they will perceive their own contribution as better although it might not objectively have been so. Similarly, their engagement might bias their view of how useful the method is. Between the engagement felt by the participants and how useful they found the method, an $R^2=0.43$ was found. Previous research
concerning brainstorming found that group members enjoyed their work more and were more satisfied with their own performance than individual brainstormers, although they were actually less productive (e.g., Stroebe et al., 1992), and that one reason for this might be a tendency to appropriate ideas of others in the group (Paulus et al., 1993). One cannot make a straightforward comparison between these findings and the findings from the application of ideation methods in the OptiKrea Group because of the different circumstances of the studies, but this underlines the importance of taking both objective and subjective results into account when determining what method is most promising, since the views of the participants might be biased by their engagement. Since most participants had similar levels of agreement with most of the statements in the questionnaire, the pair-wise plotting of the statements against each other in most cases resulted in a clustering of most of the data points on which each linear regression was performed, as is exemplified in Figure 6.2.

The absence of correlations for some pairs of statements is interesting. How useful the participants found the method and how satisfied they were with their own contribution during the workshop did not correlate ($R^2=0.18$). A possible explanation for this is that the participants tried to judge how useful the method would be in general, while they rated their own contribution in the specific instance. How engaged the participants were in the ideation topic did not correlate to how engaged they felt during the workshop ($R^2=0.10$) or how useful they found the method ($R^2=0.01$). Other researchers (e.g., Isaksen, 1998; Bolin and Neuman, 2006) have suggested that ownership of the task or the strength of the incentive for high performance is important for the outcome of a brainstorming session. One would intuitively think that a high engagement in the topic should lead to a high engagement during the workshop and a better outcome compared to a lower motivation to address the topic on the part of the group. The absence of a correlation between the engagement in the topic and the engagement felt during the workshop indicates that other factors are more important. However, other factors that could be expected to influence the engagement felt during a workshop, for example the extent to which the method resulted in ideas that could be used in practice ($R^2=0.06$), how many ideas were new to the participants ($R^2=0.01$) or how interesting the generated ideas were ($R^2=0.06$), were also uncorrelated. Neither were these factors correlated to how useful the participants found the method ($R^2=0.18$, 0.00 and 0.12, respectively). This may be because the right factors were not investigated in the questionnaire or because the relations between these parameters are more complex and cannot be explained by linear correlations between different pairs of statements. In any case, more instances of participants not being engaged in the topic would be required to obtain a reliable correlation.
6.3.3 Ideation outcome

The group generated concepts that included physical solutions to the problem of the ideation topic, organizational solutions and maintenance solutions. The concepts were described by words, a sketch, or by a combination of a sketch and words. Examples of concepts from each ideation session are shown in Figure 6.3 and Figure 6.4, respectively. In all the examples of concepts and ideas presented throughout this chapter, the text or parts of text have been translated from Swedish into English. Examples from the tests of Method 635 and the gallery method are to be found in Chapter 5. A popular-scientific overview of most of the generated concepts from the OptiKrea project is to be found in Part II of this thesis.

It was expected that the developed method would outperform the other methods with regard to the quantity of ideas generated, and, as is found from Table 6.8, this expectation was fulfilled if one considers the total number of ideas. However, one can question whether the developed method outperformed the other methods in terms of the overall concept and idea generation rates presented in Table 6.8. The generation rate for new ideas during IP1 tended to decrease as the length of this step increased, which is consistent with previous research on verbal brainstorming showing that the idea flow declines with time (e.g. Paulus and Dzindolet, 1993; Byron, 2012). However, although the idea generation rate decreased somewhat, not least due to the minutes with less activity during the last rounds of IP1, the second version of the novel method had a surprisingly good idea generation rate during this step, considering the fact that the time was doubled compared to that allocated for the first version. All of the participants thought that it was better to spend some extra time, despite a lower idea generation rate, to achieve a greater absolute number of ideas.

Method 635 and the similarly executed IP1 of OKMv1 were of equal length, but the former generated more ideas. Several parameters can explain this result, but the fact that two of the six participants experienced a “bad day” when testing OKMv1 and were unhappy with their own effort probably contributed. Nevertheless, the group was very satisfied with the results of the workshop, especially with the breadth of the ideas generated during the workshop. Therefore, another explanation may be that the scope of the topics dealt with in Method 635 and OKMv1 differed somewhat in breadth, despite efforts to find relevant topics of a similar scope, and that in relation to the topic the performance of IP1 of OKMv1 was good.
Figure 6.3 Examples of concepts from the OKMv1 ideation session on how to adjust the position of a modular turnout.

- Fixed point for screw nut
- Introduce a ballast suction apparatus to reduce the amount of ballast in the crossing section.
- Cradle under the rail that can adjust the angle of the rail.

Excavation and filling to frost-safe depth. Requirements for drainage.
Standardized way of working when placing a turnout, providing a good, stable substructure that meets requirements for fixed track and involves the use of a crane during placement. Renovation of worn-out turnouts so that a "replacement system" can be arranged. Document, train, certify, procure based on this.

Knock in place with a sledge hammer or fasten/adjust with an adjustment screw bolt.

Allow heating of ballast (kill roots) combine with environmentally approved anti-plant chemicals.
Pack a first layer of ballast so that the stone surfaces get an outward angle.
Drainage mats with a space of 1 dm between them. The material must be wear-resistant so that it's not cut to pieces by the ballast.

Several pipes with filling such as hawser suck water over ridge. Thirsty trees or especially thirsty bushes are planted.

Do these mats exist? Interesting.

Place wires in the ground/substructure to be able to measure the drainage function. (Double wires are used in the isolation of district heating pipes. When the conductivity increases, there is a leakage.)

Figure 6.4 Examples of concepts from the OKMv2 ideation session on drainage in turnouts.
Table 6.8 Number of concepts and ideas, length of time and idea and concept generation rates for each phase of each method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Phase</th>
<th>Number of concepts</th>
<th>Number of new ideas</th>
<th>Average number of ideas/b concept</th>
<th>Time (min.)</th>
<th>Concept generation rate (concepts/min.)</th>
<th>New ideas generation rate (ideas/min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>IP1</td>
<td>47</td>
<td>125</td>
<td>3.43</td>
<td>30</td>
<td>1.56</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>47</td>
<td>125</td>
<td>3.43</td>
<td>30</td>
<td>1.56</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td>IP1</td>
<td>25</td>
<td>66</td>
<td>2.92</td>
<td>15</td>
<td>1.67</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>GP1</td>
<td>2</td>
<td>15</td>
<td>-</td>
<td>25</td>
<td>0.08</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>IP2</td>
<td>2</td>
<td>5</td>
<td>NA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5</td>
<td>0.40</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>29</td>
<td>86</td>
<td>NA&lt;sup&gt;c&lt;/sup&gt;</td>
<td>45</td>
<td>0.64</td>
<td>1.91</td>
</tr>
<tr>
<td>Gallery</td>
<td>IP1</td>
<td>36</td>
<td>100</td>
<td>3.31</td>
<td>30</td>
<td>1.20</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>GP1</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>37</td>
<td>-</td>
<td>0.32</td>
</tr>
<tr>
<td>OKMv1</td>
<td>IP2</td>
<td>13</td>
<td>31</td>
<td>5.46</td>
<td>10</td>
<td>1.30</td>
<td>3.10</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>-</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>53</td>
<td>144</td>
<td>3.85</td>
<td>110</td>
<td>0.48</td>
<td>1.31</td>
</tr>
<tr>
<td>OKMv2</td>
<td>IP2</td>
<td>10</td>
<td>32</td>
<td>7.80</td>
<td>10</td>
<td>1</td>
<td>3.20</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>-</td>
<td>15</td>
<td>-</td>
<td>31</td>
<td>-</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>87</td>
<td>247</td>
<td>3.60</td>
<td>158</td>
<td>0.55</td>
<td>1.56</td>
</tr>
</tbody>
</table>

<sup>a</sup>“New ideas” means those ideas which were presented for the first time in the ideation session in the step in question.

<sup>b</sup>Here “ideas” means all ideas, including ideas from earlier steps which have been reused and/or ideas suggested in other concepts.

<sup>c</sup>Not applicable (NA), because the participants continued working with their sheet of paper from IP1, while in OKMv1 and OKMv2 they received new sheets.
Although some ideas were expressed verbally, it can be seen from Table 6.9 that most of the spoken words were dedicated to the elaboration of ideas during all the group phases.

**Table 6.9** Distribution of spoken words between activities (%) during each group phase of the ideation sessions.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Going off at a tangent</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Elaboration</td>
<td>80</td>
<td>78</td>
<td>80</td>
<td>74</td>
<td>82</td>
</tr>
<tr>
<td>Ideas</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Criticism</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The total number of words spoken during each phase of the method applications is shown in Table 6.10. The longer time the participants had to ideate individually, the longer was the time it took to present, explain and discuss the ideas in the subsequent step (see Table 6.8).

**Table 6.10** Total number of interpretable spoken words during each group phase of the ideation sessions.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken words</td>
<td>3781</td>
<td>5356</td>
<td>5032</td>
<td>7374</td>
<td>4783</td>
</tr>
</tbody>
</table>

From Table 6.8 it is obvious that there is a great difference in the idea generation rate between the individual and group phases. During IP1 and IP2, where the participants work individually, significantly more ideas were generated than during the group phases, as can be seen in the far right column in Table 6.8. It might be tempting to draw the immediate conclusion that, to maximize the total idea generation rate, only individual phases should be included. However, as was observed during IP1 of OKMv2, the participants eventually ran out of ideas. After the group phases, the intention was that the participants would be able to create associations with what had been presented and come up with new ideas. Although it is difficult to quantify the extent of inspiration, it was found that the ideas that had turned up in IP2 in many cases had been based on some issue that had been brought up during GP1. The average concept generation rate was lower in IP2 than in IP1, but the average number of ideas per concept was higher in IP2, reflecting the fact that several of the participants aimed to combine ideas from previous phases with new ideas to achieve a complete solution to the problem of the ideation topic (as opposed to producing several unrelated ideas). An example of such a combination of ideas is shown in Figure 6.5. The fraction of ideas from earlier phases that each participant reused in concepts generated in
IP2 varied from 0 to 100%, but was on average 57% for OKMv1 and 54% for OKMv2, i.e. more than half. It is especially interesting that, although the participants had difficulties creating any more ideas during the last rounds of IP1 of OKMv2, despite exchanging ideas through rotational viewing, after the verbal interaction in GP1 they could squeeze out 32 new ideas in IP2. From these results it is concluded that the verbally interactive steps including the presentation and discussion of ideas are relevant in the method. Kohn et al. (2011) found that interactive brainstorming groups generated combinations of previously generated ideas of higher utility than nominal groups. Nijstad et al. (1999) found that when groups brainstorm verbally without a time constraint, the productivity loss in relation to the productivity of nominal groups decreases, because the groups continue to ideate for a longer time. These findings indicate that verbal group interaction may be beneficial in group ideation if it is applied in the right instance, and may perhaps be especially beneficial in the longer ideation sessions that may be required in the field of engineering design. This may be especially fruitful when the absolute number of ideas is more important than the idea generation rate. One would expect that industrial companies would rather use some extra hours during ideation to assure that as many ideas as possible would be collected before starting the selection process. These extra hours would be much fewer than the total number of hours that would be used to develop a new product. Although the idea generation rate is an interesting parameter for evaluation purposes, the total number of ideas in a day or even a week is likely to be more relevant to companies and other organizations. It might therefore be more interesting to focus on how to make a group create more ideas on a topic which they think they have exhausted than mainly to consider the idea generation rate.

![Figure 6.5 Example of a concept generated during IP2 that combined several ideas and concepts from IP1 and GP1](image)

6.3.4 Behaviour of the participants

In Table 6.11 it can be observed that the supervisor spoke much less during GP1 and GP2 of OKMv1 and OKMv2 than during GP1 of the gallery method, indicating that the group worked more independently when applying the developed method. Although this is positive, since it suggests that the group can work
through the method without a supervisor, it must be acknowledged that each member of the studied group had had sufficient time to understand and learn the methods (despite the differences between them), as well as acquaint themselves better with the other group members. Therefore, at this stage one cannot conclude with certainty that a group with “beginners” would be able to accomplish similar results when applying OKMv2.

Table 6.11 Distribution of spoken words between the participants (%) during each group phase of the ideation sessions.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>19</td>
<td>9</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>16</td>
<td>6</td>
<td>9</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>16</td>
<td>23</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>E</td>
<td>23</td>
<td>26</td>
<td>29</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>F</td>
<td>26</td>
<td>23</td>
<td>22</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Supervisor</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.12 shows the average deviation from an equal distribution of the spoken words and speaking frequency during the group phases (calculated using Equation 6.1). If the number of spoken words had been completely equally distributed among the participants, the deviation would have been zero.

Table 6.12 Average deviation from an equal distribution of the spoken words and speaking frequency (absolute %) during each group phase of the ideation sessions.

<table>
<thead>
<tr>
<th></th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free discussion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoken words</td>
<td>7.3</td>
<td>6.4</td>
<td>9.4</td>
<td>4.3</td>
<td>5.8</td>
</tr>
<tr>
<td>Speaking frequency</td>
<td>6.0</td>
<td>7.2</td>
<td>10.0</td>
<td>4.1</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Table 6.12 shows that the average deviation from an equal distribution of the spoken words and speaking frequency among the participants decreased when the participants took turns to comment on the ideas (in OKMv2), suggesting that this might be a feasible procedure, although the same people as before still spoke the most (see Table 6.11). The distribution of spoken words is not only dependent on the design of the method, but also, for example, on the personality of each group member and how they are affected by the behaviour and personalities of the other group members. Therefore, it is not likely that the observed speaking patterns would disappear completely by taking turns. Another explanation for the decrease in the deviation from an equal distribution of the spoken
words and speaking frequency may be that the participants had become aware of the unequal distribution of spoken words and therefore tried to adjust their behaviour (e.g. the group members who had spoken a great deal previously consciously tried to speak less). All the participants thought that letting all the group members around the table take turns to speak was a good improvement. The participants themselves jointly initiated a procedure whereby the sheet of paper containing suggestions presented by one participant was circulated among the other participants for them to comment on the ideas in turn. This made it clearer when the next person should start giving feedback and easier for the participants to see which concepts they were supposed to comment on.

When testing OKMv1, the participants felt that they had not had enough time to follow through on their own ideas created during the first round, but had been interrupted by the ideas on the sheet of paper given to them by their neighbour after 5 minutes. Linsey et al. (2011) hypothesized that “rotational viewing encourages the participants to spend more time understanding other [team-mates’] ideas”, which is in line with the experience from the present research. Attending to the ideas of others is a critical factor for the observation of stimulation effects in groups (Paulus et al., 2015). The first IP of the developed method actually consists of a) attending to the content of the sheet of paper and b) generating new ideas. To allow the participants both to comprehend the existing ideas on the sheet of paper received and to generate more ideas, the length of time for each round was successfully increased to ten minutes. With each successive round, each sheet of paper included more and more information, and it took a longer time for the participants to acquaint themselves with this increased amount of information during the remaining rounds. However, it is important not to prolong the duration of each round unnecessarily, since previous research has shown that individuals under time pressure work at a faster rate and that the less time there is available, the greater is the task focus shown by groups; however, the quality of the performance may suffer if too high a pressure is applied, since information is processed less thoroughly then and a narrower range of alternatives are considered (Karau and Kelly, 1992). Too long a time might also allow the participants to start thinking about disadvantages of their ideas and decide not to share them with the group, and to dedicate too much effort to a limited part of the solution space through design fixation. It was concluded that ten minutes was the right length of time for the topics considered in the present research. Linsey et al. (2011) used ten minutes for individual idea generation prior to four subsequent 7.5 minute periods where each group of five participants viewed either all the ideas (gallery viewing) or a new subset of ideas during each period (rotational viewing). The total time of 40 minutes in the study performed by Linsey et al. (2011) is shorter than the time used for IP1 in the present research, but longer than the total ideation time reported for the typical study in psychology literature (examples of ideation times can be found in Parnes, 1959; Paulus et al., 1993; Nijstad et al., 2002; Wegge and Heslam, 2005; Paulus et al., 2015), which does not deal with design problems. An intui-
tive assumption is that the appropriate total and cycle times will depend on the ideation topic and the size and experience of the group concerned, and developing guidelines for setting these times as a function of these parameters is an interesting area for future research.

After the test of OKMv1, two design principles were added and put into practice in OKMv2 (see Table 6.4). In the test of OKMv2 it was confirmed that the addition of these design principles resulted in an improvement. As a result of the continuous reflection taking place during the study, it was realized that one of the procedures that had been put into practice entailed yet another important design principle which was added to the set of principles. All the design principles to be used for the development of ideation methods to be used in cross-functional inter-organizational groups on the Swedish deregulated railway market are to be found in Table 6.4.

All the design principles were put into practice in OKMv2. Table 6.13 summarizes the evidence suggesting that OKMv2 is the best method for the given context with respect to the requirements (connected to each goal). OKMv2 was found to generate the highest number of concepts and ideas. Qualitative evidence suggests that the participants found that OKMv1 and OKMv2 gave concepts with more breadth and potential than the other tried methods. However, this was difficult to assess quantitatively, and the evidence in this case derives from the earlier findings in the literature stating that quantity and quality are correlated (Dugosh et al., 2000; Paulus et al., 2015). The participants did not report having and were not found to have any difficulties using the method. The participants found OKMv2 to be the best method that they had tried, and it scored highest regarding how useful the participants found the method to be, and regarding the total method score. Other advantages of OKMv2 are the continuous documentation during the ideation session, the fact that all the participants obtain an overview of the entire set of concepts (which, however, also applies to some of the other tested methods), and the fact that the speaking distribution was more equal.

However, the participants in the OptiKrea Group became more and more familiar with each other and the use of ideation methods during the course of the project, and this might have affected the findings from the application of the different methods. It is also possible that the participants in the OptiKrea Group were positively biased towards the developed method because of their participation in its development. It would therefore be desirable to test the method in other groups, and especially to obtain the views of other participants on the method and to ascertain how easy it would be for them to apply it.
Table 6.13 Overview of the different requirements formulated for the method to be developed and evidence showing that OKMv2 met these requirements best.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goal a) The method should produce a great number of ideas with breadth and potential.</strong></td>
<td></td>
</tr>
<tr>
<td>Generate a great number of ideas</td>
<td>OKMv2 had the greatest number of concepts and ideas (Table 6.8).</td>
</tr>
<tr>
<td>Generate ideas with breadth and potential</td>
<td>Some qualitative evidence that OKMv1 and OKMv2 gave concepts with more breadth and potential (see Section 6.3.2.1 and 6.3.2.2). Not measured quantitatively since it is difficult to assess (see Section 5.3.2). Reliance on findings in the literature stating that quantity and quality are correlated (see Section 2.2.2).</td>
</tr>
<tr>
<td><strong>Goal b) The method should be implemented by STA after the project finishes.</strong></td>
<td></td>
</tr>
<tr>
<td>Be easy to use and understand</td>
<td>The participants did not report having and were not observed to have any difficulties using the method.</td>
</tr>
<tr>
<td>Be attractive to the users</td>
<td>The participants thought that OKMv2 was the best method of those they had tried (see Section 6.3.2.2). OKMv2 scored highest regarding how useful the participants found the method (see Figure 6.1). OKMv2 had the highest total method score (see Section 6.3.2.3).</td>
</tr>
</tbody>
</table>

6.4 Conclusions

The goals for the method to be developed were found to be that the method should:
- produce a great number of ideas with breadth and potential,
- be implemented by STA after the project finishes.

The requirements for the method were found to be that it must:
- generate a great number of ideas,
- generate ideas with breadth and potential,
- be easy to use and understand,
- be attractive to the users.

The participants thought that the developed method was the best method that they had tried. The method was found by the participants in the group to be more useful and to generate more ideas that could be used in practice compared to the other methods that they had tried.
The developed method outperformed the other methods with regard to the total number of generated concepts and ideas, but not with regard to the overall concept and idea generation rates.

The participants did not report having and were not observed to have difficulties understanding and applying the method. Most of the spoken words were dedicated to the elaboration of ideas during all the group review phases. The distribution of spoken words between the participants became more equal when the procedure of taking turns to speak was put into practice in the second version of the method.

Considerably more ideas were generated during the individual phases than during the group review phases. The group review phases were mainly used to present and elaborate on concepts and ideas. The average fraction of ideas from earlier phases that each participant reused in concepts generated in the second individual phase was more than half.

The learning derived from the development of an ideation method in a single group was formalized into the following design principles for ideation methods to be used in cross-functional inter-organizational teams in the context of the Swedish deregulated railway market:

- development of the method based on methods proven to be easy to understand,
- the initial generation of a pool of ideas with great variety,
- exposure to stimulus ideas generated internally within the group,
- mixing idea-exchanging strategies,
- the use of external representation in the form of sketches and text,
- including time for discussion and debate,
- optimization of the cycle time,
- development of a strategy for distributing the verbal interaction,
- paying attention to stimulus ideas.

The developed method was found to be the best of the methods tried with respect to the requirements formulated for the specific context.
Appendix 6.1

- What did you think was good and bad about the method?
- What things can be improved?
- Other points of view?
- How do you compare this ideation method to the other methods the group has tried?
- Did you feel that the concepts were developed further?
- Would you like to work with others during one or both of the individual ideation steps, e.g. with one or two partners?
  - Why/why not?
- Would you like to have additional individual ideation steps?
  - Why/why not?
Appendix 6.2

- How did you feel about the time for each 635-round being increased from 5 to 10 minutes?
  - Too much/too little time?
  - Was it less stressful?
  - Did you have time to generate more ideas?
- Today we chose to take turns when giving feedback after each presentation. How did you experience that?
- How did you feel about the time length of the gallery step (step [IP2])?
- How did the structuring of ideas and concepts work?
  - Did it give additional value?
- How did you experience this workshop compared to the [previous one]?
- Do you have any suggestions about how the method could be changed to work better?
7. Field tests of the final version of the developed ideation method

7.1 Introduction

This chapter presents Stage 4 of the research performed for this thesis (see Table 1.1), involving field tests of the final version of the ideation method (OKMv2) developed in Stage 3 (Chapter 6). The opportunities for conducting field tests were given in two different settings: one of the work packages in an EU project called In2Rail and a Swedish collaboration pilot project called UPPSAMT. The In2Rail project aimed to create the technology necessary to complete the Single European Railway Area. The UPPSAMT project investigated how data from rolling stock can be used to assess the condition of the infrastructure, with a focus on the overhead line.

One of the requirements set for the development of the ideation method was that the method must be easy to use and understand. Therefore, it was expected that the method would be easy to apply during the field tests. Further, it was expected that the participants of the field tests would share the views of the OptiKrea Group on the method. In the case of the In2Rail test, it was expected that the groups would generate a comparable number of suggestions to the number generated by the OptiKrea Group. The topic of the UPPSAMT workshop was directed more towards data management, the handling of information and organizational issues than the topics that had been applied in the OptiKrea project. Therefore, it was less meaningful to compare the number of suggestions generated in the UPPSAMT groups with that generated in the OptiKrea Group. On the other hand, this gave the opportunity to see if the method worked well with this type of topic, and it was expected that it would.

The objective of the field tests was to evaluate the ideation method which had been developed within the OptiKrea Group (see Chapter 6) in other cross-functional inter-organizational groups on the railway market. The studies aimed to explore how the method would perform in these groups and to judge if the findings from the development of the method in the OptiKrea Group could be verified in other groups in similar contexts.

The following research questions were posed to meet the objective.

- Were the field test groups able to apply the method as intended?
What did the participants in the field test groups think about the method?

What was the ideation outcome and, for each field test, how did the outcome differ between the groups within that field test?

How did the results compare to those from the evaluation of the method in the OptiKrea Group?

Can the findings from the development in the OptiKrea Group be verified in the field test groups, especially considering the fact that one of the field tests took place in a European context and the other applied the method to an ideation topic of a different nature?

### 7.2 The In2Rail field test

#### 7.2.1 Method

**7.2.1.1 Participants and groups**

A total of 17 people participated in the In2Rail workshop. On the first day, 16 people took part in the workshop and on the second day the number of participants was 15. Fourteen people took part on both days and three people only during one day. The participants came from six different European countries (Sweden, Great Britain, Germany, France, Italy, and Austria) and represented five different infrastructure managers (STA, Network Rail, SNCF, Deutsche Bahn and ÖBB), four different universities (two Swedish universities and two British universities) and four different companies. Two of the participants were also members of the OptiKrea Group.

The participants were asked to answer a questionnaire regarding their background and prior experience of ideation methods (see Appendix 7.1). The results of the questionnaire are shown in Table 7.1. The ten participants who had used ideation methods previously were asked which ideation methods they had experience of. Apart from the two people who had been members of the OptiKrea Group, none of the participants reported having tried any other method that resembles the developed method.

During the first day, the participants were assigned by the work package leader to two different groups according to what task group they belonged to in the In2Rail project. Some participants took part in both task groups, and then they were assigned to the group considering the topic which they were judged to have most knowledge about. During the second day, the participants were split into two new groups with the aim of reaching a good balance between the types of participating organizations in each group. This assignment to different groups was carried out by one of the task leaders, taking the participants’ opinions into consideration. Table 7.2 shows the details of each In2Rail group, as well as the
Table 7.1 Participants’ replies to questions about their background and prior experience of ideation methods.

1. What type of organization do you represent?

<table>
<thead>
<tr>
<th>Organization Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure manager</td>
<td>8</td>
</tr>
<tr>
<td>Supplier</td>
<td>2</td>
</tr>
<tr>
<td>Academia</td>
<td>4</td>
</tr>
<tr>
<td>Another organization</td>
<td>3</td>
</tr>
</tbody>
</table>

2. How many years of experience do you have from the railway sector?

<table>
<thead>
<tr>
<th>Experience Period</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>6</td>
</tr>
<tr>
<td>5-10</td>
<td>5</td>
</tr>
<tr>
<td>10-15</td>
<td>1</td>
</tr>
<tr>
<td>15-20</td>
<td>3</td>
</tr>
<tr>
<td>20-25</td>
<td>0</td>
</tr>
<tr>
<td>&gt;25</td>
<td>2</td>
</tr>
</tbody>
</table>

3. Have you used ideation methods before?

<table>
<thead>
<tr>
<th>Answer</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td>No</td>
<td>7</td>
</tr>
</tbody>
</table>

4. If you answered “Yes” to Question 3, how often have you used ideation methods?

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have used them a few times</td>
<td>4</td>
</tr>
<tr>
<td>A few times per year</td>
<td>4</td>
</tr>
<tr>
<td>Every month</td>
<td>1</td>
</tr>
<tr>
<td>Every week</td>
<td>0</td>
</tr>
<tr>
<td>Other frequency</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7.2 Details of the groups (G=Group, OKG=OptiKrea Group).

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
<th>OKG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants representing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure manager</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Supplier</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Academia</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other organizations</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of organizations represented</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Number of participants</td>
<td>7</td>
<td>9a</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
</tbody>
</table>

*aOne participant in Group 2 arrived 7.5 minutes after the group had started applying the ideation method.
details of the OptiKrea Group for comparison. The participants who were members of the OptiKrea Group were assigned to different groups throughout the workshop, so that each of the four In2Rail groups had one participant from the OptiKrea Group.

7.2.1.2 Procedure

The workshop was planned in accordance with the prerequisites given by the leaders of the work package. Each of the two groups selected on the first day worked on a different ideation topic. Each of the two groups selected on the second day worked simultaneously on both the topics chosen for that day, and the reason for this was that the two topics had points in common and synergies could be expected. The ideation topics are presented in Table 7.3.

<table>
<thead>
<tr>
<th>Group</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Novel turnout locking mechanism</td>
</tr>
<tr>
<td>2</td>
<td>Embedded and integrated sensors</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>Radical turnout concept</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>Self-inspecting/-correcting/-adjusting turnout</td>
</tr>
</tbody>
</table>

According to the plan drawn up, the different task groups in the In2Rail project were supposed to prepare the ideation session related to their task by executing a topic analysis and goal specification prior to the workshop, and each day was to begin with a presentation of the outcome of the preparation. Thereafter, the ideation method was to be applied to each topic on the first and second day. On each day, after the ideation method had been executed, each group was to organize their ideas into different categories and finally present these to the other group.

Before the first ideation session, the ideation method was presented to the participants and its instructions were shown by means of a projector and explained to the participants for about ten minutes. The instructions had also been sent to the participants prior to the workshop, but few participants appeared to have read them. The instructions, as given to the participants, are shown in Table 7.4. The participants were told that only six rounds would be performed in Stage 1. Group 2 and 4 were supervised by one of their participants, who had taken part in the OptiKrea project and was familiar with the method. Group 1 and 3 were supervised by the author of this thesis, who did not take part in the ideation, but gave instructions, when necessary, and answered questions. Each group worked in a separate meeting room, and was provided with sheets of A3 paper, pens, tape and pairs of scissors.
Table 7.4 Ideation method instructions (OKMv2).

General directions
a) The instructions herein presuppose that the group has understood and agreed upon the ideation topic.
b) Maximize the number of ideas (old and new).
c) Avoid negative criticism during ideation (the ideas will be thoroughly examined after the ideation).
d) Appoint one person to be responsible for keeping track of the time with the help of a timer.

Stage 1: Rotational viewing
a) During ten minutes each participant comes up with at least three ideas on how to address the ideation topic. Each participant documents their ideas with sketches and/or text on a sheet of A3 paper.
b) Each participant gives their sheet of paper to the neighbour on their left.
c) The neighbour reads through the ideas and adds at least three improvements, combinations of the ideas and/or new ideas on the sheet of paper during ten minutes. One is allowed to ask the neighbour on one’s right what they meant by an idea that is incomprehensible.
d) The sheets pass between all the participants (i.e. Stage 1 ends when one receives the sheet of paper which one started out with).

Stage 2: Presentation of ideas and feedback from other participants
a) The participants use five minutes to read through the ideas that have been added to the sheet of paper which they started out with.
b) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.
c) After each presentation, the presented sheet is circulated among the participants and each participant gives feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions.
d) A maximum of ten minutes is allowed per sheet of paper (for presentation and feedback).

Stage 3: Gallery viewing
a) The sheets of paper from Stage 1 are put up on a wall or some other place where all the participants can easily view them.
b) Each participant works individually during ten minutes, developing or combining ideas from the collection of ideas from Stage 1. New ideas are also welcome. New sheets of A3 paper are used to document the ideas by means of sketches and/or text.
c) The aim is still to collect as many ideas as possible!

Stage 4: Presentation of ideas and feedback from other participants
a) Each participant presents their own ideas from Stage 3.
b) After each presentation the presented sheet of A3 paper is circulated among the participants and each participant gives their feedback on the ideas. The remaining available time is used for discussions.
c) A maximum of five minutes is allowed per participant (for presentation and feedback).
7.2.1.3 Behaviour of the groups

Each group was audio-recorded and the recordings were used to analyse how the participants had interpreted the ideation method instructions and applied them to the topic. The recordings were listened to and after any instance of words spoken concerning the instructions of the method and how they were carried out in practice by each group, the recording was stopped and a note of that instance was made. The note included the time of the recording of that instance, a description of what had been said and reflections on the part of the author, if any. After the entire recording had been listened to, the notes were examined and summarized.

7.2.1.4 Views of the participants

After the ideation method had been applied each day, each participant answered a questionnaire presenting different statements about the method. One participant did not fill in the questionnaire on the first day. This was the same questionnaire as had been applied in the OptiKrea Group and is described in Section 5.2.2 and shown in Appendix 5.1. The average group value for each statement was calculated, as well as the average values for the subgroups of infrastructure managers (IMs), suppliers, academia and other organizations. It was known that competitors in different market segments participated in the In2Rail project. Therefore, in addition to the statements from the original questionnaire applied in the OptiKrea Group, a question was posed which asked the participants whether they had kept ideas to themselves due to the presence of competitors, since the participants in the OptiKrea Group had mentioned that they would have kept ideas to themselves if competitors had been present. This effect had been observed in a parallel project conducted by Prof. Jan Lundberg of Luleå University of Technology and has also been reported by other researchers (Almefelt and Claesson, 2015). However, it turned out that there were no direct competitors participating in the In2Rail workshop. However, representatives of different actors might still have withheld ideas, since the outcome of the workshop was to be reported to the other participants in the In2Rail project.

A linear regression was performed between each pair of statements in the questionnaires to find out if any correlation existed. The available questionnaire results from the In2Rail workshop resulted in 30 data points (15 participants filling in a questionnaire after the first and second application of the ideation method, respectively).

A linear regression was performed between the average levels of agreement with the different statements in the questionnaire from the OptiKrea Group and the In2Rail groups, respectively. Statement 1 and 10 in the questionnaire were excluded from this linear regression, since they were not considered to be dependent on the method (in addition, statement 10 was not answered by the OptiKrea Group).
The participants were able to give written comments on the method in free text under the headings “I liked these things about the method”, “I found that the following things did not work well when applying the method” and “Other comments”. The free-text fields in the questionnaire were analysed according to the procedure in Table 7.5.

**Table 7.5** The procedure for analysing the free-text comments in the questionnaire.

1. The comments from those participants who had written any in the free-text fields at the end of the questionnaire concerning the method (as opposed, for example, to the composition of the groups or the workshop in general) were collected in a single word processing document according to the three free-text comment headings (“I liked these things about the method”, “I found that the following things did not work well when applying the method” and “Other comments”).

2. The comments were examined and comments concerning the same matter were listed together and tentative categories were formed. Comments concerning matters mentioned only by one participant were sorted in an “Other comments” category.

3. A check was performed to ensure that all the comments concerning matters mentioned by more than one participant fitted into a category, otherwise the comments which did not fit were split, if possible, or the categories were changed. When all the comments had been placed in a category, the categories were given descriptive labels. For each matter described in a category, a count was made of the number of participants who had mentioned that matter.

4. Comments that did not directly concern the method were examined and sorted according to the matter which they concerned. If these comments were found to concern matters that could contribute indirectly to an understanding of how the method had performed during the test, they were processed according to point 2 and 3.

**7.2.1.5 Ideation outcome**

The sheets of paper that the participants had used for documentation while applying the ideation method were copied and analysed. The sheets had contents of different types. An inductive approach was used in developing guidelines for categorizing the content. The entire content was perused several times. After the first round, the categories of content were tentatively defined. In the second round, a check was performed as to how the content fitted into these categories and the categories were more clearly defined. New categories were defined, as some content did not fit into the original categories. The third round involved formulating guidelines as to what category a certain content should belong to, and a final check was performed to ensure that all the contents belonged to a category. The guidelines are shown in Table 7.6. The number of pieces of content in each category generated by each group was counted.
Table 7.6 Guidelines for categorizing pieces of content.

1. As a starting point, the pieces of content are identified on the sheets as text and sketches that belong together and are marked off from other text and sketches by the participants.

2. The pieces of content belong to one of three main categories:
   A. Topic analysis and requirements: This category includes content that only treats issues related to the background of the topic, goals and requirements for the solution.
   B. Suggestions on how to address the ideation topic: Each suggestion constitutes a solution or an approach to a solution of the problem or a subproblem of the ideation topic on its own.
   C. Irrelevant content: This type of content does not have any relevance to the ideation topic.

3. Suggestions (main category B) are categorized into three subcategories:
   a. Physical suggestions: Physical solutions suggested to address the ideation topic.
   b. Organizational suggestions: Suggestions attempting to address the ideation topic by making some change in the organization which the ideation topic belongs to.
   c. Data management suggestions: Suggestions attempting to address the ideation topic by making some change concerning how data are managed.

4. Suggestions (main category B) may be a mix of physical, organizational and data management suggestions. If this is the case, the major contribution in the suggestion determines what category it should belong to.

5. Physical suggestions are either abstract or concrete suggestions. If the suggestion has no information on how to realize the solution, it is an abstract suggestion. If there is any information on how to realize the solution, it is a concrete suggestion.

6. Some suggestions mention alternatives. If an alternative is a variant of a minor function or detail that does not change the principal idea of the solution, the alternative belongs to the original suggestion. Otherwise it is counted as a separate suggestion.

7. Those suggestions are counted which reframe the problem and, although they do not specifically address the ideation topic as described, still meet the higher level needs.

8. Pieces of content that include the same information are counted only once.

Group 2 was supposed to ideate on embedded sensor technology, but re-framed the topic to include the properties which they considered important to measure and any type of method for measuring them. Consequently, a considerable part of the documented content had the character of a list. Two additional subcategories had to be used for the content generated by Group 2, i.e. two subcategories in addition to the subcategories needed to categorize all the pieces of content generated by the other groups. All the pieces of content generated by Group 2 that listed properties and measurement techniques were counted in an additional subcategory. Several concepts had the character of a list, but had additional information such as implications, questions and objections. These
were counted separately. The total number of properties and measurement techniques was also counted. Thereafter, the remaining content was categorized according to Table 7.6.

7.2.2 Results and discussion

7.2.2.1 Behaviour of the participants

Most of the participants had not contributed to the topic analysis and goal specification stages in advance. Therefore, more time than planned had to be dedicated to the common introduction during the workshop and, consequently, less time was available to perform the ideation sessions. This led to the following deviations from the method instructions in Table 7.4.

- Group 1 performed only Stage 1 and 2.
- Group 2 combined Stage 3 and 4 with categorization. The participants used Stage 3 to summarize, based on Stage 1 and 2, what they had found to be most promising and/or important. During the presentation of the summaries during Stage 4, the group simultaneously categorized the presented content.
- Group 3 and 4 completed five rounds during Stage 1.
- Group 4 set a maximum limit of five minutes per participant for presentation and feedback during Stage 2.
- Group 3 and 4 performed Stage 3 and 4 as one group. Stage 3 and 4 were used as they were used by Group 2 during Day 1. The participants were told, however, that they could also add any new ideas which they had thought of.

The ideal time consumption (as specified by the instructions) and the actual time consumption of each stage in each group are shown in Table 7.7. Between the different stages, there was typically a short break while the participants stretched their legs or fetched a cup of coffee. A few times, participants also left the group during a stage, e.g. to take a phone call.

On average, Group 1, 2 and 3 used a little less time than the ideal time for presentation and feedback during Stage 2. However, the participants presenting suggestions at the beginning of Stage 2 in these groups used more than ten minutes. As the remaining participants realized that the group was running out of time, many of them chose to present their ideas very quickly and not to give any comments or feedback on other participants’ suggestions. Therefore, the participants presenting suggestions at the end of Stage 2 had less time at their disposal than those presenting suggestions at the start, and unfortunately this means that interesting suggestions and viewpoints might not have been presented to the group. Very little time was available for Stage 4 in Group 3 and 4, and consequently the presentation assumed more the character of a discussion. Nevertheless, the group was able to classify its suggestions into several categories during the approximate time of 20 minutes which was available.
Table 7.7 Ideal time consumption and actual time consumption of each stage in each group (G=Group).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Idealb</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of rounds</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Ideation time/round (min.)</td>
<td>10:00</td>
<td>10:00</td>
<td>10:00</td>
<td>10:00</td>
<td>10:00</td>
</tr>
<tr>
<td>Stage 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review</td>
<td>05:00</td>
<td>05:00c</td>
<td>10:00</td>
<td>05:00</td>
<td>05:00</td>
</tr>
<tr>
<td>Presentation and feedback (min./participant)b</td>
<td>10:00</td>
<td>08:34</td>
<td>09:24</td>
<td>07:42</td>
<td>04:31</td>
</tr>
<tr>
<td>Stage 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideation time (min.)</td>
<td>10:00</td>
<td>-</td>
<td>10:00</td>
<td>10:00d</td>
<td>10:00d</td>
</tr>
<tr>
<td>Stage 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation and feedback (min./participant)b</td>
<td>05:00</td>
<td>-</td>
<td>04:43</td>
<td>01:19d</td>
<td>01:19d</td>
</tr>
</tbody>
</table>

aAs specified in the instructions.
bAverage time from when the participant starts presenting suggestions until when the next participant starts, including the time for feedback and discussion (excluding the time for disturbances such as a person entering the room).
cThis time is uncertain since the review of ideas was performed in connection with a break.
dGroup 3 and 4 performed Stage 3 and 4 as one group.

The deviations from the method instructions in Table 7.4 gave insights. The design of the method allowed it to be adapted to a specific situation through easy adjustments. An advantage of the method from this point of view is that it is built of stages that can be viewed as different modules. By removing modules, changing the number of rounds or the time per round, or the time for presentation and feedback, the method can rapidly be adjusted to the current situation.

Stage 3 and 4 were used in another way than intended. Instead of generating new suggestions, the majority of the participants in the In2Rail workshop chose to use the time for summarizing the content of Stage 1 and 2, and either put together a “complete” solution from different suggestions or listed their favourite suggestions. From an ideation perspective these stages appeared almost useless, since very few participants contributed new suggestions. Interestingly, the participants did not make any comments regarding the possibility of excluding Stage 3 and 4 from the method, either verbally or in the questionnaire. It appears that they found these stages useful despite the lack of new suggestions and used them as a bridge from ideation to categorization of suggestions.

One important goal of the method design was that it should be possible for participants who are unfamiliar with the method, or new to ideation methods in
general, to use the method directly. The participants were introduced to the method instructions shown in Table 7.4 for about ten minutes and were then asked to start the ideation session. They received further short directions from the person supervising each group before starting the ideation session and between stages when necessary. Although the participants were slightly hesitant at the start of the ideation sessions during the first day, they were able to apply the method directly to the ideation topics. There were few questions about Stage 1 and 2 of the method, whereas Stage 3 appeared to be more confusing to the participants, as in that stage they asked more questions about what they were supposed to do. A more useful approach would probably be to give paper copies of the instructions to every participant, as this would remove the need for remembering the instructions shown before the ideation session.

7.2.2.2 Views of the participants

Figure 7.1 shows the questionnaire results from the In2Rail workshop compared to those from the OptiKrea Group.

![Figure 7.1](image_url)  

**Figure 7.1** Average questionnaire results for each In2Rail group and the OptiKrea Group.

Figure 7.2 shows the questionnaire results from the In2Rail workshop according to the types of organization which the respondents represented. Since the number of participants representing each type of organization was quite small (see Table 7.1), it is not possible to draw any certain conclusions from Figure 7.2. Table 7.8 shows the statements included in the questionnaire. Overall it can be concluded from the questionnaire that the In2Rail project participants were satisfied with using the method. Figure 7.1 shows that the average level of agree-
ment in each group varies, but that the OptiKrea Group awarded higher levels of agreement with respect to statement 6-9. Compared to the OptiKrea Group, the participants in the In2Rail groups on average thought that they would have come up with more ideas if they had had more time, and reported having had more ideas than they presented to a higher degree than the OptiKrea Group.

Figure 7.2 Average questionnaire results according to the types of organization which the respondents represented.

Table 7.8 The statements in the questionnaire.

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am engaged in the topic.</td>
</tr>
<tr>
<td>2</td>
<td>If I had had more time, I would have come up with more ideas.</td>
</tr>
<tr>
<td>3</td>
<td>I had more ideas than I presented.</td>
</tr>
<tr>
<td>4</td>
<td>Many interesting ideas were presented during the workshop.</td>
</tr>
<tr>
<td>5</td>
<td>Many ideas were new to me.</td>
</tr>
<tr>
<td>6</td>
<td>I think ideas that were presented can be used in practice.</td>
</tr>
<tr>
<td>7</td>
<td>I felt engaged during the workshop.</td>
</tr>
<tr>
<td>8</td>
<td>I am satisfied with my own contribution during the workshop.</td>
</tr>
<tr>
<td>9</td>
<td>I think the method is useful.</td>
</tr>
<tr>
<td>10</td>
<td>I kept ideas to myself during the workshop due to the presence of competitors.</td>
</tr>
</tbody>
</table>

In Figure 7.2, the results indicate that the supplier representatives kept more ideas to themselves than the other subgroups (statement 3). This may, however, be misleading, since only two participants reported that they represented sup-
pliers. Of these two, one agreed to a great extent with the statement, pulling up the average, whereas the other did not award a conspicuous score for this statement. On the other hand, one of the participants representing an IM awarded a conspicuously high score for this statement, and two other IM representatives awarded a fairly high score compared to the other participants. It was not expected that the IM representatives would keep ideas to themselves, and rather the opposite was expected. One reason could be that these individuals had had novel ideas presented to them by manufacturers, for example, and had agreed to be confidential about these ideas. The supplier representatives on average agreed to a much greater extent than the other subgroups with the statement that they would have come up with more ideas if they had had more time (statement 2), and both the participants representing a supplier agreed to a similar degree with that statement.

No significant correlation could be found between the statements in any pair of statements, the linear regression between the statements in most pairs of statements had an $R^2$ below 0.1 and the highest $R^2$ was found to be 0.36. This implies that the correlation (of $R^2=0.64$) between how engaged a participant felt during the workshop and how satisfied that participant was with his own contribution which was found in the OptiKrea Group could not be verified in the questionnaire results from the In2Rail workshop (an $R^2=0.20$ was found).

Figure 7.3 shows that there is a positive correlation between the extent to which the OptiKrea participants and the extent to which the In2Rail participants agreed on average with the statements in the questionnaire. This is an indication that the views of the participants in the OptiKrea project can be generalized to other cross-functional inter-organizational groups within the railway sector.

![Figure 7.3](image_url)

**Figure 7.3** Linear fit ($R^2=0.83$) between the extent to which the OptiKrea participants and the extent to which the In2Rail participants on average agreed with each statement in the questionnaire.
Fifteen participants made free-text comments in the questionnaire. The comments about the method were mainly positive. A number of themes emerged. Examples of quotes representing these different themes are shown in Table 7.9 for replies to the question “I liked these things about the method” and in Table 7.10 for replies to the question “I think that the following things did not work well when applying the method”.

Table 7.9 Examples of quotes from the free-text comment field in the questionnaire, regarding things which the participants liked about the method, and relating to different themes that were mentioned by at least two participants (the number of participants mentioning the theme is given in brackets after each theme).

### I liked these things about the method

**Structure of method (5):**
- The structure helps stimulate conversations and exchange of ideas.
- Structured process.
- The structure helps people to concentrate on their own ideas and at the same time have the possibility of listening to other people.

**Contributions from several participants in idea generation (5):**
- The possibility of sharing ideas step by step.
- Reacting to and improving others’ ideas.
- The formation of ideas with contributions from several people. This works better here than with brainstorming or the discussion of ideas.

**Sharing of information/ideas (4):**
- The sharing of ideas and concepts in real time between all the participants.
- One adopts a mindset where all ideas are welcome and have a place.
- It avoided the risk of ideas not being proposed.

**Different competences/functional backgrounds among the participants (3):**
- Useful way to discuss different perspectives.
- Utilised the critical knowledge of those present.
- Opens discussion in unexpected areas.

**Involvement of all the participants (2):**
- Participation and consideration by everyone in the session.
- Everybody is involved.

**Time pressure (2):**
- Time pressure (“we have a limited amount of time!”).
- Limited time for each task.

Five participants gave positive comments connected to the structure of the method. Five participants made positive comments relating to contributions coming from several participants in the generation of ideas. Four participants
made comments relating to the sharing of information/ideas in the group. Three people made positive comments concerning the different competencies/perspectives of the participants in the group. Two participants mentioned the involvement of all the participants as a positive feature.

Table 7.10 Examples of quotes from the free-text comment field in the questionnaire, regarding things which the participants thought did not work well when applying the method, and which concerned different themes that were mentioned by at least two participants (the number of participants mentioning each theme is given in brackets after the theme in question).

<table>
<thead>
<tr>
<th>Things I do not think worked well with the method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ideation session preparation (4):</strong></td>
</tr>
<tr>
<td>- The topic analysis probably needed more preparation and more focus should be aimed at the objective of the session (i.e. ideation and not detailed design).</td>
</tr>
<tr>
<td>- More guidance could be given regarding the specification of goals, for example as to the difference between effect &amp; function goals. We need to be more specific (as in the SMART goals, for example).</td>
</tr>
<tr>
<td>- It would have helped to have had sufficient time for a short presentation of the current status and for identification of the importance criteria for judging a good solution.</td>
</tr>
<tr>
<td><strong>Too little time (3):</strong></td>
</tr>
<tr>
<td>- A total time of six hours including lunch is too short.</td>
</tr>
<tr>
<td>- Too little time at the end.</td>
</tr>
<tr>
<td>- Not enough time.</td>
</tr>
<tr>
<td><strong>Training (2):</strong></td>
</tr>
<tr>
<td>- Some confusion about moving between phases; not sure exactly how we were moving on, or what we were moving on to. Perhaps some really basic examples would be good.</td>
</tr>
<tr>
<td>- Maybe one can have a warm-up round with a simple case so that the participants can begin to find their footing (in case they do not know each other from before).</td>
</tr>
<tr>
<td><strong>Too many participants in the groups (2):</strong></td>
</tr>
<tr>
<td>- Too many people?</td>
</tr>
<tr>
<td>- Maybe too many people in the group. I think six would be better.</td>
</tr>
</tbody>
</table>

Two participants mentioned the time pressure during the application of the method as positive. On the other hand, three participants mentioned that the time was too short. In the case of two of these participants, the comments appeared to concern the time for the entire workshop. One possible interpretation of the (limited number of) comments concerning the time allocated is that it was beneficial to have a limited amount of time during each step of the method, but the total time for the workshop was too short to make it possible for the participants to wrap up its content. This interpretation would be in agreement with the other findings of the study. The perception that there was too little time for the workshop was not due to an inherent problem of the ideation method, but rather
correlated to the preparation made before the workshop. Four participants were of the opinion that the pre-workshop preparation did not work well, which is reflected in the fact that the preparation had partly to be done under time pressure during the workshop instead of being completely prepared by each task group before the workshop.

Two participants suggested performing a training exercise before starting the actual ideation session and working on the ideation topic selected for the session. This indicates that they did not find the application of the method in the group to be straightforward. Finally, two participants considered the number of participants in each group to be too high. This is not related to the method itself, but affects how it was executed.

7.2.2.3 Ideation outcome

Since Stage 3 and 4 were used for summarizing and categorizing the outcome of Stage 1 and 2, rather than generating new suggestions, this section concerns the content from Stage 1.

Table 7.11 shows how the pieces of content generated by each group are distributed between the main categories specified in Table 7.6.

<table>
<thead>
<tr>
<th>Table 7.11 Number of pieces of content from Stage 1 in the different main categories according to group (G=group).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
</tr>
<tr>
<td>Topic analysis and requirements</td>
</tr>
<tr>
<td>Suggestions on how to address the ideation topic</td>
</tr>
<tr>
<td>Irrelevant content</td>
</tr>
</tbody>
</table>

*Suggestions with the character of a list are not included here.

Irrelevant content only appeared once. The conclusion can therefore be drawn that the participants focused on the ideation topic at hand. Some content concerned topic analysis, including the discussion of goals and requirements for the solution to be developed. This content is relevant to the ideation topic, but should ideally have been treated in the topic analysis phase before the ideation started. Therefore, this content is a sign that the topic was not adequately clarified and the goals of the workshop were not adequately specified before the ideation started, which is also reflected in the fact that most participants had not contributed to the topic analysis and it was not possible to assign enough time for a thorough problem clarification during the workshop. This was also reflected in free-text comments made by the participants in the questionnaire. Nevertheless, most of the content concerned suggestions on how to address the ideation topic.

Group 1 worked with novel mechanisms for locking turnouts, and dealt with the three functions of actuation, locking and detection. Table 7.12 shows that most suggestions concerned locking and detection, and only one suggestion
addressed actuation. Six suggestions involved the performance of two or all of the three functions with the same feature. During the discussion of the topic before ideation, the group focused on a frequent problem involving a stop signal being triggered at a turnout without any apparent fault having occurred. It seems reasonable to surmise that this led the group to focus on ways to lock the turnout and detect whether it is in the correct configuration, in order to avoid traffic disturbances, and this might explain why actuation received less attention. Another possible reason could be that the participants in the group had more knowledge in the area of locking and detection.

Table 7.12 Number of suggestions generated by Group 1 in Stage 1, according to subcategory and function.

<table>
<thead>
<tr>
<th>Function</th>
<th>Physical (concrete)</th>
<th>Organizational</th>
<th>Data management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actuation</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Locking</td>
<td>10</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Detection</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shared</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*No abstract physical suggestions were generated by Group 1.

Table 7.13 shows the distribution of the suggestions generated by Group 2 between the subcategories. In total, 178 different measurable properties and 34 measurement techniques were given by Group 2 (dispersed among all the suggestions).

Table 7.13 Number of suggestions generated by Group 2 during Stage 1, according to subcategory.

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggestions in list-form</td>
<td>21</td>
</tr>
<tr>
<td>Suggestions in list-form with additional information such as implications, questions and objections</td>
<td>19</td>
</tr>
<tr>
<td>Physical suggestions</td>
<td>33</td>
</tr>
<tr>
<td>abstract</td>
<td>1</td>
</tr>
<tr>
<td>concrete</td>
<td>32</td>
</tr>
<tr>
<td>Organizational suggestions</td>
<td>4</td>
</tr>
<tr>
<td>Data management suggestions</td>
<td>6</td>
</tr>
</tbody>
</table>

Group 3 and 4 worked on the same topics during the same length of time in Stage 1. From Table 7.11 it can be seen that Group 3 generated 40 suggestions, while Group 4 generated 63 suggestions, despite the fact that Group 3 had eight
participants and Group 4 had seven participants. In Table 7.14 the distribution of these suggestions between the two tasks and the different subcategories and groups is shown. It can be observed that Group 4 generated a considerably higher amount of suggestions concerning the radical design of a turnout compared to Group 3. Group 3 actually generated more suggestions on self-adjustment of the turnout than Group 4, but the difference in quantity here is not nearly big enough to make up for the big difference found in the case of radical turnout design. The following are possible explanations for the difference in the quantity of generated suggestions between Group 3 and Group 4.

- Compared to Group 4, Group 3 might have put more effort into creating suggestions on self-adjustment, and this might have been a more difficult topic to generate ideas on.
- Group 3 might have generated suggestions with many more inherent ideas and details than Group 4, which is not captured in the counting of suggestions.
- The personalities and knowledge of the participants differed between the groups.

Overall, the groups generated mostly concrete physical suggestions. However, Group 3 and 4 generated more abstract suggestions than Group 1 and 2, which is probably due to the nature of the ideation topics.

Table 7.14 Number of suggestions generated by Group 3 and 4 in each subcategory according to group and ideation topic (G=Group).

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Radical design</th>
<th>Self-adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G3</td>
<td>G4</td>
</tr>
<tr>
<td>Physical suggestions</td>
<td>21</td>
<td>49</td>
</tr>
<tr>
<td>abstract</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>concrete</td>
<td>17</td>
<td>47</td>
</tr>
<tr>
<td>Organizational suggestions</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Data management suggestions</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>53</td>
</tr>
</tbody>
</table>

“Suggestions” in the present study correspond to “concepts” in the tests of the developed method in the OptiKrea Group (see Chapter 6) and the terms are regarded as synonymous in the present chapter, although slightly different guidelines for counting “concepts” were used in the OptiKrea tests. Table 7.15 compares the number of suggestions generated during Stage 1 in the In2Rail groups with the number of concepts generated in the OptiKrea Group. Since the time used and the number of participants varied between the groups, the suggestion generation rate per participant has been calculated. One participant in Group 2 arrived 7.5 minutes into Stage 1 and it is uncertain how this affected the quantity of suggestions. However, calculating the suggestion generation rate
per person for Group 2 for eight participants instead of nine only changes the rate slightly from 0.08 to 0.09. On the other hand, the fact that Group 2 used an unknown part of their time for generating suggestions in list-form makes it less meaningful to compare their suggestion generation rate with the rate of the other groups. Group 4 and the OptiKrea Group had approximately the same suggestion generation rate per person, whereas the other groups in the In2Rail project had a lower rate than the OptiKrea Group. At least four different aspects are important to highlight with respect to this comparison.

- Previous research on brainstorming has found that larger groups are less productive (Mullen et al., 1991), and a similar effect might exist during non-verbal ideation.
- The participants in the OptiKrea Group had worked together during the development of the method and had previously tried similar ideation methods together, and therefore they were more familiar with each other and the method, which probably had a positive effect on the number of generated concepts.
- The ideation topics were different and might have given different opportunities for generating suggestions.
- In the OptiKrea Group, all the participants used their native language when applying the ideation method, whereas in the In2Rail groups the majority of the participants did not. Generating and discussing ideas in another language than one’s native language might be more difficult.

Table 7.15 Comparison of the quantity of suggestions and the suggestion generation rate between the In2Rail groups and the OptiKrea Group (G=Group, OKG=OptiKrea Group).

<table>
<thead>
<tr>
<th></th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
<th>OKG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of suggestions*</td>
<td>35</td>
<td>43</td>
<td>40</td>
<td>63</td>
<td>77</td>
</tr>
<tr>
<td>Time (minutes)</td>
<td>60</td>
<td>60</td>
<td>50</td>
<td>50</td>
<td>63</td>
</tr>
<tr>
<td>Number of participants</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Suggestion generation rate (suggestions/minute)</td>
<td>0.58</td>
<td>0.72</td>
<td>0.80</td>
<td>1.26</td>
<td>1.22</td>
</tr>
<tr>
<td>Suggestion generation rate per person (suggestions/minute/person)</td>
<td>0.08</td>
<td>0.08</td>
<td>0.10</td>
<td>0.18</td>
<td>0.20</td>
</tr>
</tbody>
</table>

*Suggestions with the character of a list are not included (Group 2).
7.3 The UPPSAMT field test

7.3.1 Method

7.3.1.1 Participants and groups

Twenty-three people participated in the UPPSAMT workshop, and one of them was also a member of the OptiKrea Group. The participants were asked to answer the same questions as those in the In2Rail workshop (see Section 7.2.1.1 and Appendix 7.1) regarding their background and prior experience of ideation methods. Under question 1, the option “Maintenance contractor” was changed into the two alternatives “Contractor: infrastructure maintenance” and “Contractor: rolling stock maintenance”, and “Train operator” was replaced by the two alternatives “Train operator: passengers” and “Train operator: freight”. The questionnaire was written in Swedish and was filled in by 22 participants (one participant had to leave the workshop during the application of the ideation method and did not fill in the questionnaire, which was handed out after the ideation session); the results are shown in Table 7.16. The 12 participants who had used ideation methods previously were asked what ideation methods they had experience of. In addition to the participant who had been a member of the OptiKrea Group, there were three participants who reported having tried other methods that had some similarities to the developed method.
Table 7.16 Participants’ replies to questions about their background and prior experience of ideation methods.

1. What type of organization do you represent?

<table>
<thead>
<tr>
<th>Organization</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure manager</td>
<td>8</td>
</tr>
<tr>
<td>Contractor: infrastructure maintenance</td>
<td>3</td>
</tr>
<tr>
<td>Contractor: rolling stock maintenance</td>
<td>1</td>
</tr>
<tr>
<td>Train operator: passengers</td>
<td>4*</td>
</tr>
<tr>
<td>Academia</td>
<td>2</td>
</tr>
<tr>
<td>Another organization</td>
<td>4</td>
</tr>
</tbody>
</table>

2. How many years of experience do you have from the railway sector?

<table>
<thead>
<tr>
<th>Experience</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>2</td>
</tr>
<tr>
<td>5-10</td>
<td>4</td>
</tr>
<tr>
<td>10-15</td>
<td>5</td>
</tr>
<tr>
<td>15-20</td>
<td>3</td>
</tr>
<tr>
<td>20-25</td>
<td>3</td>
</tr>
<tr>
<td>&gt;25</td>
<td>5</td>
</tr>
</tbody>
</table>

3. Have you used ideation methods before?

<table>
<thead>
<tr>
<th>Answer</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>12</td>
</tr>
<tr>
<td>No</td>
<td>10</td>
</tr>
</tbody>
</table>

4. If you answered “Yes” to Question 3, how often have you used ideation methods?

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have used them a few times</td>
<td>7</td>
</tr>
<tr>
<td>A few times per year</td>
<td>3</td>
</tr>
<tr>
<td>Every month</td>
<td>0</td>
</tr>
<tr>
<td>Every week</td>
<td>0</td>
</tr>
<tr>
<td>Other frequency</td>
<td>2</td>
</tr>
</tbody>
</table>

*One participant representing a train operator did not fill in the questionnaire. Therefore, the total number of participants representing a train operator for passenger trains was 5.

The participants were assigned to different groups by the UPPSAMT project leader. The aim of this assignment was to obtain a mix of competences in each group, and to ensure that the participants were from different organizations and that there was at least one expert on overhead contact wires in each group. There were three groups with six participants and one group with five participants. Table 7.17 shows the details of each UPPSAMT group, as well as the details of the OptiKrea Group for comparison. The person who had been a member of the OptiKrea Group was assigned to Group 2.
Table 7.17 Details of the groups (G=Group, OKG=OptiKrea Group).

<table>
<thead>
<tr>
<th>Number of participants representing</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
<th>OKG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure manager</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Supplier</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Academia</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Contractor: Infrastructure maintenance</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Contractor: Rolling stock maintenance</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Train operator: Passengers</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other organizations</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of organizations represented</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Number of participants</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

*a*One participant left the workshop during Stage 2 of the ideation method.

*b*One participant from the OptiKrea project participated in this group.

7.3.1.2 Procedure

The workshop was prepared in close cooperation between the author, the project leader and the project group. The ideation topics emerged as a consequence of the UPPSAMT project and their scope was discussed in detail between the project group and the author during the preparation meetings, to make sure that there was a good definition and clear formulation of what the ideation groups were supposed to create ideas about. The project leader invited relevant participants that the project group thought would be able to contribute to the fulfilment of the aim of the workshop and benefit from its results. The duration of the workshop was one day. Although most of the participants were already familiar with the UPPSAMT project, the workshop started with a presentation of the project to make sure that all the participants had a common base to start from. The background to the ideation topics and the need for the workshop were explained. The participants were then introduced to the ideation topics, which were as follows:

“Main issue: Improved infrastructure condition monitoring and improved maintenance (and consequently punctuality and regularity) through the utilization of data collected by rolling stock to obtain information about the condition of the infrastructure.

Topic 1: What information can be collected from the rolling stock today with the existing configuration, and how can we use that information and to what purpose?

Topic 2: What information would we like to have in addition to the information that can be collected with the existing configuration? How can we get the rolling stock to collect and transmit that information [to the interested parties]?”
The ideation method instructions (Table 7.4) were shown (in Swedish) by means of a projector to all the participants and explained for about five minutes. Each participant received a paper copy of the instructions. The instructions had also been sent to the participants prior to the workshop. Group 1 and 2 worked in two different sections of a larger meeting room. Group 3 and 4 worked in two separate smaller meeting rooms. The ideation topics were written on a whiteboard in the smaller meeting rooms and shown by means of a projector in the larger room. The groups were provided with sheets of A3 paper, pens and tape, and applied the method on their own. The author visited each group to check on their progress once in a while, gave instructions as necessary and answered questions if there were any. Between Stage 1 and Stage 2 (see Table 7.4) there was a lunch break. During the application of the method, participants sometimes left their group, for example to go and grab a cup of coffee or take a phone call.

After executing the ideation method, the groups performed another task given by the project group, as described in Section 4.3.1.2.

7.3.1.3 Behaviour of the groups

Each group was audio-recorded and the recordings were used to analyse how the participants had interpreted the ideation method instructions and applied them to the topic. The analysis was performed in the same way as the analysis in the In2Rail workshop and as described in Section 7.2.1.3.

7.3.1.4 Views of the participants

Directly after the ideation method had been applied, each participant answered the same questionnaire as that used in the In2Rail workshop, see Section 7.2.1.4 and Appendix 5.1.

A linear regression was performed between each pair of statements in the questionnaires to find out if any correlation existed. The available questionnaire results from the UPPSAMT workshop resulted in 22 data points (22 participants each filling in one questionnaire). The same procedure was applied to all the available questionnaire data from the OptiKrea ideation workshops (concerning the tests of five different methods) and the In2Rail and UPPSAMT ideation workshops (a total of 82 data points), and to the questionnaire data from only the UPPSAMT and OptiKrea ideation workshops (a total of 52 data points).

A linear regression was performed between the average levels of agreement with the different statements in the questionnaire from the OptiKrea Group and the UPPSAMT groups, respectively. Statement 1 and 10 in the questionnaire were excluded from this linear regression since they were not considered to be dependent on the method (in addition, statement 10 was not answered by the OptiKrea Group).

The free-text fields in the questionnaire were analysed according to the procedure in Table 7.5. After the free-text comments from each field test had been
analysed separately, on different occasions with several weeks in between, the results from each analysis were compared and examined to determine whether the themes that had emerged were the same or not, and whether the comments from either group could be fitted into the themes of the other group.

7.3.1.5 Ideation outcome

The total number of suggestions generated in each group was counted according to guideline 1 and 8 (for each group) in Table 7.6.

7.3.2 Results and discussion

7.3.2.1 Behaviour of the participants

The following deviations from the method instructions occurred in the groups during the application of the ideation method.

- Group 1 had some problems operating their timer, and therefore each round was not exactly ten minutes, but very close to ten minutes (up to 32 seconds more). During the last round, Group 1 decided to stop two minutes in advance.
- Group 4 stopped their timer in advance during the fourth and fifth round of Stage 1.
- Group 1 circulated the sheets of paper during Stage 2. However, each presenter only presented the content which he himself or she herself had contributed on each sheet. Then the next person explained what they had added and so on.
- Group 2-4 did not circulate the sheets of paper or take turns to comment on what the other group members had presented during Stage 2 and 4. Most presentations were, however, followed by a discussion.
- Group 1 interpreted Stage 3 as meaning that they, as a group, should make a list of their favourite suggestions.

None of the groups seemed to understand the intention of the instruction stating that they were to circulate each presented sheet of paper directly after it was presented while taking turns to comment on what had just been presented (step c in Stage 2 and step b in Stage 4). Group 1 interpreted this instruction as meaning that each person should say the things that they had written on the sheet. This worked quite well and everything was explained well, but the drawback was that the participants did not make any comments about the overall content, but focused on what they themselves had written. On the other hand, this led to discussions. Group 3 interpreted the instructions as meaning that, after the presentation of a sheet of paper, anyone who commented should write down their comments on the sheet, but they chose not to put their interpretation into practice.
Two groups were confused by the wording of the instruction for step a in Stage 4, which states, “Each participant presents their own ideas from Stage 3.” (Table 7.4). These groups thought that it might mean that they should not present any ideas which they had written down and/or sketched on their sheet of paper which had originated in ideas from another participant than themselves during Stage 3. However, they did not put this interpretation into practice, but decided that each person should present all the ideas on the sheet of paper which they had started out with (which was the actual intent of the instructions).

Two groups stopped the timer in advance during at least one round of Stage 1. Probably the time-keeper had the impression that no one had anything more to add. This represents a trade-off between finishing when no new ideas seem to appear and trying to squeeze out some more ideas after some reflection. Anyone who had wanted to work a bit more would probably not have objected to the wishes of the rest of the group. On the other hand, it is good when groups adjust the method to their needs. The method variables will not be perfect for every situation, as the time required will depend, for example, on the type of ideation topic.

The extent varied to which the groups discussed the instructions and their interpretations of what they were supposed to do. There were some instances of the method instructions being interpreted in surprising ways, and there is a need to investigate how the instructions can be improved and possibly clarified by means of pictures. Stage 1 and 2 were mostly straightforward to perform for the groups. Stage 3 and 4 were more confusing and it was found from the free-text comments (further described in Section 7.3.2.2), that some participants found that these stages were lengthy and were characterized by the repetition of information. In these particular cases, it might have been better to terminate the ideation after Stage 2 and continue by applying the classification and concept screening methods. However, some new suggestions turned up during Stage 3 and 4 and additional relevant discussions took place.

Despite the fact that the groups performed some of the stages in other ways than intended, they all took part in engaged discussions about the ideation topic and delivered a range of suggestions on the topic after the ideation method had been applied; therefore, in this sense the application of the method fulfilled its goal.

Table 7.18 shows the ideal time consumption (as determined by the method instructions) and the actual time used for each step of the method in each group. Overall, the actual time used for each step by the groups was on average fairly close to the ideal times. The groups used a timer during the presentation and feedback stages (Stage 2 and 4), but it was used more as a reminder to round off the discussion than to cut it off immediately. The time for presentation and feedback was not always evenly distributed among the participants. This was because some of the presentations led to much longer discussions among the groups, and not because the presenter in question used much more time than the other participants. On average, the ideal time of ten minutes during Stage 2
appears to be a suitable guideline for the time needed for each participant to present and receive feedback in this case. During Stage 4, the participants used less time than had been assigned, and this was probably because they felt that many of the contributions were repetitions of what had been covered in Stage 2, and because they wanted to finish in time to be able to perform the last task.

Table 7.18 Ideal time consumption and actual time consumption of each stage in each group (G=Group).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Ideal</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>G4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of rounds</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Ideation time/round (min.)</td>
<td>10:00</td>
<td>09:53</td>
<td>10:00</td>
<td>10:00</td>
<td>09:35</td>
</tr>
<tr>
<td>Stage 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review</td>
<td>05:00</td>
<td>01:30</td>
<td>05:00</td>
<td>05:00</td>
<td>05:00</td>
</tr>
<tr>
<td>Presentation and feedback (min./participant)</td>
<td>10:00</td>
<td>10:32</td>
<td>11:34</td>
<td>09:51</td>
<td>11:55</td>
</tr>
<tr>
<td>Stage 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideation time (min.)</td>
<td>10:00</td>
<td>N/A</td>
<td>10:00</td>
<td>10:00</td>
<td>10:00</td>
</tr>
<tr>
<td>Stage 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presentation and feedback (min./participant)</td>
<td>05:00</td>
<td>N/A</td>
<td>03:42</td>
<td>03:29</td>
<td>02:32</td>
</tr>
</tbody>
</table>

aAs specified in the instructions.
bAverage time from when the participant starts presenting suggestions until when the next participant starts, including the time for feedback and discussion (excluding the time for disturbances such as a person entering the room).
cThis time is uncertain since the review took place in connection with the lunch break.
dSome of the participants discussed the topic during five minutes of this time. One participant left for about 20 minutes after five minutes.
eThe person who left during Stage 3 was still not present. The content of his sheet of paper was presented by another group member.
fThe group cut off the fourth and fifth round before ten minutes had passed.
gIt appears that the group kept on writing for a few more minutes (after the timer had beeped), but this is not evident from the audio recordings.
hTwo participants did not present their sheet of paper explicitly, but they participated in the discussion during Stage 4. It is not evident from the recordings why they did not have an explicit slot for presenting their proposals; they may perhaps have thought that their proposals had already been mentioned by someone else.
iThe group had some problems operating the timer and chose to end the last round in advance.
jOne person had to leave the group during the presentation of the fourth presenter. This person had already presented their contributions.
7.3.2.2 Views of the participants

The average level of agreement with each statement in the questionnaire for each UPPSAMT group is shown in Figure 7.4. The results from the OptiKrea Group are shown for comparison. The statements are to be found in Table 7.8.

Figure 7.4 Average questionnaire results for each UPPSAMT group and the OptiKrea Group.

The questionnaire results show that the engagement in the ideation topic was very high in all the groups. Compared to the OptiKrea Group, the participants in the UPPSAMT groups on average thought that they would have come up with more ideas if they had more time and reported having had more ideas than presented to a higher degree than the OptiKrea Group. Three of the four UPPSAMT groups reported having found many interesting ideas to be presented during the workshop to a higher degree than the OptiKrea Group. However, not that many ideas were new to the participants, except for Group 4, who on average reported to a very high degree that many ideas were new to them. Although there was some variation in the level of agreement, all the groups reported having found that many of the ideas could be used in practice. There was an average high level of engagement in all the groups, and the participants were satisfied with their own contribution and found the method to be useful. However, for statement 7, 8, and 9, it can be seen in Figure 7.4 that Group 1 had a lower average than the other groups. Focusing on the individual levels of agreement, it was found that one person in Group 1 had awarded much lower scores for these statements compared to the other group members. This person's scores also explain the lower average for Group 1, compared to the other groups, regarding statement 4. This person explained in the free-text comment...
field in the questionnaire that this way of working (and other similar methods) hampered their creativity. From the recordings, it was noted that this person did not contribute very many ideas, but still contributed several insights into problems and issues related to the ideation topics that helped to illuminate the measures required. On average, the participants in the groups reported a low level of agreement with the statement that they had kept ideas to themselves due to the presence of competitors.

The average questionnaire results for each type of organization represented in the UPPSAMT workshop are shown in Figure 7.5. Since the number of participants representing each type of organization was quite small (see Table 7.16), it is not possible to draw any certain conclusions from Figure 7.5.

![Figure 7.5 Average questionnaire results for each type of organization.](image)

Statement 2, “If I had had more time, I would have come up with more ideas”, had quite a large dispersion in the average scores for each type of organization. However, when studying the individual scores, one can observe that there was a larger dispersion between the participants within each type of organization than between the different types of organizations. Regarding statement 3, however, it is interesting to note that several of the participants representing the IM reported a level of agreement of more than 50% with the statement that they had had more ideas than they had presented. Nevertheless, none of the IM representatives reported a higher level of agreement than 10% with the statement that they had kept ideas to themselves because of competitors. In the case of the infrastructure maintenance contractors, two of them reported a level of agreement of more than 90% with the statement that they had had more ideas than they had presented (whereas the third participant in this group reported a very low level of agreement). One of the infrastructure maintenance contractors who said that
they had had more ideas than they had presented reported a level of agreement of 35% with the statement that they had kept ideas to themselves because of competitors, whereas the others in that group reported a very low level of agreement with the statement that they had kept ideas to themselves because of competitors participating. There is a similar pattern for the views of some of the other participants in the other groups. This raises the question as to why the participants did not share all their ideas, if this was not due to the presence of competitors. Possible reasons are that they thought that their ideas had been bad, that they had not felt confident enough to share them in the group because they were afraid to be criticized, that they had not had enough time, or that they had thought that other group members had had more important ideas.

The linear regression between different statements in the questionnaire found some correlations with some significance. How engaged a participant felt during the workshop and how satisfied that person was with their own contribution was correlated with $R^2=0.64$ (see Figure 7.6), which was close to the correlation found in the case of the OptiKrea Group ($R^2=0.68$).

![Figure 7.6](image)

**Figure 7.6** Linear fit (solid line) between the UPPSAMT participants’ average levels of agreement with the statements “I felt engaged during the workshop” and “I am satisfied with my own contribution during the workshop” ($R^2=0.64$).

Between how engaged the participant had felt and how useful they had found the method to be, a correlation of $R^2=0.59$ was found (the corresponding figure for the OptiKrea Group was $R^2=0.44$). A correlation of $R^2=0.57$ was found between a participant having found that many interesting ideas were presented during the workshop and how useful the participant had found the method to be; between how many interesting ideas had been presented and how engaged the participant had felt during the workshop, a correlation of $R^2=0.43$ was found. Finally, a correlation of $R^2=0.69$ was found between how satisfied a participant
had been with their own contribution and how useful they had found the method, see Figure 7.7.

Figure 7.7 Linear fit (solid line) between the UPPSAMT participants’ average levels of agreement with the statements “I am satisfied with my own contribution during the workshop” and “I think the method is useful” ($R^2=0.69$).

Figure 7.8 brings some clarity to the relations between the statements that were found to have an $R^2$ above 0.5 in the UPPSAMT test. The corresponding $R^2$ values for the OptiKrea and In2Rail tests are included in Figure 7.8 for comparison. In addition, the connection between “I felt engaged during the workshop” and “Many interesting ideas were presented during the workshop” is indicated for comparison in Figure 7.8. The numbers connecting each statement show the $R^2$ values for the correlations between the statements in the UPPSAMT, OptiKrea and In2Rail groups, respectively. Firstly, it can be observed from Figure 7.8 that the $R^2$ values for the correlations between the statements were low in the In2Rail groups. In the UPPSAMT groups, on the other hand, most correlations with an $R^2$ above 0.5 between statements were found. The OptiKrea Group had one correlation with an $R^2$ above 0.5. The three most significant correlations were found in a “loop” between how engaged a person felt during the workshop, how satisfied they were with their own contribution, and how useful they found the method to be. These connections suggest that the participants might have found the method useful because they had felt good during the workshop or because they had had such a good time in the group, or that the method itself might have been the reason for the participants having a good time. This pattern strengthens the findings from the OptiKrea Group (which are given in Section 6.3.2.3), where, however, only a participant’s satisfaction with their own contribution had a fairly high correlation with how
engaged a participant felt. Although the OptiKrea Group had more data points in the analysis, these points were based on 6 individuals. The UPPSAMT workshop had 22 points from 22 different participants, and in this sense might be considered to give a better data set for statistical analysis. Further, the participants in the OptiKrea Group can be considered to be comparable to the population in the UPPSAMT workshop (the Swedish deregulated railway market), whereas they were less comparable to the In2Rail groups, which had representatives from several different EU railway markets. However, the findings from the UPPSAMT groups indicate that the number of interesting ideas affected their satisfaction with the method and their satisfaction with their own contribution and possibly how engaged they were, which was not found in the OptiKrea Group. One possible interpretation of this is that the larger number of different participants in the UPPSAMT workshop gives a more valid result and that this pattern was not found in the OptiKrea Group because the number of participants was too small. At least three things differentiated the In2Rail groups from the OptiKrea Group and the UPPSAMT groups and might explain why corresponding findings were not made in the different test contexts. Firstly, the In2Rail groups consisted of individuals who came from different EU railway markets (which are deregulated to different extents and have different cultures) and not solely from the Swedish market. Secondly, the size of the groups during the In2Rail workshop was larger than that during the OptiKrea and UPPSAMT workshops; this might have affected how well the group cooperated, how “intimate” the workshop felt and, consequently, the experience which the partici-

Figure 7.8. The relations between the statements that were found to have an $R^2$ above 0.5 in any of the test contexts. The first figure is the UPPSAMT $R^2$ value, the figure in parenthesis is the OptiKrea $R^2$ value and the figure in bold is the In2Rail $R^2$ value.
pants had, among other things. Thirdly, the majority of the participants in the In2Rail workshop did not take part in their native language. In addition, the OptiKrea and UPPSAMT workshops dealt with different types of ideation topics. To summarize, it can be concluded that a fairly strong correlation between how engaged a person had felt during the workshop and how satisfied they had been with their own contribution was found in two of the three test contexts. However, there are many question marks regarding the cause of the results of this analysis and the analysis suffers from the fact that there is a small number of participants who awarded a low level of agreement with most statements (or a high level of agreement in the case of statement 2 and 3), as is exemplified in Figure 7.6 and 7.7. In some cases, a correlation might not have been possible to detect due to the lack of data points with a low level of agreement. The results can, however, assist in finding directions for future research.

Taking the findings in this regard from each test context into account, it is not surprising that a correlation of $R^2=0.66$ was found between how engaged a person had felt during a workshop and how satisfied they had been with their own contribution when applying linear regression only to the questionnaire results from the UPPSAMT and the OptiKrea tests. Between how engaged a person had felt and how useful the method had been found to be, an $R^2=0.47$ was found.

Further, applying linear regression to the questionnaire data from all the tests taken together (the In2Rail, UPPSAMT and OptiKrea tests) did not provide any interesting results. The only correlation between statements with an $R^2$ above 0.5 was that between how engaged a person felt during a workshop and how useful they found the method ($R^2=0.51$), which is not surprising, since this correlation was fairly high in the OptiKrea Group and the UPPSAMT groups, but was low in the In2Rail groups.

Figure 7.9 shows the linear fit between the extent to which the OptiKrea participants and the extent to which the UPPSAMT participants on average agreed with each statement in the questionnaire. A correlation of $R^2=0.84$ was found. This indicates that there is a good agreement between the views of the OptiKrea participants and the UPPSAMT participants regarding the method, and indicates that the method can be found to be useful by other similar groups than the group in which it was developed, and for other types of ideation topics.
Figure 7.9 Linear fit ($R^2=0.84$) between the extent to which the OptiKrea participants (the average of the group) and the extent to which the UPPSAMT participants (the average of the groups) agreed with each statement in the questionnaire.

Twenty participants made free-text comments in the questionnaire which covered quite a wide range of different opinions and were mostly positive. Some themes emerged, and examples of quotes representing these different themes are shown in Table 7.19 for replies to the question “I liked these things about the method”, in Table 7.20 for replies to the question “I found that the following things did not work well when applying the method”, and in Table 21 for comments written in the field “Other comments”.

Eight participants made comments that were connected to the advantages of sharing information and building upon other group members’ ideas. Six participants mentioned things connected to the structure of the method as positive. Four participants stated that the discussions that had taken place in connection with the presentation of ideas had been a positive outcome of applying the method. The sharing of information was regarded by three participants as a positive feature of the method.

Four participants mentioned that the same ideas/points had been repeated during some of the steps and the procedure had become lengthy. Two participants mentioned things that concerned the screening of concepts to prioritize some suggestions. However, such concept screening is not an inherent part of the ideation method which they tested.

Three participants mentioned that they had found the group composition to be good. This is a feature of the workshop which is not an inherent part of the method, but the group composition is likely to affect how effective the method is. Finally, two participants mentioned that they had had fun during the applica-
tion of the method. This does not provide any indication of how effective the method is regarding the production of suggestions, but is important, since one of the goals in the development of the method was to make it attractive to its users.

An interesting idea put forward by one of the participants was to mix the participants into new groups between Stage 2 and 3 or perhaps between Stage 3 and 4. In this way, it would be possible to receive even more viewpoints on the ideas that were taken forward through the different stages. This would probably also avoid the repetition and lengthiness of the method experienced by some of the participants.

Table 7.19 Examples of quotes from the free-text comment field in the questionnaire, regarding things which the participants liked about the method and which relate to different themes that were mentioned by at least two participants (the number of participants mentioning each theme is given in brackets after the theme in question).

<table>
<thead>
<tr>
<th>I liked these things about the method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Building upon other group members’ ideas (8):</strong></td>
</tr>
<tr>
<td>• Being able to make additions to the ideas of others.</td>
</tr>
<tr>
<td>• Other people’s ideas gave rise to new ideas or reminded one of things one had forgotten to bring up.</td>
</tr>
<tr>
<td>• The idea of refining each other’s suggestions I found rather brilliant. :)</td>
</tr>
<tr>
<td><strong>Structure of method (6):</strong></td>
</tr>
<tr>
<td>• Efficient. Enough time to come up with ideas, but without having the time to criticize one’s own ideas too much. :)</td>
</tr>
<tr>
<td>• Clear-cut conditions. Mixed competence. Time control.</td>
</tr>
<tr>
<td>• More structure in the idea generation since one starts with individual work.</td>
</tr>
<tr>
<td><strong>Discussion in connection with the presentation of ideas (4):</strong></td>
</tr>
<tr>
<td>• Good discussions, exchange of ideas and experience, and insight into different areas of interest.</td>
</tr>
<tr>
<td>• It turned out that there were good “catalytic” discussions. Generated descriptions of possibilities on a suitable level.</td>
</tr>
<tr>
<td>• It turned out that there were many interesting discussions.</td>
</tr>
<tr>
<td><strong>Sharing of information (3):</strong></td>
</tr>
<tr>
<td>• The exchange of information and ideas.</td>
</tr>
<tr>
<td>• Good way to get everybody to speak, to extract all the ideas.</td>
</tr>
<tr>
<td>• The combination of individual work and checking [the contributions in the group].</td>
</tr>
</tbody>
</table>
Table 7.20 Examples of quotes from the free-text comment field in the questionnaire, regarding things which the participants thought did not work well when applying the method, and which relate to different themes that were mentioned by at least two participants (the number of participants mentioning each theme is given in brackets after the theme in question).

I found that the following things did not work well when applying the method

Ideas/points were repeated through some of the steps, and the procedure became lengthy (4):
- Too long a time spent looking at other people’s sheets of paper. The same ideas appear on all the sheets so the time [is used inefficiently].
- Somewhat tedious, especially Stage 4. A lot of repetition, but on the other hand, a number of new [ideas] appeared too, of course…
- Felt that there were a lot of recurrent points and [it was] a bit forced.

No concept screening/prioritizing (2):
- It was a bit confused during Stage 3; I feel that the group together should have prioritized a number of suggestions which we then could have processed further in the large group; in that way we could have got three ideas which everyone could have agreed upon; [we could have continued working on] the rest of the ideas in the form of a backlog.
- I rather missed having a step where the group “skimmed off the cream” of the individual suggestions, i.e. an additional refinement step.

Table 7.21 Examples of quotes from the free-text comment field in the questionnaire, regarding “other comments” on the method that relate to different themes that were mentioned by at least two participants (the number of participants mentioning each theme is given in brackets after the theme in question).

Other comments

Good group composition (3):
- Good combination of different stakeholders in the groups of the workshop.
- Good mix of participants.
- Well composed group with different (but not too different) areas of interest and ideas.

Had fun (2):
- We had fun in the room at the same time as we performed the exercises, probably thanks to the format.
- Fun way of working! Very good way to get to know each other, build relations!
In the free-text comments, three themes occurred which were touched upon by both the In2Rail and the UPPSAMT field test participants. Table 7.22 shows an overview of the common themes and the number of participants in each field test mentioning each theme.

<table>
<thead>
<tr>
<th>Theme</th>
<th>In2Rail (17)</th>
<th>UPPSAMT (21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contribution from several participants in idea generation</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Structure of method</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Sharing of information</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

About a third of the participants in each field test who had filled in the questionnaire stated that the contribution from several participants in idea generation (building on each other’s ideas) and the structure of the method were advantages of the method. This number is regarded as quite high, considering the fact that the questions were open-ended and could comprise many different types of answers. It is therefore believed that these are important properties of the method, and properties that could be interesting to follow up on in future research.

### 7.3.2.3 Ideation outcome

Table 7.23 shows the number of suggestions generated by each group during Stage 1 and Stage 3, respectively.

<table>
<thead>
<tr>
<th>Group</th>
<th>Stage 1</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of concepts</td>
<td>Number of concepts&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Group 1</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>Group 2</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>Group 3</td>
<td>47</td>
<td>27</td>
</tr>
<tr>
<td>Group 4</td>
<td>24&lt;sup&gt;a&lt;/sup&gt; (29)</td>
<td>20&lt;sup&gt;a&lt;/sup&gt; (24)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Group 4 had five participants while the other groups had six. The figure in brackets is the result for five participants normalized to the result for six participants (rounded off to the closest integer).

<sup>b</sup>Including concepts that were repeated from Stage 1.
It can be seen that the number of suggestions during Stage 1 ranges from 24 to 47. Group 4 had one participant less than the other groups, and the numbers in brackets show the number of suggestions normalized to the result for six participants. During Stage 3, Group 1 listed a much lower number of suggestions than the other groups, and this was because they performed Stage 3 jointly as one group, listing their prioritized ideas, instead of generating suggestions individually as the participants did in the other groups.

The topic of the UPPSAMT workshop was directed more towards data management, the handling of information and organizational issues than the topics that had been used in the OptiKrea project, and therefore it was not meaningful to compare the suggestion generation rates. On the other hand, the workshop demonstrated that the method could be applied successfully to this type of topic.

7.4 Comparing the outcome with expectations

There were expectations as to the outcome of the field tests. Firstly, it was expected that it would be easy to apply the ideation method. It turned out that during the In2Rail workshop, the groups were able to use the method without training, but asked several questions, especially regarding Stage 3. The groups might have been able to work more independently if the method instructions had been available to each participant all the time. Further, having the supervisor available makes it easy for the participants to ask questions, and it is possible that if the groups had been on their own, they would have found out what to do without asking. During the UPPSAMT workshop, each group worked more independently compared to the groups in the In2Rail workshop and were given paper copies of the method instructions. Each group was able to work through the method on their own, but they interpreted some parts of the instructions in another way than was intended. As in the In2Rail workshop, Stage 3 was the one that met with most confusion in the UPPSAMT workshop. There is a need to improve the method instructions and performing the last two stages of the method might not always be suitable; this appears to depend, among other things, on the nature of the ideation topic.

Secondly, it was expected that the field test participants would share the views of the OptiKrea Group participants on the method, since the latter had been chosen carefully to be representative of a typical use setting in the deregulated railway sector. A correlation of \( R^2 = 0.83 \) was found between the average questionnaire scores of the In2Rail groups and the OptiKrea Group, and between the average scores of the UPPSAMT groups and the OptiKrea Group, a correlation of \( R^2 = 0.84 \) was also found. This indicates that an agreement exists between the views of the In2Rail groups and those of the OptiKrea Group. However, a more thorough statistical analysis is required to confirm this finding, and one limitation in this connection is the small number of participants. Figure 7.10 shows a comparison of the average questionnaire scores of the par-
participants in the UPPSAMT, In2Rail and OptiKrea workshops. The statements are given in Table 7.8.

![Bar chart showing the average questionnaire scores of the participants in the UPPSAMT, In2Rail and OptiKrea workshops.]

**Figure 7.10** The average questionnaire scores of the participants in the UPPSAMT, In2Rail and OptiKrea workshops.

The OptiKrea participants on average judged the method to be more useful and were more engaged during the workshop and more satisfied with their own contribution than the In2Rail and UPPSAMT participants did and were on average. This might be explained by the fact that the participants in the OptiKrea Group were more familiar with each other and similar ideation methods compared to the participants in the In2Rail and UPPSAMT workshops. Another possible explanation is that the OptiKrea Group was positively biased towards the method since they had taken part in its development. Interestingly, the participants in the UPPSAMT workshop on average had higher scores than the In2Rail participants regarding these three statements. From Figure 7.10, it is found that, compared to the OptiKrea Group, the participants in both the UPPSAMT and In2Rail workshops agreed on average to a greater extent with the statement that they would have been able to come up with more ideas if they had had more time. This is interesting (although all the average scores are quite low) since the participants in these workshops were aversive towards using Stage 3 to generate more ideas, whereas the OptiKrea Group used Stage 3 with this intention and also reported to a lesser extent that they believed that they would have been able to come up with more ideas, if given more time. In consideration of this, it seems rather odd that the participants in the field tests did not use Stage 3 to try to add more ideas. At the same time, the field test participants also reported having kept more ideas to themselves than the OptiKrea participants. In the free-text comments of the questionnaires, the field test participants focused on
other aspects of the method than the OptiKrea participants did in their questionnaires and interviews. It is believed that this is because the field test participants were trying such an ideation method for the first time and compared it to a “normal meeting”, while the OptiKrea participants, in their replies, compared the method to the methods which they had tried earlier, and especially to the earlier version of the developed method.

Thirdly, it was expected that the number of suggestions generated by the In2Rail groups would be comparable to that generated by the OptiKrea Group. Jackson and Poole (2003) found in their study of ideation procedures in groups in naturally occurring contexts that their idea generation rate was lower than the rates reported in laboratory studies. During Stage 1, one of the In2Rail groups had a suggestion generation rate per person that was comparable to that of the OptiKrea Group, whereas the other In2Rail groups had lower rates than that of the OptiKrea Group. Because the ideation topic was of a different nature during the UPPSAMT workshop, the number of suggestions generated in that workshop was not compared to that generated by the OptiKrea Group. It was found, however, that the method worked well when applied to this different type of ideation topic.

These findings suggest that developing ideation methods for a certain context in a small group can be feasible, but a prerequisite may be that the participants should be carefully chosen to form a group that is representative of the context. It would be very interesting to perform future studies where the outcome of methods developed among non-representative participants (e.g. students) would be compared to the outcome of methods developed in a representative group after both these types of methods have been applied in a typical use setting. Nonetheless, it can be concluded that testing the present method in a typical use setting which was different from the development context provided valuable insights into the method and suggestions for future research directions.

The findings of the two field tests did not provide any evidence which indicated that the design principles for ideation methods to be used in cross-functional inter-organizational groups that were formulated in Chapter 6 (see Table 6.4) should be modified. Rather, insofar as it was possible to judge from the field tests, the design principles formulated in Chapter 6 form a relevant basis for designing such ideation methods. However, the application of the method developed on the basis of the design principles in the two field settings gave insights into further improvements regarding how the design principles should be most efficiently put into practice.

The research in Stage 4 has also highlighted challenges that arise when applying a method outside an experimental environment. The primary goal of the participants in the field test workshops was not to provide the researcher with data, but to accomplish one of their project goals, and as a researcher one must respect this to obtain access to this setting. Despite rigorous preparations before the In2Rail workshop, a number of deviations from the planned method had to be made in order to complete the workshop, which affected the robustness of
the study. Partly due to the learning derived from the In2Rail workshop, the UPPSAMT workshop could be performed with a higher level of robustness, although it is still far from a controlled laboratory study. However, such studies are still worthwhile, since previous research on idea generation (Sutton and Hargadon, 1996; Jackson and Poole, 2003; Gish and Hansen, 2013) suggests that, in order to understand how ideation can be successful in reality, the context must be taken into account.

7.5 Conclusions

Overall, Stage 1 and 2 were applied as intended, whereas Stage 3 and 4 were partly used in other ways than intended.

The participants in the field test groups were satisfied with using the method. During the In2Rail workshop, most of the contents from the ideation sessions consisted of suggestions on how to address the ideation topic and the groups generated mostly concrete physical suggestions. There was a large dispersion in the number of suggestions (or the suggestion generation rate per person) between the groups within each field test.

The suggestion generation rate per person of one of the In2Rail groups during Stage 1 was comparable to that of the OptiKrea Group, whereas the other three In2Rail groups had lower rates. A positive correlation was found between the views of the OptiKrea Group participants and the views of the field test group participants regarding the method. However, the OptiKrea participants on average judged the method to be more useful and were more engaged during the workshop and more satisfied with their own contribution than the In2Rail and UPPSAMT participants did and were on average.

Some findings from the development of the OptiKrea Group could be verified in the field tests. Evaluation of the method in a typical use setting gave additional insights into the method and revealed possibilities for further refinement.
Appendix 7.1

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) What type of organization do you represent?</td>
<td>Infrastructure manager, Supplier, Maintenance contractor, Train operator, Academia, Other:</td>
</tr>
<tr>
<td>2) How many years of experience do you have from the railway sector?</td>
<td>0-5, 5-10, 10-15, 15-20, 20-25, &gt;25</td>
</tr>
<tr>
<td>3) What is your area of expertise?</td>
<td></td>
</tr>
<tr>
<td>4) Have you used ideation methods before?</td>
<td>Yes, please proceed to 5), No</td>
</tr>
<tr>
<td>5a) If Yes on 4), can you mention what methods you have used?</td>
<td></td>
</tr>
<tr>
<td>5b) How often have you used ideation methods?</td>
<td>I have used them a few times, A few times per year, Every month, Every week, Other:</td>
</tr>
</tbody>
</table>

Comments on questions 1-5:
8. Development of concept selection method

8.1 Introduction

This chapter presents Stage 5 of the research conducted for this thesis (see Table 1.1). The goal was to develop a practical collaborative concept selection method that the Swedish Transport Administration (STA) could use to decide on which concept(s) to take forward in a product development or procurement process. The modified Pahl and Beitz (MPB) method (see Section 2.3.1.3) was chosen as the starting point for the development since the MPB method is an adaption of the well-known method by Pahl and Beitz (Pahl et al., 2007) to the needs of the STA’s predecessor the Swedish Rail Administration. The method developed by Pahl and Beitz is systematic, and it is able to combine greatly different aspects that are not related to each other. However, the developer of the MPB method suspected that the method had become too cumbersome since variants of it had been presented to about 20 different organizations and no one had implemented it, despite the fact that all these organizations had been positive towards it. Further tests could reveal why it had not been implemented.

This research stage had the following objectives:

- to apply the MPB method in the OptiKrea Group to explore and evaluate how it works in practice;
- to use the findings from the application of the MPB method to develop an improved method for the context of the Swedish deregulated railway market;
- to apply and evaluate the developed method in a case study;
- to find out how the diversity of the group would manifest itself during the application.

To meet the objectives, the following research questions were posed.

- What problems (if any) are associated with the MPB method with respect to the given context?
- How are these problems addressed in the developed method?
- How does the diversity of the group manifest itself?
- What do the participants think about the MPB method and the developed method?
- What are the general design principles for the developed method?
8.2 Method

8.2.1 Procedure

The OptiKrea Group (see Section 3.4) applied the MPB method during a one-day workshop. The workshop started with an introduction to and discussion around concept selection methods in the specific context of the Swedish deregulated railway market. Thereafter the method was applied to the topic of track geometry in turnouts and to six random concepts that had been generated during ideation workshops on this topic earlier during the project (the concepts were drawn by one of the participants from a box containing all the concepts that had been generated during the ideation session on the same topic). These concepts were quite different and had been developed to a different extent. The actual application of the method took about 2.5 hours. This was quite a short time for an application of the MPB method, and therefore the focus was on completing all the steps, limiting the number of calculations and evaluated concepts, and, rather, exploring how the principles of the method worked in practice. The application of the method was supervised by Prof. Jan Lundberg of Luleå University of Technology, the developer of the MPB method.

The application of the MPB method was evaluated (see Section 8.2.2) and the findings from the evaluation were used to set up a number of requirements for the method to be developed, as well as to make suggestions as to how the MPB method could be modified to address these requirements. The requirements and modification suggestions were presented to and approved by the OptiKrea Group.

The newly developed method was then tested in the OptiKrea Group two months after the MPB method had been tried. The application was performed during a two-day workshop. During the first day, the new method was explained to the group and discussed, and the problem clarification stage was executed. During the second day, the group performed the goal specification, requirement specification, concept screening and concept scoring stages. The total time used for applying the method was approximately six hours. The group continued working on the topic of track geometry in turnouts during this workshop. After going through the problem clarification, goal specification and requirement specification stages, the group was provided with five concepts (V1-V5) which had been conceptually developed by Jan Lundberg to a similar extent as possible solutions to the topic within the scope of another project that also concerned turnouts (“Improved Availability and Reduced Life Cycle Cost of Track Switches”). These concepts were used during both the concept screening and the concept scoring stages. The application of the developed method was supervised by the author.
8.2.2 Evaluation

8.2.2.1 The behaviour of the participants

The entire workshops were audio-recorded and the recordings were used to analyse the behaviour of the participants during the workshops. The procedure is described in Table 8.1. Being present at the workshops allowed the author to act as an observer, although during the second workshop, this role was combined with the role of supervising the workshop and not systematically fulfilled during the entire workshop. However, if observations were made of participants behaving unexpectedly or in a manner that was interesting in other respects, this was noted down. The notes were used to support the analysis of the recordings. The recordings were, however, the main source of the findings regarding the behaviour of the participants.

Table 8.1 The procedure used to analyse the audio recordings from the application of the concept selection method

<table>
<thead>
<tr>
<th>Step 1:</th>
<th>The recordings of all the ideation sessions were transcribed by the author.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 2:</td>
<td>The recordings were listened to while the transcriptions were simultaneously being perused. The excerpts that were interesting with respect to the research questions were marked. Notes were made about the author’s reflections concerning these data, for the purpose of assisting in defining categories (themes).</td>
</tr>
<tr>
<td>Step 3:</td>
<td>Step 2 was repeated for each application or part of an application one or several times, depending on the complexity of the text. (Sometimes this involved either listening to the recordings or reading the transcript.)</td>
</tr>
<tr>
<td>Step 4:</td>
<td>All the marked excerpts were collected in a single word processing document. For each quote contained in an excerpt, information was provided as to which participant had contributed the quote.</td>
</tr>
<tr>
<td>Step 5:</td>
<td>Excerpts that were found to represent the same phenomena were listed together. Tentative categories were formed. The data were perused as many times as required to find exhaustive and mutually exclusive categories (i.e. so that all the quotes could fit into a category and no quotes could fit into more than one category). The categories were given descriptive labels and a summary of each category was written.</td>
</tr>
</tbody>
</table>

For the second and more comprehensive workshop, it was investigated whether the participants had contributed different viewpoints. To this end, the contributions from each participant during the topic clarification and requirement specification stages were compared. Firstly, all the contributions were identified from the workshop material and the transcribed recordings from the application of the method. A contribution made in the topic clarification stage was identified as a “piece of information”, which in turn was defined as an input from one of the participants that was limited to one matter. A contribution made in the requirement specification stage was identified as a suggested requirement. Examples of raw data in the form of quotes that were counted as pieces of in-
formation are to be found in Table 8.2-8.4; see Table 8.8 for a list of the topic clarification questions.

### Table 8.2 Examples of replies to question 2: “Who wants the problem to be solved and why?”

<table>
<thead>
<tr>
<th>Who</th>
<th>Why</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>STA</td>
<td>Wants to provide a reliable railway.</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>An irregular track geometry generates impacts that deteriorate the turnout.</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>Wants a predictable deterioration so that maintenance work can be planned.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Wants a method for performing track geometry adjustment with a low LCC.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Maintenance costs are too high and traffic standstill is too long.</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>They have to pay to fix the failures.</td>
<td>F</td>
</tr>
<tr>
<td>Train operators/</td>
<td>Do not want speed reductions or disruptions in traffic.</td>
<td>F</td>
</tr>
<tr>
<td>the customers of STA</td>
<td>Rolling stock breaks down.</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Make demands on comfort and safety.</td>
<td>C</td>
</tr>
<tr>
<td>Contractors</td>
<td>Viability, be able to carry out calculable maintenance with the right competence. It should be clear to the machine drivers what they are supposed to do.</td>
<td>D</td>
</tr>
</tbody>
</table>

### Table 8.3 Examples of replies to question 3: “What are the (root) causes of the problem?”

<table>
<thead>
<tr>
<th>Cause</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Settlements in the substructure/subballast. Drainage problems and similar problems.</td>
<td>A</td>
</tr>
<tr>
<td>Irregularities in the steel in the turnout propagate down into the macadam, which in turn deteriorates, resulting in instability. (Also, one needs to find the right standard level of maintenance of steel and tamping.)</td>
<td>B</td>
</tr>
<tr>
<td>The deterioration of the system is due to the load from usage, which leads to deformations. Uncertain where the load originates.</td>
<td>C</td>
</tr>
<tr>
<td>The geometry changes so that the ballast pressure is exceeded.</td>
<td>D</td>
</tr>
<tr>
<td>Either the tamping methodology or the engineering design, together with the design of the area of the sleepers, or a combination.</td>
<td>E</td>
</tr>
<tr>
<td>Different load distributions between the turnout and the ordinary track. Different tamping methods for the turnout and the ordinary track.</td>
<td>F</td>
</tr>
</tbody>
</table>
Table 8.4. Examples of replies to question 12: “What are the future trends regarding technology, environmental aspects and ergonomics?”

<table>
<thead>
<tr>
<th>Trend</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modularisation</td>
<td>B</td>
</tr>
<tr>
<td>Short time on the track for maintenance and installation</td>
<td>B</td>
</tr>
<tr>
<td>Lower energy usage and environmental emissions during manufacture and operation</td>
<td>D</td>
</tr>
<tr>
<td>Recyclability</td>
<td>D</td>
</tr>
<tr>
<td>The self-inspecting turnout</td>
<td>E</td>
</tr>
</tbody>
</table>

Contributions with the same meaning were considered to be the same. If a person mentioned a specific contribution more than once, it was only counted once for this participant. A count was made of the number of participants who had made a certain contribution. The total number of contributions per participant was counted, as well as how many contributions were made in each round of talking when the participants took turns to present their contributions. The information that was actually presented to the group during the workshop was included in this analysis, and any answers which the participants had written down but had not shared with the group were excluded. Although the participants were encouraged to provide any information which they thought of during the topic clarification and requirement specification stages, in order to capture a wide range of aspects, only the contributions whose relevance was agreed upon and which the group decided to take further were included in the analysis. Otherwise the analysis would have given an unrealistic picture of the method, as participants might have generated a large number of different contributions which would have given a positive impression in the analysis, but would actually have been useless for developing the desired product.

8.2.2.2 The views of the participants

After the test of each method, the participants answered a questionnaire and a group interview took place regarding the participants’ experience of the applied method to capture the participants’ immediate reactions to the method that they had tried. The questionnaire presented different statements about each method, and the participants assessed their level of agreement with these statements, answering on a continuous scale from “Do not agree at all” to “Agree completely”, by making a mark on a line which was about 100 millimetres in length and where “0” represented “Do not agree at all”. The position of the mark was measured in millimetres from “0” with a ruler and divided by the total length of the line. The average value for each statement was calculated and the individual results were also compared to find out if any interesting patterns were present. The questionnaire is to be found in Appendix 8.1 (translated from Swedish into English). The interview questions that were used after each method test (trans-
lated from Swedish into English) are to be found in Appendix 8.2. The group interviews were audio-recorded and analysed in the same way as described in Section 5.2.2 and Table 5.5.

8.3 Results and discussion

8.3.1 Test of the modified Pahl and Beitz method

Applying the MPB method in the OptiKrea Group revealed a number of issues that are described and discussed in this section. The two major issues were the assessment of values and the determination of weights.

8.3.1.1 Issues related to the assessment of values

In the MPB method, requirements are separated from wishes. Whereas requirements must be fulfilled, wishes are used to evaluate further the concepts that have been found to fulfil all the requirements. Since the parameters related to the different wishes have different dimensions, they must be transferred into a value with a common unit, in this case by assigning 0-10 points, where “0 points” represents “the parameter magnitude is equal to the requirement limit” and “10 points” represents “a theoretically perfect solution”. Pahl et al. (2007) originally proposed that one should award “0 points” if the concept was absolutely useless and “10 points” if the concept was an ideal solution. The upper limit was found to be problematic for several reasons. Firstly, “a theoretically perfect solution” was found to be interpreted differently by the participants in the group. Some interpreted it very literally and suggested very extreme and non-achievable parameter magnitudes, whereas others suggested those magnitudes which the best product on the market could achieve or magnitudes which they thought would be possible to achieve under favourable circumstances. Secondly, even if the participants had been able to resolve the issue of what “a theoretically perfect solution” actually meant in terms of the parameter magnitudes and had been able to establish upper limits for each wish that were perfectly equivalent, there was still the issue of the different parameters related to each wish having non-equivalent parameter magnitude ranges. For example, the weight of a turnout from the lower limit of “useless” to the upper limit of “theoretically perfect” does not have a broad range. The possibility of lifting a turnout with a crane must be balanced against the mass necessary to keep the turnout in place. Dividing this range into ten steps means that each step is quite narrow. On the other hand, other parameters might have a much wider magnitude range, and therefore a larger change in the parameter magnitude is required to change the number of assigned points. In the applied method, “0 points” is regarded as representing the requirement limit instead of “an absolutely useless solution”, and then the parameter magnitude change required to change the number of assigned points will be affected by the choice of the requirement
limit (and be smaller compared to using “an absolutely useless solution” to de-
fine the lower limit). Further, a change in the requirement limit will change the
parameter magnitude that is required to achieve a certain number of points, and
this is especially pronounced for the lower part of the value scale (assuming that
the same upper limit is used). This implies that apparently reasonable variations
in the parameter magnitudes assigned to the upper and lower value scale limits
can be of decisive importance for the outcome. It should be noted that the con-
cepts are judged in the same relative way, but, as will be illustrated in an exam-
ple, the extent of the parameter magnitude ranges can impact the final decision,
which means that, if one is not conscious of how the values are assigned, the
importance of a certain wish might be unconsciously increased or decreased in
the calculation with respect to the other wishes. A simplified example is pre-
sented in Table 8.5 and Table 8.6. For clarity, only two wishes stated during the
application of the MPB method were used in the above example, namely the
wish of buying a turnout with a low purchase price and the wish of this turnout
having a long lifespan. The maximum purchase price (the requirement) was set
to 3 MSEK. The upper limit of “10 points” was determined by the participants
during the workshop to be 0 MSEK. This was assigned despite a discussion
among the participants where it was argued that it was not realistic that the
turnout supplier would give turnouts away for free, and perhaps it would be
even more ideal to be paid when receiving the turnout. The demand limit con-
cerning the life span was set to an accumulated load of 450 MGT. The group
agreed that it would be non-realistic and unfeasible to select the ideal life span
of infinity as the upper limit. Instead they discussed a realistic ideal lifespan
between 600 and 900 MGT. In Table 8.5, both these alternatives have been
included. In Table 8.6, a comparison is made of the results obtained using an
upper limit of 600 MGT and those obtained using an upper limit of 900 MGT. It
is found that these two upper limits result in different ordinal rankings of the
three concepts considered. In the “narrow range scenario”, where the upper
limit is 600 MGT, concept 45 receives the highest score. However, if the upper
limit is changed to 900 MGT (the “broad range scenario”) concept 25 will be
the concept with the highest score. This example illustrates the challenges of
this procedure.

Another difficulty in assigning points is that the value function in many cas-
es is actually not linear (Pahl et al., 2007). For example, an initial improvement
can be worth more than an improvement close to the maximum limit.
Table 8.5 The relation of different purchase prices and life spans (narrow range/broad range) to points on the value scale

<table>
<thead>
<tr>
<th>Value scale</th>
<th>Parameter magnitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points</td>
<td>Low purchase price (MSEK)</td>
</tr>
<tr>
<td>0 (demand limit)</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td>3</td>
<td>2.1</td>
</tr>
<tr>
<td>4</td>
<td>1.8</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>1.2</td>
</tr>
<tr>
<td>7</td>
<td>0.9</td>
</tr>
<tr>
<td>8</td>
<td>0.6</td>
</tr>
<tr>
<td>9</td>
<td>0.3</td>
</tr>
<tr>
<td>10 (ideal)</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 8.6 Two different ranges of the life span magnitude give different ordinal rankings of the evaluated concepts.

<table>
<thead>
<tr>
<th>Concept No.</th>
<th>Narrow range scenario</th>
<th>Broad range scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase price (weight 0.198)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost (MSEK)</td>
<td>3 2.3 2</td>
<td>3 2.3 2</td>
</tr>
<tr>
<td>Value</td>
<td>0 2 3</td>
<td>0 2 3</td>
</tr>
<tr>
<td>Subscore*</td>
<td>0 0.40 0.60</td>
<td>0 0.40 0.60</td>
</tr>
<tr>
<td>Life span (weight 0.173)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accumulated load (MGT)</td>
<td>600 500 500</td>
<td>600 500 500</td>
</tr>
<tr>
<td>Value</td>
<td>10 3 3</td>
<td>3 1 1</td>
</tr>
<tr>
<td>Subscore*</td>
<td>1.7 0.52 0.52</td>
<td>0.52 0.17 0.17</td>
</tr>
<tr>
<td>Total score**</td>
<td>1.7 0.92 1.1</td>
<td>0.52 0.57 0.77</td>
</tr>
</tbody>
</table>

*The subscore is obtained by multiplying the weight by the value.
**The total score is obtained by adding the subscores.
8.3.1.2 Issues related to the determination of weights

The MPB method involves assigning weights to each wish by means of a matrix where the wishes are compared in pairs. According to Johannesson et al. (2013), an advantage of this method is that it avoids subjective evaluation. The procedure, which is described in Section 2.3.1.3, leads to each wish in the matrix having a maximum possible weight which is determined by the number of wishes. The maximum weight of each wish as a function of the number of wishes (up to 30 wishes) is shown in Figure 8.1.

![Figure 8.1](image)

**Figure 8.1** The number of wishes determines the maximum weight which the wishes can have when using the weight matrix.

For example if there are ten wishes, the maximum weight that a wish can achieve is 19%. In reality, a project development group may have certain wishes that are much more important that the others and this cannot be reflected by using this type of weight matrix.

During the workshop, it was found that it was difficult for the group to distinguish between requirements and wishes. For example, when comparing the wish of a low purchase price with the wish of having as few maintenance personnel on the track as possible, they did not take into account the fact that they had a purchase price limit of 3 MSEK and that they should compare the importance of having a purchase price below 3 MSEK to the importance of having a small number of maintenance personnel on the track. Demands are not subjected to weighting since they are either fulfilled or not. Another problem was that the stated wishes were not independent.

Alternative pair-wise comparison procedures can be used. For example, in the analytical hierarchy process, the importance of each criterion in relation to
the other criteria is found by approximating the maximum left eigenvector by using the geometric mean of each row and normalizing these values by dividing them by their sum (Triantaphyllou and Mann, 1995). However, this process is also connected with certain difficulties. Triantphyllou and Mann (1995) demonstrated that replacing one of the evaluated alternatives can lower the ranking of the top-ranked concept, even though the alternative which has been removed is a less-ranked alternative and is replaced with an even worse alternative. Further, using this approach in the MPB method would not solve the confusion concerning the difference between requirements and wishes.

8.3.1.3 Other considerations

Other issues emerged during the application of the MPB method. Throughout the MPB workshop, the distribution of speech was unequal. The group was intentionally composed of participants from different organizations and with different knowledge and experience. When a few people dominate the discourse, viewpoints held by the more silent people will not be taken into account. This same experience occurred during the development of the ideation method, see Section 6.3.4.

The concepts that were screened during the concept screening stage had been developed to a different extent, and some of them had much more detail. Pugh (1990) advised that concepts need to be equally developed before entering the concept screening stage, and the experience from this workshop is that this would have contributed to a fairer comparison of the concepts. (However this is not a problem inherent in the method itself.)

An additional concern is that the amount of calculations seemed to give the participants a sense of confidence in the outcome, although the calculations might be associated with several problems such as those described above in Section 8.3.1.1 and 8.3.1.2 (in addition to the uncertainty of input data, which is, however, beyond the scope of the present study). Moreover, it was difficult for the participants to grasp all the steps in the method. This might, however, be partly explained by the short time used for its application.

The issues mentioned so far are not specific to the railway context. However, there are a few issues which ensue from this specific use setting. Railway products to be used on the track infrastructure are ultimately purchased to provide a public utility. Failure to provide this utility leads to societal costs in the form of people being delayed for work or goods not reaching their destination on time, among other things. The group found it difficult to incorporate these societal costs in a meaningful way in the total concept score. Further, the considered products are financed through taxes and STA has a limited budget for buying new products or modifying existing products. Therefore, it is important to be able to motivate a purchase economically. For example, if a product is of high quality, it is typically more expensive, and then it must be shown that this increased quality implies a lower cost of ownership for a certain time span before
one considers buying it instead of a product with a lower purchase price. It is easier to propose changes in the design and maintenance strategy for assets through quantified values of cost instead of just using failure rates and the number of inspection remarks (Nissen, 2009). The professional participants agreed that, to gain a hearing for their suggestions, it is an advantage to be able to present the outcome of the concept selection method (the concept score) as a monetary sum. They found that this was easier for non-engineers to understand, as is exemplified by the following excerpt from the workshop transcript:

“But then […] if we can express it as [a sum] in [Swedish] crowns, then the rest of the organization will accept that. It’s that simple. Sums given in crowns, they represent the truth.” (D)

“Everyone understands that.” (B)

The representatives from STA found it to be an advantage that, when the concept selection method has been applied once to a product, it can be reused on later occasions with relevant modifications and can be updated with the learning gained from each project or procurement.

8.3.2 Design of a new concept selection method

8.3.2.1 Requirements for the method

When applying the MPB method, it was found that it was difficult for the participants to grasp (see Section 8.3.1.2 and 8.3.1.3), and that a small change in the procedure had a disproportionate impact on the outcome in terms of the score which each concept received (see Section 8.3.1.1 and 8.3.1.2). Otto (1995) reported that his personal experience showed that many users do not believe in the mechanics of concept selection methods and question what the differences in numbers really mean. The experience from the present project, however, is that the calculations give some of the participants a disproportionate sense of confidence in the results. In reality, the method is limited both by the uncertainty of the input data and by its procedures. Therefore, the author’s view is that non-expert users should apply methods where they are in control of the concept selection process and understand how and why differences arise. In conclusion, this requires a method where each step is easy to understand and follow. The greater the number of mathematical operations which a method includes, the more difficult it is to understand for the user(s), and therefore the more difficult it is for the users to judge if the method is appropriate for the problem at hand. Complex methods should be applied by expert users who can judge when they are appropriate and understand why a result emerges. The method developed in Stage 5 of the present research was to be used by non-experts. One of the requirements for the method to be developed was therefore that it must be easy to understand and apply. The method must avoid calculations where it is easy to
change the outcome through minor changes in the input data without the user being aware of this. An additional requirement for the method was that it must encourage all the members of the collaborative team to express their contributions and share information during the execution of the concept selection method.

In consideration of the railway-specific context, a few special requirements for the method emerged. Firstly, it must incorporate the societal cost in the evaluation to force the decision-maker(s) to take it into account. Secondly, the concept score must be expressed as a monetary sum, preferably as a cost of ownership, to make it easier to communicate the outcome to other stakeholders in the organization. Thirdly, it must provide documentation of the process so that one can return to the documentation and understand why a certain concept was chosen, and so that the process can be reused. In addition, the proposed method should be general enough to be possible to use for the procurement of existing railway products and products to be developed, as well as for development projects concerning railway products.

8.3.2.2 Life cycle and societal costs

The focus on costs and especially the cost of ownership justifies a life cycle cost (LCC) approach. The LCC approach is well known in the railway community and allows the calculation of the cost of ownership on a yearly basis (based on the expected life span of the asset), which can be compared between different alternatives. The LCC calculation can be performed at different levels of complexity. Once an LCC-model has been established, it can be evaluated, developed and adapted as the organization learns more about maintenance and other costs, including their interdependence. An LCC analysis is typically made some time before a new product or a modified product is introduced, and time may be saved by incorporating it at an early stage. The LCC can be used in the topic clarification stage to identify the cost drivers of the current solution, thus providing important input for deciding where efforts should be directed during the conceptual design of the product to be developed.

Using the LCC to estimate the yearly cost of ownership makes it a straightforward task to add on the yearly societal cost (SC) predicted to be the result of introducing a concept. In one sense, the societal cost can be viewed as part of the LCC, as it describes costs that originate from a failure of the product. In the proposed method, however, these costs are separated, so that one can distinguish between direct and indirect costs and choose when to present them separately and when to combine them into a concept score. The direct costs include all the costs that STA actually has to cover within its own budget. The indirect societal costs are not paid directly by STA, but are rather an estimate of how much society suffers economically from a failure of the asset. Depending on the purpose of applying the method, the values can be presented separately, which
gives a more nuanced understanding of the implications of the evaluated concepts.

STA has rules and regulations regarding how the LCC and SC are to be calculated and what input values should be used. These rules stipulate, for example, the hourly societal cost for different types of track closures. In projects and during procurements, there might be a wish to attach greater importance to the SC in relation to the LCC, or vice versa, for different reasons (e.g. to steer the development in a certain direction). Therefore, the SC and LCC can be assigned different relative weights, or can be considered to be of equal weight, when calculating the concept score.

Since infrastructure managers in different countries apply different sets of rules and regulations concerning societal and life cycle costs, and because different levels of detail and complexity are required on different occasions, no SC or LCC model was developed for the present research (although a case study presenting one example of an application is presented in Section 8.3.3). LCC and SC models need to be adapted to the specific cases.

The LCC and SC are combined with requirements stating the minimum values for a range of properties. Typically one does not expect STA to be prepared to pay more than what the stated required minimum values imply if the LCC and SC costs are not decreased. However, in some cases, when relevant soft parameters influence neither the LCC nor the SC in a meaningful way and STA wants to reward a better solution than a solution merely meeting the minimum requirements, the values of the properties representing these parameters are transferred into values on a monetary scale representing a monetary value that is included in the concept score. However, this procedure was not tested in the present study, since this need was identified after the study, but the procedure is included in the instructions.

It should be highlighted that presenting a concept score as a monetary value does not mean that this sum of money is the actual cost that will be paid for the concept. The score is affected by how the calculations are performed and how the SC and LCC are weighted in relation to each other and reflects how the concepts differ with respect to how the model has been set up. The model should therefore put emphasis on what the decision-maker(s) find(s) important. Just as in the case of the MPB method, the model is not capable of making a better decision than the quality of its input and its procedure allow, and it is important to be aware of this in the interpretation of the outcome.

8.3.2.3 Individual and group work

In a group composed of participants representing different actors, the topic clarification stage is especially important, since the participants have different backgrounds and typically different apprehensions of the issue. This diversity makes it important to achieve a common basis to start from, but also represents a major advantage of a cross-functional group. A cross-functional group with
participants from different relevant organizations is more likely to cover a wide range of issues related to the topic compared to a homogeneous group. However, some strategy to facilitate information sharing between the participants is required, since an expert in one domain often does not know what information is of use to an expert in another domain (Straus et al., 2011) and since it has been found that less information is shared within a team when its members possess diverse sets of knowledge (Mesmer-Magnus and DeChurch, 2009). Therefore, the topic clarification, goal specification and requirement specification stages all begin with a step where the participants individually answer questions concerning the topic and project goal, and formulate requirements according to the headings provided before discussing with the other participants. In this way, the participants cannot influence each other until they have written down their initial thoughts about the topic. After this individual step, the participants share their contributions with the group and use them as a starting point for clarifying the topic, formulating a goal specification and compiling a requirement specification, respectively. To avoid a few participants steering the discussion and to encourage all the participants to share the information which they have provided during the individual work, a structured procedure should be implemented to let each participant speak (see further in Section 5.3.3 and Section 6.3.4 regarding the speaking distribution between the participants during the tests of ideation methods). In the developed concept selection method, the same procedure is applied as in the developed ideation method; i.e. the participants take turns to present their contributions and it is not until all the participants have spoken that a free discussion commences.

To make it easier for the participants to keep track of the contributions from each participant during the topic clarification, goal specification and requirement specification stages, the contributions should be recorded in such a way that all the participants can see them, for example by writing them on a whiteboard or by projecting them on a screen. Relevant information can then be shown at each stage of the method and facilitate for the group’s discussions and decision making. This procedure also facilitates documentation.

Some modifications were introduced concerning the questions asked during the topic clarification during the MPB method and concerning the screening procedure and the screening chart applied during the MPB method. These are described in detail in the presentation of the method instructions in the next section.

8.3.2.4 Method instructions

The instructions for the new concept selection method, incorporating the changes suggested in Sections 8.3.2.2-8.3.2.3 into the MPB method, are shown in Table 8.7 and each stage is described in Sections 8.3.2.4.1-8.3.2.4.5.
Table 8.7 Overview of the instructions for the new concept selection method.

**Topic clarification**
1. Short introduction to the topic by the initiator.
2. All the participants work individually with the questions in Table 8.8 during 10-20 minutes.
3. One question is dealt with at a time. The participants take turns to present their response to the question. The contributions are documented in such a way that everyone can see them (e.g. using a computer connected to a projector, or a whiteboard), and are discussed and summarized before the group proceeds to the next question.
4. Continue until all the questions have been dealt with.
5. Establish preliminary LCC and SC models and make an LCC and SC calculation with respect to today’s existing solution (if relevant) to identify the largest cost drivers.
6. Check that the topic clarification and especially the (root) causes of the situation are well enough illuminated and understood to justify ideation or procurement. Otherwise analyse what must be done before proceeding.

**Goal specification**
1. The objective of formulating a goal specification is to make sure that all the participants have the same interpretation of what the project should achieve and to use the specification as a reminder of the goals during the project. The goal specification should be 1-3 sentences long (Pahl and Beitz, 2007).
2. All the participants individually formulate a goal specification (their idea of the outcome of the project).
3. The participants take turns to present their goal specifications.
4. The participants discuss the contributions and agree upon a joint goal specification.

**Requirement specification**
1. All the participants work individually on the formulation of a requirement specification according to the (relevant) headings in the checklist described in Section 2.3.1.3 (Table 2.2) for 10-20 minutes.
2. One heading is dealt with at a time. The participants take turns to present their requirement specifications. The requirements are documented in such a way that everyone can see them and are then discussed. Similar requirements are merged, and a check is performed to ensure that all the contributions are indeed requirements (a requirement is something that is either fulfilled or not fulfilled) and to determine whether they are all necessary. The content under each heading is summarized before proceeding to the next heading.
3. Continue until all the headings have been dealt with.
4. If relevant, identify “soft parameters” which are not valued by the LCC or SC models and for which a value above the stated minimum requirement value is sought. Transfer the values for the soft parameters to a monetary scale.
   - The requirement value should correspond to zero in the monetary scale.
   - The maximum limit of the monetary scale should be the maximum sum which the purchaser is prepared to pay for an improvement, and should correspond to the value of the property connected to the soft parameter which they would expect at this price.
   - The values between the minimum requirement value and the maximum sum that the purchaser is prepared to pay are extrapolated.
Decide weights
The weights for the evaluation criteria, the LCC and the SC, are determined through agreement between the participants. The participants reflect on how big an influence they desire the LCC and SC, respectively, to have on the concept score in the specific case. If soft parameters are included in the concept score, the weights of these in relation to the weights of the LCC and SC are also determined (as a starting point, it is reasonable to set the weight of the soft parameters as equal to that of the LCC).

Concept generation and development or concept compilation
This stage is not part of the method developed in Stage 5 of the present research. The ideation method developed in Stage 3 (see Chapter 6) can be used in the case of research projects and in the case of procurement the concepts are compiled through tenders. The output from this phase is a number of non-prioritized concepts that are developed to a similar degree of detail.

Concept screening
1. Go through the requirements list and choose the requirements to be used for the screening of concepts.
   - Prioritize the requirements that will not be based solely on guesses.
   - Prioritize the most important requirements.
   - Choose as few requirements as possible, preferably a maximum of ten requirements.
2. The group uses a screening chart (adapted from Pahl and Beitz (2007), an example of which is shown in Figure 8.5) to select the concepts that should be developed further. As soon as a concept receives a minus for any criterion, it is eliminated and not screened further. The reason for elimination is given in the screening chart.
3. The information available is often insufficient at this stage, and then the group has to make a decision as to whether it has the time to gather the needed information or whether the concept should be selected/eliminated despite the lack of information at the current stage.

Concept scoring
1. A concept must fulfil all the requirements to proceed further.
2. Calculation of the LCC according to the agreed LCC model.
3. Calculation of the SC according to the agreed SC model.
4. Calculation of the values connected to soft parameters according to the specified monetary scales (when applicable).
5. Calculation of the concept score according to Equation 8.1 (see Section 8.3.2.4.5).
6. The lowest concept score indicates which concept(s) is/are the most promising for further development.
7. An analysis should be made of the reasons why the concept with the lowest score received that score, in order to gain an understanding of what distinguishes this concept from the other concepts. One should analyse why the concept with the lowest score outperformed the other concepts and investigate whether there is some area in which the concept with the lowest score is weaker than the other concepts despite receiving the lowest score. This knowledge can be used to improve the most promising concepts in an iteration of the development process, in the case of a development project.
Table 8.8 Checklist with questions to be answered by each participant individually during the topic clarification.

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What is the core of the problem and wherein lies the greatest need?</td>
</tr>
<tr>
<td>2</td>
<td>Who wants the problem to be solved and why?</td>
</tr>
<tr>
<td>3</td>
<td>What are the (root) causes of the problem?</td>
</tr>
<tr>
<td>4</td>
<td>What hidden wishes, demands and exceptions are involved?</td>
</tr>
<tr>
<td>5</td>
<td>What are the tasks which the product has to solve?</td>
</tr>
<tr>
<td>6</td>
<td>Which possibilities are open and which possibilities are closed for the achievement of the new concept?</td>
</tr>
<tr>
<td>7</td>
<td>What properties should the concept have/not have?</td>
</tr>
<tr>
<td>8</td>
<td>What alternative products exist today?</td>
</tr>
<tr>
<td>9</td>
<td>What are the demands according to standards and regulations?</td>
</tr>
<tr>
<td>10</td>
<td>What are the wishes and demands regarding possibilities for product upgrading?</td>
</tr>
<tr>
<td>11</td>
<td>Are there other aspects that should be taken into account?</td>
</tr>
<tr>
<td>12</td>
<td>What are the future trends regarding technology, environmental aspects and ergonomics?</td>
</tr>
</tbody>
</table>

8.3.2.4.1 Topic clarification

The method starts by letting the participants individually answer the questions shown in Table 8.8. The questions constitute a refinement of those used in the MPB method and originally most of the questions were inspired by Pahl et al. (2007, p. 150). All the questions may not be relevant to all product development situations, but by checking this list, one can be confident that one has covered many important aspects. During the group review of the participants’ replies to the questions, the participants take turns to state and explain their contributions.

In railway contexts, the product to be developed can be compared to an existing solution. Therefore, it is valuable to make an analysis of the LCC and SC of the existing solution to identify the major cost drivers involved. This will provide important information for the team whose task it is to generate concepts or highlight those areas where extra effort should be made when writing the tender request. In situations involving procurement, the LCC and SC models to be used in the tender request cannot be changed later in the evaluation process and therefore require special attention. It is important to remember that the method selected for calculating the LCC and SC will guide the development of concepts/products by the suppliers. The development of LCC and SC models should not be seen as an isolated event during a specific project, and these models should be reused and continuously modified to reflect the learning gained from different projects and procurements, as well as the day-to-day operation of the asset concerned. In this way, the procurement process can be refined to be-
come a powerful tool in directing the suppliers to develop products that meet the objectives of the infrastructure manager.

8.3.2.4.2 Goal specification

The goal specification stage is important to make sure that the participants have actually established a common basis and agreed on the overall goal of the process. The goal specification should briefly, in a maximum of three sentences, specify the overall goal of the project (Pahl et al., 2007). Optimally, the participants are in agreement at this stage, but, if not, the group needs to discuss and agree upon a goal specification before continuing. Compromises might be necessary, but the purchaser (in this case the infrastructure manager) generally has the last word. Another important point about the goal specification is that it should specify a vision of what the end-product will fulfil, but it should not specify how this vision should be realized, as this will lock the development process to certain solution principles.

8.3.2.4.3 Requirement specification

After working on the topic clarification and goal specification, it is time for the group to consider the requirements regarding the product. This stage is performed in a similar way to the topic clarification stage. To their help, the participants have the list of headings described in Section 2.3.1.3 and shown in Table 2.2. The group can restrict itself to the requirements that concern the development/modification of the product.

After the requirement specification, the process branches in two directions depending on whether one is performing a research project or a procurement. In a research project, the group moves on to the ideation phase, where solutions addressing the topic are collected and/or generated. These solutions need to be organized and a concept screening is typically performed to sort out a number of ideas that should be developed further conceptually. Different methods can be used during this phase, and the result should be a number of different concepts that are developed to a similar degree. In Section 8.3.2.4.4, one such method is suggested. The extent to which the concepts are conceptually developed determines which method is most suitable. The method described in Section 8.3.2.4.4 assumes that the concepts are developed to a similar extent and that some information about them was gathered after the ideation took place.

In the case of a procurement, a workshop with representatives from relevant actors will provide information on a wide range of aspects of the product to be purchased. However, for procurements, the infrastructure manager needs to review the requirements further and add other relevant requirements before issuing a tender request. The requirement list and the LCC and SC models need to be included in the tender request. Each supplier’s individual process results in a tender offer delivered to the infrastructure manager. The tenders each repre-
sent a concept to be evaluated in comparison with other tenders according to the specified procedure.

8.3.2.4 Concept screening

Concept screening is performed using a screening chart which is based on that created by Pahl et al. (2007, p. 108) and which has been modified based on the findings made during the test of the MPB method. The elimination criteria are described in Table 8.9. The reason for the elimination of a concept should be given in the screening chart to facilitate later understanding of why a certain concept was chosen or discarded.

<table>
<thead>
<tr>
<th>No.</th>
<th>Elimination criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Addresses the core problem</td>
<td>Check if the concept actually has the potential to solve the main problem connected to the product development process.</td>
</tr>
<tr>
<td>2</td>
<td>Realisable in principle</td>
<td>Check if the concept is realizable in respect of performance, layout, etc. (Pahl et al., 2007). Cost is also considered under this criterion.</td>
</tr>
<tr>
<td>3</td>
<td>Selected requirements are fulfilled</td>
<td>Check whether the concept fulfils a subset of requirements that has been selected from the requirements list by the group because of their high importance and because they offer the possibility of judging whether the concepts fulfil them or not at the current stage.</td>
</tr>
<tr>
<td>4</td>
<td>Special reasons for preferring or discard-</td>
<td>Check whether there are special reasons for preferring or discarding a concept. Examples of possible such reasons are the possibility of using an earlier developed product, the fact that the concept is known to fit into the existing infrastructure, and the fact that the organization has no knowledge of the design principles used in the concept.</td>
</tr>
</tbody>
</table>

An important modification of the screening procedure proposed by Pahl et al. (2007) is that, in the newly developed method, the product development team does not consider if a concept fulfils all the demands of the requirements list. Instead, the team selects a subset of requirements which are considered to be of high importance and which offer the possibility of judging (at least tentatively) whether the concepts fulfil them or not at the current stage. The group found that this modification streamlined the concept screening, since many properties, and hence the fulfilment of related requirements, are difficult to judge at the current stage of development and therefore unnecessary to consider.
8.3.2.4.5 Concept scoring

The concept scoring procedure is similar for both research projects and procurements. During research projects, however, the stage can be applied iteratively, and new concepts or modifications of concepts may be generated in the evaluation phase. In the procurement case, the information delivered by the suppliers in tenders needs to be of high enough quality to act as input in the LCC and SC models, and all the requirements must be addressed.

The concept score is calculated by:

\[ TC_i = w_{LCC}LCC_i + w_{SC}SC_i - w_{SP} \sum_{j=1}^{n} SP_{i,j} \]  

Equation 8.1

where \( TC_i \), \( LCC_i \), \( SC_i \) and \( SP_{i,j} \) are the total cost, the LCC, the SC and the estimated monetary worth of the soft parameter(s) of the \( i^{th} \) concept, respectively, \( w_{LCC} \) is the weight of the LCC, \( w_{SC} \) is the weight of the SC, \( w_{SP} \) is the weight of the soft parameters, and \( n \) is the number of soft parameters. The worth of the soft parameters is subtracted from the LCC and SC, since, by definition, it is better the greater it is, whereas the total cost is better the smaller it is. For example, passenger comfort is considered to be better the greater it is, but is a parameter that is difficult to incorporate into the life cycle cost. Note that \( TC_i \) is not a real cost, but an aggregation of different parameters on a monetary scale and represents a score designed to pick out the most suitable concept based on the needs of the infrastructure manager.

If there are restrictions concerning how the LCC and SC can be calculated, the team responsible for the purchase have the opportunity to assign weights to the LCC and SC if they think that one of them has an inappropriately low or high value compared to the other, considering the needs in the case in question. The weight of the soft parameters, if there are any, must then also be considered so that the impact of the soft parameters does not become negligible compared to the impact of the LCC and SC. As a starting point, it is reasonable to use the same weight for the LCC and the soft parameters.

The concept score should also be calculated for the existing product or solution, so that the proposed concepts can be compared to this “benchmark” solution. The concept with the lowest total cost is the winner of this evaluation. In the case of a procurement, this winner is strictly regarded as the optimal solution, unless the purchaser has specified special rules in the tender request. This may be the case if, for example, two bids are so close that the uncertainty in the calculation is greater than the difference between them.
8.3.3 Application and evaluation of the new concept selection method

8.3.3.1 Topic clarification and requirement specification

The group used about 17 minutes to write down their replies to the questions in Table 8.8. Examples of their replies are shown in Table 8.2-8.4. These examples indicate that the participants contributed different viewpoints on the questions. The quantitative analysis of all the contributions shows a similar pattern. The group in total generated 92 different pieces of information during the topic clarification. As can be seen in Table 8.10, 80 pieces of information were mentioned by only one person and none were mentioned by more than four people.

The group used approximately ten minutes to write down requirements individually and compiled 58 relevant requirements. As is shown in Table 8.10, 48 requirements were mentioned by only one person. No requirement was mentioned by more than four participants. From these findings, the conclusion can be drawn that the participants had come to think of different aspects regarding the topic when answering the questions for the topic clarification and formulating the requirements under the different headings, and in this sense the method instructions worked as intended.

Table 8.10 The number of participants mentioning specific pieces of information during the topic clarification stage and requirements during the requirement specification stage, respectively.

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>Topic clarification stage</th>
<th>Requirement spec. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>92</td>
<td>58</td>
</tr>
</tbody>
</table>

It was found that most contributions made during the topic clarification and requirement specification stages were unique for a certain participant. It is possible that the participants might have written down more requirements than those which they mentioned, and in particular that they might have refrained from mentioning a requirement that had already been contributed by someone else. When giving an account of their answers to all the topic clarification questions and the requirements which they had formulated for most of the requirement headings, the participants took turns to present their contributions, and they rotated the task of starting the presentation session. It is therefore of interest to investigate if speaking in a certain order affected how many contributions a participant made.
Figure 8.2 shows the number of contributions made depending on the order of speaking.

![Figure 8.2](image)

**Figure 8.2** The total number of pieces of information (during topic clarification) and requirements (during requirement specification) presented, according to the order of speaking.

It appears that during the topic clarification the first three people to speak contributed more information than the last three. During the requirement specification, it appears that only the person who spoke last contributed less than all the others. Intuitively, one would expect that the number of contributions would decline later on in the order of speaking, and this would especially be the case if the participants had a homogeneous background regarding the topic. In theory, if all the participants had exactly the same knowledge, it would be expected that the first person would have mentioned all the relevant information (assuming a perfect memory) and none of the following group members would have anything to add. Although the number of contributions declined, new contributions were also made by the last participants in each round, and this indicates that they were able to provide relevant information that had thus far not been thought of and presented by another participant. It is believed that this was possible due to their different knowledge of and perspectives on the topic. However, not all the participants spoke the same number of times in each speaking sequence and this might have influenced this finding. Figure 8.3 shows how many times each participant (A-F) (see Table 3.1) spoke in a certain order during the topic clarification.
Figure 8.3 The number of times each participant (A-F) (see Table 3.1) spoke in each position (1-6) in the order of speaking (“1” denotes the first position in the speaking order and “6” the last position) during the topic clarification.

For example, participant F spoke three times in position 5 and five times in position 6, whereas participant D spoke only once in each of those positions. Figure 8.4 shows (for all the categories) the total number of pieces of information contributed by each participant and (for all the headings) the total number of requirements formulated by each participant.

Figure 8.4 The total number of pieces of information and demands mentioned by each participant during the topic clarification and requirement specification stage, respectively.
Participant D was able to contribute more pieces of information than participant F, and this might have been because participant F spoke late during each round more often and hence might have refrained several times from mentioning contributions that other participants had already mentioned. On the other hand, participant E, for example, spoke first in the speaking order several times more than participant F and spoke last fewer times than participant F, but still contributed fewer pieces of information. It is likely that several factors play a role in the contribution of pieces of information. These results might have been affected by both the order in which each person spoke and other factors such as the relevance of their knowledge to the topic at hand.

Each participant did not contribute answers to every question. As can be seen in Table 8.11, the number of pieces of content and contributing participants was higher for questions 1-7 than for questions 8-12. This probably reflects the fact that questions 1-7 were general and open-ended, whereas questions 8-12 asked more specifically about different aspects which all participants might not have had knowledge about and which might have had a limited number of possible answers.

<table>
<thead>
<tr>
<th>Question</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pieces of content</td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>22</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Number of contributing participants</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

In conclusion, the results suggest that the participants were able to contribute different information about the topic, and it is reasonable to assume that this was due to their different knowledge regarding the topic.

8.3.3.2 Goal specification

After reviewing and discussing the goal specifications written down by each participant, the group agreed on the following goal for the product development process:

“A turnout which maintains its track geometry and gives few [traffic] disturbances, [and] which is possible to adjust in a cost-efficient way (with reference to the LCC) when the track geometry changes.”
8.3.3.3 Concept screening

During the concept screening, discussions in the group around one of the concepts led to the inclusion of one additional concept, V₆.

The group selected a subset of requirements from the requirements list to be used during the concept screening, see Table 8.12 (see also Section 8.3.2.4.4). Figure 8.5 shows the screening chart that was used. Three concepts (V₂, V₃ and V₄) were eliminated because they were found by the group not to fulfil at least one of the selected requirements in Table 8.12.

Table 8.12 Subset of requirements chosen by the group to be used during concept screening (quotes)

<table>
<thead>
<tr>
<th>No.</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>It must be possible to adapt it to all types of turnouts with a fixed crossing.</td>
</tr>
<tr>
<td>2</td>
<td>It must fulfil the minimum limits for track geometry (EN, TDK).</td>
</tr>
<tr>
<td>3</td>
<td>The speed of deterioration (the vertical and lateral movement as a function of time) must be less than that of the turnout used today.</td>
</tr>
<tr>
<td>4</td>
<td>The base design must be able to absorb energy without suffering plastic deformation.</td>
</tr>
<tr>
<td>5</td>
<td>It must be prefabricated.</td>
</tr>
<tr>
<td>6</td>
<td>It must be delivered in sections (modularized).</td>
</tr>
<tr>
<td>7</td>
<td>It must be easy to transport to the installation location.</td>
</tr>
<tr>
<td>8</td>
<td>It must be easy to adjust the track geometry.</td>
</tr>
<tr>
<td>9</td>
<td>The maximum LCC must be the same as that today.</td>
</tr>
<tr>
<td>10</td>
<td>The prototype must be manufactured within one year.</td>
</tr>
<tr>
<td>11</td>
<td>The mean time between failures must be at least 20 years ([for the parts made of] steel).</td>
</tr>
<tr>
<td>Concept</td>
<td>Addresses the core problem (+/-/?</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>V1</td>
<td>+*</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>+*</td>
</tr>
<tr>
<td>V3</td>
<td>+*</td>
</tr>
<tr>
<td>V4</td>
<td>+*</td>
</tr>
<tr>
<td>V5</td>
<td>+*</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>V6</td>
<td>+*</td>
</tr>
</tbody>
</table>

**Figure 8.5** Screening chart used during the concept screening process.
8.3.3.4 Concept scoring

The group decided to take only the costs connected to the track geometry into account in the LCC and SC calculations. The LCC was calculated according to:

\[
LCC = \frac{C_c}{l} + \frac{C_{inst}}{l} + C_{geo} + C_m
\]

Equation 8.2

where \(C_c\) is the cost price, \(C_{inst}\) is the installation cost, \(l\) is the life span, \(C_{geo}\) is the yearly track geometry adjustment cost and \(C_m\) represents the other maintenance costs per year. \(C_{geo}\) was estimated based on the cost of performing one track geometry adjustment and the frequency of track geometry adjustments expected for the concept in question (based on its design). \(C_m\) was estimated by considering the size of the reduction in the other maintenance costs which could be expected due to the improvement in the track geometry achieved through the concept design in question, in comparison with the existing solution, \(V_0\).

The SC consisted of the costs to society caused by the failure of a turnout (leading to corrective maintenance) and connected to the closing of the track when performing planned maintenance. (Maintenance work is considered to be planned in advance if it is decided upon three months before it takes place; this is the point in time when the train operators finalize their train schedules.) The SC caused by the failure of a turnout was calculated by estimating the percentage of track-geometry-related failures that would be the case with a certain concept compared to \(V_0\), the average time length for remedying such a failure and the SC for unplanned stops per time unit (this cost is part of the rules applied to SC calculations at STA). The estimation of the percentage of track-geometry-related failures was based on the assumption that a design that implied a more stable track geometry would result in a lower number of failures. The SC originating from planned maintenance stops was based upon the frequency of track geometry adjustments expected on the basis of the design of the concept, the time for completing a track geometry adjustment and the SC for planned stops per time unit (again determined by rules applied to SC calculations at STA). All the parameters were given on a yearly basis.

It was decided that the weight of the SC should be 2 (and that the weight of the LCC should be 1) in this specific case, and this decision was taken to give the SC some extra weight compared to the LCC, since some participants thought that the rates for a standstill were too low and did not reflect the actual effects on society. Table 8.13 shows the resulting LCC and SC and the concept score.
Table 8.13 The LCC and SC, and the score for each concept.

<table>
<thead>
<tr>
<th></th>
<th>V₀</th>
<th>V₁</th>
<th>V₅</th>
<th>V₆</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCC (SEK/year)</td>
<td>170000</td>
<td>140667</td>
<td>131111</td>
<td>114545</td>
</tr>
<tr>
<td>SC (SEK/year)</td>
<td>22400</td>
<td>14000</td>
<td>1000</td>
<td>500</td>
</tr>
<tr>
<td>Score (SEK/year)</td>
<td>214800</td>
<td>168667</td>
<td>133111</td>
<td>115545</td>
</tr>
</tbody>
</table>

The input data were estimated by the group. According to these calculations, V₆ is the most promising concept. However, V₅ is not far behind and it is interesting to investigate what distinguishes V₅ and V₆ from the other concepts and from each other. Despite a higher purchase price for V₅ and V₆ compared to V₀ and V₁, the former pair of concepts still had a lower LCC, mainly due to the expectation that no tamping would be required during their lifetime. The decisive difference between the concepts V₅ and V₆ was that V₆ was estimated to have a ten year longer life span than V₅. When comparing V₅ and V₆, the effect of the SC on the outcome was small.

8.3.3.5 The views of the participants

8.3.3.5.1 Questionnaire results

Figure 8.6 shows to what extent the participants on average agreed with different statements about applying the MPB method and the developed method. Table 8.14 shows the statements in the questionnaire. Overall, the questionnaire results indicate that the participants were more satisfied using the newly developed method than using the MPB method. Especially, the participants found the new method to be considerably more time-efficient and easier to carry out. Interestingly, the participants found the newly developed method to have greater potential for improvement than the MPB method, which seems odd since the developed method was based on the MPB method. One reason could be that the developed method was easier to understand and therefore it was easier for the participants to envisage how it could be improved. Another reason could be that, when applying the MPB method, the participants did not have any other method to compare it with.

Some interesting patterns were detected in the detailed analysis of the individual participants’ levels of agreement. Regarding the statement that the participants were confident that they had selected the concepts that had the greatest potential to solve the problems associated with the topic, the participants representing STA had a considerably lower level of agreement compared to the other participants (on average 0.08 compared to 0.68 in the case of the application of the MPB method and 0.30 compared to 0.82 in the case of the developed method). The participants representing STA also had a considerably lower level of agreement with the statement that the group had made an objective choice regarding which concepts should be developed further (on average 0.56 compared
to 0.85 concerning the application of the MPB method and 0.36 compared to 0.82 concerning the application of the developed method). One possible explanation is that more time might have been needed for working through the steps of the method, and especially for creating the requirement list and the LCC and SC models, and that the STA representatives had most experience of the effect on the final outcome of deficiencies in the requirement list and the cost models.

![Figure 8.6 Questionnaire results](image)

**Table 8.14 Statements in the questionnaire**

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am engaged in the topic.</td>
</tr>
<tr>
<td>2</td>
<td>I think the method was easy to carry out.</td>
</tr>
<tr>
<td>3</td>
<td>I feel that the criteria we chose for evaluating the concepts are relevant.</td>
</tr>
<tr>
<td>4</td>
<td>I feel confident that we have selected the concepts that have the greatest potential to solve the problems associated with the topic.</td>
</tr>
<tr>
<td>5</td>
<td>Eventually, we as a group agreed on which concepts should be selected for further development.</td>
</tr>
<tr>
<td>6</td>
<td>I think that we made an objective choice regarding which concepts should be developed further.</td>
</tr>
<tr>
<td>7</td>
<td>I think the method is time-efficient.</td>
</tr>
<tr>
<td>8</td>
<td>I felt engaged during the workshop.</td>
</tr>
<tr>
<td>9</td>
<td>I am satisfied with my own contribution during the workshop.</td>
</tr>
<tr>
<td>10</td>
<td>I think the method is useful.</td>
</tr>
<tr>
<td>11</td>
<td>I think the method can be improved.</td>
</tr>
</tbody>
</table>
8.3.3.5.2 Interview results

Overall, the participants during the interviews expressed that they were satisfied with the newly developed method. The participants’ comments regarding the developed method can be summarized by highlighting three features which all the participants regarded as positive. Firstly, the method provided a structured way of thinking, as is exemplified by the following quote:

“I like this structure, it fits fairly well, so to speak.” (D)

Secondly, the topic clarification gives a good picture of the problem, covering a large number of aspects, and helps the group quickly to develop a common understanding of the issue, define the problems and their root causes, and envisage the characteristics of possible solutions, etc. The participants found that the topic clarification lays the foundation for making a good evaluation of the concepts and that, if the pre-stages are performed satisfactorily, then the concept screening and scoring will run smoothly. This is exemplified by the following quote:

“And so [the topic clarification] is firstly an individual [phase] and then a common [phase]; there’s an enormous spread [of ideas] and […] the amplitude of those [ideas], the way in which one becomes acquainted with different issues, becomes gigantic. I often focus like a nerd on small concrete M8 threads, whereas [another participant in the group] might focus higher and think about STA on the whole, and in some way we babble on […] [and think that] this is what we should all be thinking about and I think we get […] a good picture of it. We don’t lose so much, we gather everything together and then, when we write it down, that’s good.” (B)

Thirdly, the participants liked the fact that the method provides documentation of the concept selection process, since that makes it easy to return to the process several times and to reuse the process in a future development project or procurement. This is exemplified by the following quote:

“And then it is also the case that it becomes reusable. One can return to a process several times […]” (C)

The developed method process was found by the participants to be neither too banal nor too complex, as is exemplified by the following quote:

“It feels quite appropriate. I don’t think it’s too banal actually, but, of course, you need to be very familiar with [the topic] from the start.” (A)
One participant, however, pointed out that the LCC and SC models can be more sophisticated than those used in the application presented in this chapter. Some knowledge about the LCC and how to calculate societal costs is probably required in the team to be able to perform the method. This knowledge exists at STA.

8.4 Design principles

During the course of the project it was realized that the developed concept selection method would be relevant for expensive products with a long life span to be purchased for the provision of a public utility, or a product that is to be purchased in a large quantity for the provision of a public utility on the Swedish deregulated railway market. The design of such a product has an impact on the direct costs (e.g. investment and maintenance costs), which are financed by taxes, and the indirect societal costs (e.g. costs originating from a failure to provide public transport for citizens). Therefore, it was found that the concept selection method should belong to the class of methods that facilitate the selection of one or more large-investment products that are to contribute to the provision of a public utility on the Swedish deregulated railway market. The design principles, presented in Table 8.15, were formulated for this class of solutions.
<table>
<thead>
<tr>
<th>Design principle</th>
<th>Justification</th>
<th>Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Easy to understand and apply</td>
<td>Participants without expert knowledge of concept selection methods should understand why a result emerges and be able to judge if the method is suitable for the context concerned. To increase the probability of implementation. To make it easy to involve new participants.</td>
<td>Straightforward procedures. Straightforward calculations.</td>
</tr>
<tr>
<td>Collection of a wide range of information on the topic</td>
<td>To avoid missing important aspects of the product to be developed which might entail great costs at a later stage.</td>
<td>Involve participants with different functional backgrounds with respect to the product. The participants work individually answering the topic clarification questions and formulating requirements.</td>
</tr>
<tr>
<td>Strategy for distributing the verbal interaction</td>
<td>All the perspectives and knowledge in the group should be considered.</td>
<td>Participants take turns to present their suggestions and comment on other participants’ suggestions.</td>
</tr>
<tr>
<td>Focus on the important information</td>
<td>Make the method efficient.</td>
<td>Evaluate concepts with respect to a subset of requirements during concept screening. Only involve costs related to the modification of the product in the LCC and SC calculations.</td>
</tr>
<tr>
<td>Consider cost of ownership</td>
<td>Assets with a long life span should be evaluated concerning not only the purchase price and installation cost, but also the costs that occur during their life span.</td>
<td>The use of the LCC on a yearly basis to estimate the actual costs induced by introducing a concept.</td>
</tr>
<tr>
<td>Consider the societal costs that arise when the asset does not deliver its intended function</td>
<td>An asset installed to provide a public utility should be evaluated regarding its ability to provide the intended utility.</td>
<td>Take the societal cost into account besides the LCC.</td>
</tr>
</tbody>
</table>
8.5 Conclusions

In Stage 5 of the research for this thesis, it was found that the main problems associated with the MPB method with respect to the given context concerned how to connect the parameter magnitudes to values expressed in points, the procedure for the determination of weights, the difficulty of incorporating societal costs, and the unequal distribution of spoken words among the participants.

To address these problems, the method developed in Stage 5 of the present research uses a concept scoring based on the life cycle cost and the societal cost, and the distribution of spoken words was made more equal by letting the participants take turns to speak.

The diversity of the group manifested itself through the participants’ ability to contribute different information about the topic.

The group was more satisfied using the newly developed method than using the MPB method, and, in particular, found the developed method to be more time-efficient and easier to implement. The participants found that the method provided a structured way of thinking, that the topic clarification gave a good picture of the problem, and that the method provided documentation which could be reused. Further, the participants thought that the method was characterized by an appropriate complexity.

The learning derived from the development of a concept selection method in a single group was formalized into the following design principles for concept selection methods to be used when selecting one or more large-investment products that are to contribute to the provision of a public utility on the Swedish deregulated railway market.

- The method should be easy to understand and apply.
- A wide range of information should be collected on the topic.
- A strategy should be created for distributing the verbal interaction.
- Attention should be focused on the important information.
- The cost of ownership should be considered.
- The societal costs arising when the asset does not deliver its intended function should be considered.
Appendix 8.1

**Instructions**
For each statement, please indicate with a cross how well you think the statement agrees with your opinion.

---

I am engaged in the topic.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

I think the method was easy to carry out.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

I feel that the criteria we chose to use to evaluate the concepts are relevant.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

I feel confident that we have selected the concepts that have the greatest potential to solve the problems associated with the topic.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

Eventually, we as a group agreed on which concepts should be selected for further development.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

I think that we made an objective choice regarding which concepts should be developed further.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree

I think the method is time-efficient.
- [ ] Do not agree at all
- [ ] Agree
- [ ] Completely agree
<table>
<thead>
<tr>
<th>I felt engaged during the workshop.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not agree at all</td>
<td>Agree</td>
</tr>
<tr>
<td></td>
<td>Completely agree</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I am satisfied with my own contribution during the workshop.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not agree at all</td>
<td>Agree</td>
</tr>
<tr>
<td></td>
<td>Completely agree</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I think the method is useful.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not agree at all</td>
<td>Agree</td>
</tr>
<tr>
<td></td>
<td>Completely agree</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I think the method can be improved.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not agree at all</td>
<td>Agree</td>
</tr>
<tr>
<td></td>
<td>Completely agree</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I like these things about the method:</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>I do not think that the following things worked well when applying the method:</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Other comments:</th>
<th></th>
</tr>
</thead>
</table>
Appendix 8.2

Test of the modified Pahl and Beitz method.

- What did you think was good and bad about the method?
- What things can be improved?

Test of the newly developed method

- What is your point of view regarding the different stages of the method, and especially those stages where large modifications were made?
  - Topic clarification?
  - Goal specification?
  - Requirement specification?
  - Concept screening?
  - Concept scoring?
- What is your experience of using the method?
- How complex did the method feel? Too banal or too complex?
- Is there room for improvement? Can you think of other ways to decide which concepts to take forward for further development?
9. Conclusions, formalization of learning and limitations

9.1 Conclusions

Specific conclusions for each stage of the research (Stage 1-5, see Table 1.1) have been drawn in connection with the presentation of the respective stages. In this section, concluding answers are provided for the research questions formulated for this thesis in Chapter 1.

Research question 1: What are the prerequisites for collaborative product development on the Swedish deregulated railway market?

- The railway maintenance contracts give no incentives for maintenance contractors to conduct proactive maintenance or to develop innovations.
- Railway actors recognize the benefits of collaboration, but are hesitant to collaborate since they are afraid that the sharing of information might give their competitors advantages.
- The fact that the Swedish Transport Administration (STA) must follow the Act on Procurement within the Water, Energy, Transport and Postal Services Sectors and that their mission is to provide a public utility while spending public funds efficiently differentiates STA’s prerequisites for product development from the prerequisites for a typical product development process taking place in a profit-driven company. Further, in the case of procurement, the requirement specification and stated procedure for evaluating concepts in the form of tenders cannot be changed after the tender request has been issued.

Research question 2: How can an ideation method be designed to facilitate information sharing and idea flow in a cross-functional inter-organizational group in the context of the Swedish deregulated railway market?

The learning acquired from the development of an ideation method to be used in cross-functional inter-organizational groups in the context of the Swedish deregulated railway market was formalized in the following design principles for methods to be used in this context:

- developing the method based on methods proven to be easy to understand,
- the initial generation of a pool of ideas with great variety,
• exposure to stimulus ideas generated internally within the group,
• mixing idea-exchanging strategies,
• the use of external representation in the form of sketches and text,
• including time for discussion and debate,
• optimizing the cycle time,
• developing a strategy for distributing the verbal interaction,
• paying attention to stimulus ideas.

The ideation method developed in the research carried out for this thesis implements these design principles through the combination of rotational viewing of generated ideas in the group (as in Method 635) and gallery viewing (as in the gallery method) with steps of verbal interaction where the group members take turns to present and comment on generated ideas.

Research question 3: How can a concept selection method be designed to support STA in the selection of an appropriate concept for further development?

The learning gained from the development of the concept selection method has been formalized in the following design principles for methods that are to facilitate the selection of one or more large-investment products that are to contribute to the provision of a public utility on the Swedish deregulated railway market, or the selection of one or more concepts for a large-investment product to be developed for the same purpose. The method should:
• be easy to understand and apply,
• be based on the collection of a wide range of information on the topic,
• include a strategy for distributing the verbal interaction,
• focus on the important information,
• consider the cost of ownership,
• consider the societal costs that arise when the asset does not deliver its intended function.

The concept selection method developed in the present research combines refined versions of the procedures for topic clarification, goal specification, requirement specification and concept screening developed by Pahl and Beitz (Pahl et al., 2007) with a concept scoring procedure based on the life cycle cost and the societal cost associated with each concept that fulfils the stated requirements.

Research question 4: What effects are produced when using conceptual design methods in cross-functional inter-organizational groups in the context of the Swedish deregulated railway market, as a result of the composition of the group?
• The most important effect has been the wide range of viewpoints on the topic at hand that the participants with different functional backgrounds and from different organizations have been able to provide.

234
During verbal interaction, four different ways were found in which the diversity of the group manifested itself, and which could lead to improved ideas and concepts, as well as additional ideas and concepts: a) objections to other participants’ ideas, b) associations inspired by other participants’ ideas, c) former experience of similar solutions, and d) immediate feedback to questions.

A negative effect that might arise in cross-functional inter-organizational groups is that the participants might not share information if competitors are present in the group.

9.2 Formalization of learning

In the research performed for this thesis, the aim has been to develop tailor-made conceptual design methods to be used by actors in the specific context of the Swedish deregulated railway market. To this end, action design research (ADR) (Sein et al., 2011) was applied within the field of engineering design. Through the guidance of ADR, the findings have been generalized to design principles that can be used as a starting point for developing conceptual design methods in a similar situation that other researchers might face within other organizations.

As was described in detail in Section 3.3.4, the last stage of ADR requires the researcher to make a conceptual move from the specific and unique to the generic and abstract on three levels to accomplish generalization (Sein et al., 2011). In the present research, instance-specific solutions have been developed which have been reconceptualized into a class of solutions. The developed ideation method has been viewed as belonging to the class of ideation methods for cross-functional inter-organizational groups in the context of the Swedish deregulated railway market, and design principles for this class were developed in parallel with the ideation method and are to be found in Table 6.4. The concept selection method was found to belong to the class of methods that facilitate the selection of one or more large-investment products that are to contribute to the provision of a public utility in the Swedish railway sector. The design principles, presented in Table 8.15, were formulated for this class of solutions.

Although the methods were developed in the context of the Swedish deregulated railway market, there is reason to believe that the stated design principles are relevant for other fields as well. The methods have not been tested in other fields, but there is an analytical reason why one should expect the design principles for the ideation method to be valid for cross-functional inter-organizational groups acting in other fields. This reason is that the line of reasoning which was employed with respect to groups with participants from different organizations and with different functional backgrounds and which lay the foundation for the design principles is not specific for the railway sector. For similar reasons, it is reasonable to expect that the design principles formulated in the present research for the concept selection method will be valid for corresponding methods.
for the selection of large-investment products that are to contribute to the provision of a public utility in other fields as well.

9.3 Limitations

9.3.1 Limitations related to Chapter 4

Study 1 is based on a small number of interviews and, therefore, even though the respondents work in different sectors, an obvious limitation of the study is the generalizability of the results. The approach selected is justified by the detailed information which could be accessed, facilitating the design of further studies of the subject. Primary data gathered by interviewing experienced people regarding the situation in the maintenance sector narrows the gap in our present knowledge of the effects originating from the deregulation process. Descriptive empirical research offers an insight into the circumstances that lie behind statistics and numbers. Another limitation of the study is the fact that the different Swedish maintenance contracts have had different content, and different contract specifications might lead to different consequences. Future research can show if the conclusions from Study 1 represent the general situation within the field. The validity of the results is, however, increased by the fact that the respondents have corresponding views, although they work in different segments of the railway sector, and therefore have different perspectives on the situation. Further, several of the findings were supported by the findings in Study 2.

Study 2 can be regarded as a reconstruction after the event had taken place. The main research aim of the workshop which the study was part of was to perform a field test of the developed ideation method. The part of the workshop that constituted Study 2 was added upon a request by the UPPSAMT project group. Afterwards, the author realized that the data that had emerged from this part of the workshop were of interest for Stage 1 of the research (see Table 1.1). The available data were analysed as described in Section 4.3.1.3. However, if the author had realized before the workshop how the respondents could contribute to Stage 1 of the research for this thesis, it would have been possible to design this part of the workshop differently, for example to obtain data on the respondents’ opinions on all the findings from Study 1. On the other hand, it was interesting that the data from Study 2 regarding the overall situation on the railway market were provided spontaneously by the respondents as a response to the posed questions. This can be interpreted as a sign that these issues are indeed quite important to the respondents.

Study 3 is limited by the respondents’ experience of how product development and procurement are performed at STA. However, it was largely possible to triangulate the data from the respondents with data from documents that are valid for STA in general.
9.3.2 Limitations related to Chapter 5-8

9.3.2.1 Small number of participants in the OptiKrea Group and the field tests

The major part of the research conducted for the present thesis took place in a single group with six members, the OptiKrea Group. This approach has revealed interesting findings and insights concerning the use of conceptual design methods in real-life settings, and these findings and insights can be of value to academics and practitioners. However, the findings and conclusions presented in Chapter 5, 6 and 8 are based on the results for this particular group. The field tests of the final version of the developed ideation method described in Chapter 7 suggest that some of the findings from the OptiKrea Group can be generalized to other cross-functional inter-organizational groups acting on the railway market. However, the number of groups participating in the tests in Stage 4 was limited, and it was not possible to control all the parameters that might have influenced the results. Involving other participants or groups in the OptiKrea project and during the field tests might have resulted in different findings due to different group characteristics exerting different influences. Possible influencing factors in this connection are the personalities of the group members and their experience, backgrounds, gender, age and other personal characteristics. Further, the groups acted in the railway sector, and studying cross-functional inter-organizational groups from another sector could result in other findings.

9.3.2.2 Assessment of ideas and concepts

Concerning the tests of the ideation methods, the number of non-redundant ideas and the number of non-redundant concepts/suggestions were carefully counted separately according to a specified procedure during the different tests in the OptiKrea Group and the field test groups, to make the count as uniform as possible and facilitate comparison between the methods and between the different groups. However, even if great care is taken to count ideas in a neutral way, subjectivity will to some extent influence the counting. In the case of the tests taking place in the OptiKrea Group (Chapter 5 and 6) the different nature of the concepts made the identification of single ideas challenging. During the field tests, only suggestions were counted. This might give an unfair picture of a group’s contribution and performed work, since a suggestion might contain information corresponding to anything from a single abstract sentence to an illustrated suggestion with concrete details covering an entire sheet of paper. Further, one cannot be sure that the person analysing the content interprets the content in the same way as the group did or in the same way as was meant by its creator, and this becomes more challenging when analysing more complex and domain-specific content. Studying the recordings of the participants’ explanations of different concepts and suggestions helped in this regard. Multiple cod-
ers might have improved the results, but the coders would have needed to have knowledge about the topics. The coding system might also have influenced the findings; for example, categorizing ideas, concepts and suggestions in a different way might have revealed other findings.

9.3.2.3 Change of topic

Because only a single group was involved in the tests of the established ideation methods and the development of the OptiKrea ideation method (Chapter 5 and 6), the ideation topic had to be changed between the ideation sessions. This was because using the same topic for each method might have resulted in the solutions from the previous ideation session(s) being reused in the subsequent session(s) and comparison of the ideation outcome between the methods would not have been meaningful. Although the different topics were chosen with care so that they would be equivalent and although they were found to lead to similar engagement among the participants (see Figure 5.1 and 6.1), it cannot be excluded that their impact influenced the ideation outcomes differently. Two groups during the In2Rail field test and four groups during the UPPSAMT field test ideated on the same ideation topic. This opens up the possibility of interesting comparisons, but the sample is too small to make any statistical generalizations. Further, none of the topics during the field tests were the same as any topic ideated on in the OptiKrea Group. The advantage of changing ideation topics is, on the other hand, that it has been demonstrated that the OptiKrea method can be applied to a range of ideation topics with satisfying results.

In contrast, the topic was not changed between the two tests of the concept selection methods. However, the concepts to which the methods were applied were changed between the tests. Further, the modified Pahl and Beitz (MPB) method and the newly developed concept selection method were not compared based upon the quantitative outcomes. The use of the same topic might, however, have biased the participants’ opinions. The test of the developed method might have gone smoother because they had worked with the topic earlier. However, the tests took place with two months in between and it is not very likely that the participants remembered their contributions in detail. Further, the group had also discussed the topic before the test of the MPB method.

9.3.2.4 Length of sessions

The length of the ideation sessions differed. Therefore, the number of concepts and ideas were given as rates (and in the case of Figure 5.3, the histograms were normalized to 60 minutes) when comparing the outcome from different ideation sessions. However, this could be misleading, since it has been shown that the number of ideas can decline over time during brainstorming (Paulus and Dzindolet, 1993), and it is conceivable that this could also be the case for other ideation methods. This would imply an overestimation of the concept and idea gen-
eration rates for the applications where the ideation times were shorter compared to the rates estimated for the applications with longer ideation times. For example, in Chapter 5, this would imply an overestimation of the concept and idea generation rate for Method 635. However, taking the number of concepts generated by Method 635 in 30 minutes and dividing that number by 60 minutes (the length of the SIL session), one obtains a concept generation rate of 0.77, which is still higher than that for both the SIL (0.38) and the gallery method (0.48 over a time of 60 min). Similarly, the idea generation rate for Method 635 was also higher when calculated for a 45 minute period.

A limitation with respect to the test of the MPB method and the developed concept selection method was that the MPB method had much less time available for the test, and this might have affected how easy it was for the participants to understand the method, among other things.

9.3.2.5 Group sizes

The OptiKrea Group and the groups participating in the field tests had between five and nine participants and therefore the concept/suggestion generation rates were given per person when comparing the outcome between the different groups. However, previous research on brainstorming has found that larger groups are less productive (Mullen et al., 1991), and a similar effect might exist during non-verbal ideation, and that would imply an unfair comparison between the groups.

9.3.2.6 Point of time for tests

The tests of the methods in the OptiKrea Group (Chapter 5, 6 and 8) were performed on different occasions. This was the only possible way to organize the tests, because a) the evaluation of an earlier test had to be completed before the next test could take place and b) the participants were busy professionals located in different parts of Sweden. Therefore, the conditions of each test may have varied with respect to the mood and fitness of the participants, and this was especially evident during the test of OKMv1, where two participants explicitly mentioned that they had had a “bad day” with respect to creativity.

9.3.2.7 Choice of ideation methods

The choice of established ideation methods to be tested during Stage 2 (Chapter 5) was informal and could have been made in a more systematic way. It would have been beneficial to let the participants reflect upon the method descriptions and then interview them individually about their opinions or to arrange a vote on the preferred method. The performed procedure was influenced more than necessary by the researchers, who were more familiar with the different methods. On the other hand, taking advantage of the different competences in the
project group is one of the principles of ADR. The choice of methods represented different types of intuitive methods that appealed to the group members and there were good reasons for evaluating these methods. However, a more systematic procedure would have increased the validity of the choice with respect to the participants’ opinions, especially about what would be beneficial for their own organization.

9.3.2.8 Using participants as raters

In the test of the established ideation methods (Chapter 5), the participants were asked to judge the viability and ability of the concepts, in preference to the employment of external raters. This introduces the risk of a bias, as the participants might rate their own concepts highly to gain personal benefits. Since the gallery method is the method in which it is easiest to find out the originator of each idea, all the concepts generated using this method were checked to see if there appeared to be any bias in the ratings. Bias was considered to have occurred if the originator had rated his idea considerably higher than all the other raters. Concerning the ability of concepts, no such cases were found. With regard to the viability of concepts, there were four concepts whose originator had given them a rating which was equal to or more than 0.3 units higher than the ratings awarded by the other raters. However, when comparing the ratings for the ability of the same concept, a similar pattern was not detected. It is believed that it is more likely that the high ratings given for viability by the originators in these cases were awarded because the originators actually considered the concepts to be good. However, if any bias exists which the author is unaware of, the use of six raters implies that, even if one of them rates his own idea very highly, the idea will still receive a low average score if the other five raters think that it is a bad idea, and in this way the effect of any existing bias will be low.

9.3.2.9 Motivation for sharing information

The developed ideation method, as well as certain steps of the developed concept selection method, was designed on the basis of implications of associative memory models (AMMs, see Section 2.2.3). Although AMMs explain how information in the long-term memory is accessed, other factors exist that can further stimulate or interfere with activation, such as motivation. Stimulation requires that the subject concerned should attend to the stimulus idea (e.g. Dugosh et al., 2000), which is a process that requires the person to be motivated to do so. The motivation for a participant to create ideas or share the ideas created might be low if there is some reason why that participant might encounter negative consequences through doing so. One factor which could lower the motivation of a participant is a fear of other participants stealing good ideas from them due to the inclusion of competing actors in the group. Almefelt and Claesson (2015) applied systematic design methods in a collaborative project including
an automotive manufacturer and 35 automotive suppliers and found that, although the application of the methodology was in general successful, in some teams where the parties represented competing suppliers, the work progress was hindered by conflicts of interest. No competitors (actors competing in the same functional segment) were involved in the creative team of the OptiKrea Group, but there is still a risk that some participants might have kept some ideas for themselves owing to a fear of having them spread outside the group. On the other hand, it might have been feasible to extend the OptiKrea Group to include other different types of actors, for example train operators, as they might have been able to contribute additional viewpoints on the method. A few representatives from train operators and other types of actors were present during the field tests.

9.3.2.10 Questionnaire

Although the participants in the OptiKrea Group did not need to fill in their name in the questionnaires, it was not possible to keep their replies in the questionnaires anonymous to the author during the four-year long project, since paper versions of the questionnaire were used and it was easy to recognize the handwriting of each participant after some time. This was also realized by the participants. Not being anonymous might have led to participants refraining from expressing certain views. It is not believed that this was a great problem during the project, since the participants did not seem to care whether they were anonymous or not to the author and the data from the questionnaires were in accordance with other data that were collected.

The rating scale used in the questionnaires might have been interpreted differently by different individuals, and different individuals might have been more “positive” or “critical” when filling in the questionnaire, depending on their personality (and possibly their mood). For example, in Figure 8.6, which shows the questionnaire results from the OptiKrea Group after trying the MPB method and the developed concept selection method, respectively, one can observe that the engagement in the topic was (on average) different (although high in both cases), despite the fact that the topic was the same during the workshops. Some variation in questionnaire results that is not dependent on the methods themselves is therefore to be expected, and it is reasonable to believe that this variation is greater between groups than within the same group between different occasions. This raises the question as to when a difference in the average level of agreement is great enough for one to consider it important, especially if the sample is too small to motivate statistical approaches. For the purpose of the present thesis, the interpretation of the questionnaire results was adapted to the specific instance, taking the circumstances and other results into account. In general, differences between average levels of agreement of less than 10 absolute % were not considered.
The limitations of the linear regressions performed on the questionnaire data were discussed in Section 6.3.2.3, 7.3.2.2 and 7.4. These limitations concerned, for example, the distribution of the data points. The value of $R^2$ provides a measure of how well the linear model fits the observed data. However, other statistical methods are necessary to determine whether the model is statistically significant. The results of the linear regression should therefore be taken as input to further analysis and as a source of directions for further research.

9.3.2.11 Lack of a control group

The research conducted for this thesis has focused on developing and testing methods suitable for application in a functionally heterogeneous group. However, the methods have not been tested in a homogeneous group for comparison. One of the starting points for the development of an ideation method was that participants with different functional backgrounds would be able to provide different types of ideas as a reaction to the ideation topic in the individual ideation phase at the start of the method, and that these ideas would be used as stimulus ideas for the other participants throughout the application of the method. If the participants of a group have a similar background, competence and perspective with respect to the ideation topic, it is expected that the ideas that the group can produce at the start of the method will have less variety and will therefore not serve the purpose of functioning as stimulus ideas for the group. Consequently, the outcome is expected to be worse compared to that for a group that is heterogeneous with respect to the ideation topic. In the case of a very homogeneous group, external stimulus ideas might be a better approach. However, it has not been within the scope of the present study to investigate this, but it is a very interesting area for future research.

9.3.2.12 Other possible solutions

The proposed ideation method is a step towards taking the knowledge of different railway actors into account during the development of railway products. One major advantage of using the proposed approach is that it provides a structured way of collaborating during product development. Other methods can also be used, if they facilitate collaboration through a clear structure that makes meetings efficient and makes it possible for all the participants to contribute their knowledge. From the research carried out for the present thesis, the conclusion cannot be drawn that the developed method is the “best” method, and it can only be concluded that it is one possible solution and is considered to be the best solution of those methods which have been tried. There are probably a large number of methods which would have resulted in a better outcome than not using any method at all.
9.3.2.13 Participation through video conferencing

During the test of the established ideation methods in Stage 2 of the research (Chapter 5), one person participated through video conferencing during all three tests. During the test of the developed method in Stage 3 (Chapter 6), all the participants were physically present during the tests. Although the person participating in Stage 2 through video conferencing was found to contribute at the same level as the other participants during the workshops, it cannot be excluded that this different form of participation might have had some impact on the comparison of results between the established methods and the developed method. For example, the participants might have behaved differently when they were all physically present at the same location, compared to when they had to interact with one participant via video conferencing. Further, the participants were given a few extra short breaks during the process of scanning and electronic transmission of sheets of paper that was necessary when one person participated via video conferencing.

9.3.2.14 Choice of field test groups

The methods resulting from the present research were developed for the purpose of product development on the Swedish railway market. However, one of the field tests concerning the ideation method took place in a European context and the other field test considered a topic that was of a different type. Considering the purpose of the method development, it might have been more beneficial to perform the field test in Swedish contexts on topics strictly concerning product development. However, as a consequence of the practical and real-life nature of the project, these were the possibilities that were given. The performed field tests, on the other hand, gave the opportunity to try the method in a somewhat different context than the OptiKrea Group, which gave insights into its area of use. An important property of all the groups in the field tests was that they were inter-organizational and cross-functional, which was the basis for the development of the ideation method.

9.4 Studies in real-life contexts

Several researchers have found that studying team performance in work settings is difficult because it is hard to obtain access to such settings (Shah, 1998; Paulus et al., 2015). However, studies of ideation methods in organizations and industry are important for understanding how the use of ideation methods in such contexts can increase the effectiveness of teams working with innovations (Shah et al., 2000; Paulus et al., 2015). In the author’s experience, the company or organization must be convinced that they will profit from this type of study in order to provide the researchers with access. This implies that the researcher has to adapt to the situation and the on-going work at the company or organiza-
tion, and that the main aim of the activity for which the researcher is given access is that it should be fruitful for the organization. Further, setting up a completely controlled experiment necessarily means procedures that would never take place in real life, and therefore the study necessarily becomes less realistic, although it still takes place in a work setting. Therefore, it was found that it is critical to find a balance between letting the participants work in a realistic setting and exerting enough control of the experiment to be able to accomplish a meaningful evaluation. Performing studies of ideation methods in work settings will necessarily imply more uncertainties and a smaller sample than performing such studies in controlled experiments with students, and therefore statistical significance is a challenge. It is the author’s view that such studies are still worthwhile, since they can shed light on findings from statistical studies and indicate if these findings are relevant in real work settings, as well as identify relevant topics for future research. The empirical findings presented in the present thesis add to the cumulative body of knowledge on concept generation and selection methods applied in different types of groups, and future research can clarify to what extent the findings can be generalized and can determine whether similar findings can be made in other real-life settings.
10. Future research

10.1 Conceptual design methods in cross-functional inter-organizational groups

The findings of the present research have indicated several areas that would benefit from future studies. Firstly, how the group-analysis of topics, the presence of verbal interaction steps, and the distribution of spoken words between participants influence the ideation outcome is important to understand in order to optimize ideation methods. Secondly, by understanding how the quality of concepts changes as they receive additions or are combined, one might avoid missing valuable contributions “hidden” in a low-quality concept and facilitate the design of concept selection methods. Thirdly, studying how the distribution of old and new ideas differs between different methods can deepen one’s understanding of how to increase the probability of truly new ideas being created during ideation.

Further, it would be beneficial to test the developed ideation method in a controlled statistical study involving both cross-functional inter-organizational teams and teams with participants with similar knowledge about the concerned ideation topic. In this way, one could test the validity of the implications of associative memory models with respect to cross-functional inter-organizational groups; these implications were assumed to be valid in the development of the conceptual design methods. A similar study could be performed regarding the concept selection method.

10.2 The developed methods and the Swedish deregulated railway market

The ideation method developed in the research carried out for the present thesis has been tested in different contexts through field tests. The next logical step is to implement it at STA, which is discussed further in Chapter 12, and then to study the implementation.

The developed concept selection method was only tested within the OptiKrea Group once. Further research would include more applications to different products and in different groups. There is a very important future research task to perform with respect to the concept selection method, and that is to develop the method further so that it will take uncertainty and risks into ac-
count. It is of decisive importance to incorporate uncertainty in the process to be able to make decisions based on an understanding of the risk of choosing a certain concept. During the OptiKrea project, different types of risks that need to be considered have been identified. Firstly, the concepts are connected to a technical risk in that the solutions that have been proposed might not work in reality. Secondly, there is a risk connected to the management of product development that the product might not meet the requirements after development, and even if it meets the requirements, it might not be finished on time or at the stated cost. Further, the data that are fed into the LCC and SC models are marred by uncertainty and there is a need to take this uncertainty into account in the concept selection. For example, it is not reasonable to choose one concept in preference to another if the difference in concept scores between them is smaller than the uncertainty in the calculation. An understanding of the influence of uncertainty in the model will assist the group in determining which area is the most important area to collect information about to reach a decision. It is of great interest to understand what incentives the method will create during procurement and how the LCC and SC models can be used to identify cost drivers and drive the product development in the desired direction.

Connected work concerns how tender requests and contracts should be designed to generate trustworthy tenders and to prevent suppliers from promising more than they are able to deliver. It was found in Stage 1 of the present research (see Chapter 4) that a collaboration model that gives all the actors an incentive to work towards a common goal is lacking. Developing and implementing such models on the Swedish railway market appear to be a prerequisite for utilizing the full potential of the conceptual design methods developed in the present research and constitute an important area for future research.
11. Reflections on using action design research to develop conceptual design methods

The descriptive and practical nature of the present thesis may enable other researchers facing a similar situation in other organizational contexts to understand if action design research (ADR) can be feasible in their case. ADR has proved to be a feasible framework for the present research. Through the guidance of ADR, it has been possible to structure the research in a meaningful way. In particular, ADR has contributed an understanding of how the learning derived from the development of conceptual design methods can be formalized. However, although it was worthwhile to use the method, there are some issues that should be highlighted.

In the case of the OptiKrea project, the overarching problem was the lack of interaction during the conceptual design phase on a deregulated railway market. The OptiKrea project can be viewed as a case of the overarching problem, as it was an intervention addressing this problem limited to a group of participants representing different actors involved in turnouts. The participants in the team agreed that working in this way in the future would be very beneficial, and according to them the methods should be beneficial for other railway products as well, since the problem in the case of turnouts is very similar to that encountered with other railway products. In the beta phase of the ideation method development (see Section 3.3.2), the organizational context was reframed as a European organizational context in the first field test, which was carried out in the In2Rail project, having previously been framed as a Swedish context. In retrospect, it would have been beneficial to have defined the wider organizational context clearly from the start. The goal of the OptiKrea project was to deliver methods that would be ready for implementation at the infrastructure manager. Implementing the methods at the Swedish Transport Administration (STA) is the most appropriate wider organizational context in this case, and one should arrange for a follow-up of the outcome of the implementation. However, it seems very likely that there is now a higher probability of the implementation being successful than if the method had been developed without interaction with practitioners in different use settings. The concept selection method has so far only been tested in the OptiKrea Group, and it is very likely that a further need for improvements will be discovered upon tests in other contexts or after an implementation. Based on the learning derived from the present research, it is recommended that one should define the situation in which interventions are to be made and the wider organizational context at the start of the ADR work.
Applying ADR was successful in that the participants thought that the intervention led to a much better outcome than that which they would normally reach in other kinds of meetings dealing with a similar problem. It was encouraging that the method worked well, but from these interventions the conclusion cannot be drawn that the developed method is the “best” method of all the possible solutions, and one can merely conclude that it is one possible solution. This solution was found to be the best one, for the given context, of those methods that were tried within the scope of the project, but there are other possible solutions that were not tried and might have worked as well, or better. There are probably a large number of methods which would have resulted in a better outcome than not using any method at all. It is recommended that one should tentatively define how to measure the success or failure of the outcome. For instance, in the case of the present project, is the project to be judged as successful because the concept generation method actually generated a higher quantity of ideas than previous strategies, or merely because the participants considered the method to be a better and more useful way of working? Moreover, how is this to be measured or assessed? Rather than employing one or more control groups, ADR relies on comparing the social situation before an ADR intervention with that after the ADR intervention. It might be very difficult to measure or assess a former situation (i.e. the situation that existed before an intervention) at a later stage. It was found that testing the developed method among participants not familiar with it in a typical use setting is very fruitful in judging its feasibility and acquiring additional insights, since the group participating in the development of the method might be at least partly biased.

The organizational commitment worked very well. It is recommended that one should take the necessary steps to lay the foundation for a good collaboration between the participants and to align the participants’ goals. It was found that the interacting nature of ADR, in which the participants’ views are continuously taken into account, plays an important part in building this commitment. Nevertheless, the scientific researchers should challenge the views of the participants whenever necessary and explain why they are doing so.

In the case of the concept generation method, a more comprehensive preliminary study was conducted by testing several established methods and reviewing relevant literature before building the first version of the method. In the case of the concept selection method, the starting point was the modified Pahl and Beitz method (see Section 2.3.1.3) and, after the first cycle, the findings derived from reviewing relevant literature were incorporated. After the first cycle of the development of the concept generation method in the OptiKrea Group, only minor changes were introduced in the method. On the other hand, large parts of the concept selection method were eliminated and changed after the first cycle of its development in the OptiKrea Group. This implies that a thorough preliminary study will be rewarded by a reduction in the work needed between the BIE cycles. However, the changes made in the concept selection method were not radical. The author is concerned that ADR might lead to incremental improvements...
rather than a radical change in the artefact once the first version has been built, although it is impossible to judge this only on the basis of the present case. If this is so, it is even more important to conduct a thorough preliminary study and consider or try different approaches, because once the first version of the artefact is built, the direction of development is determined.

As the researcher who conducted the research reported in this thesis and evaluated the scientific part, it is the author’s belief that it is more important to try to acquire knowledge of the organizational context and its interplay with the artefact on the basis of unexpected events and sudden changes in plan, rather than to try to adhere to the original plan. This exemplifies how authentic and concurrent evaluation, as opposed to evaluation using controlled settings, plays an important role in the application of ADR. However, this makes it more difficult to compare methods to each other, and statistical methods are probably more capable of determining what methods are inherently better. On the other hand, ADR appears to be a better approach to adapting methods to certain contexts while simultaneously understanding how those methods perform in reality, and to identifying context-dependent factors that are not part of controlled experiments. Therefore, it was found that ADR and statistical research methods complement each other.

11.1 Conclusions

From the experience of using ADR in a project developing concept generation and selection methods in close collaboration with practitioners, the conclusion is drawn that ADR is a feasible complement to other research methods in the field of engineering design. In the present case, the main advantages of ADR were found to be the structure provided, the understanding gained of how the learning from the study can be formalized, and the authentic and concurrent evaluation. To apply the method with success, it is recommended that, at the start of the project, one should define the situation where interventions are to be made and the wider organizational context, specify how to measure the success or failure of the outcome, and take the necessary steps to secure long-term commitment from the participants. Finally, researchers aiming to test ADR should be aware that the solution developed might not be the “best” solution, but might instead be one possible solution optimized for the given context.
The research work presented in the present thesis has been conducted within an applied project with the goal of developing collaborative conceptual design methods which are tailor-made for the Swedish deregulated railway market and which will be implemented after the project ends. In this chapter, the current processes used at the Swedish Transport Administration (STA) for product development are summarized, and a strategy is presented for fitting the developed methods into the existing framework. Possible measures for adapting the developed methods further to the existing framework are suggested and a possible road map for implementation of the methods is presented. The content of this chapter is based on informal discussions that have taken place in the OptiKrea Group, the existing routines at STA and deliberations of the author on the subject. It should not be seen as a research contribution but as suggestions based on the learning acquired during the research project described in the present thesis.

12.1 Framework for product development at STA

STA follows a routine called TDK 2014:0307, “Införande av ny eller modifierad komponent i anläggningen” [Introduction of a new or modified component in the infrastructure] (Trafikverket, 2014b), when they intend to introduce a new or modified product. TDK 2014:0307 is based upon the European standard EN 50126, “Railway applications: the specification and demonstration of reliability, availability, maintainability and safety (RAMS)” (CENELEC, 1999), which is also the Swedish standard. TDK 2014:0307 (Trafikverket, 2014b) must be applied according to TDK 2014:0162, “Driftsäkerhet, säkerhet och underhåll av järnväg” [Reliability, safety and maintenance of the railway] (Trafikverket, 2014a). By following EN 50126 (CENELEC, 1999), it is assured that the actions coupled to the introduction of a new or modified component are planned taking RAMS into account. Figure 12.1 shows the process to be followed according to TDK 2014:0307 (Trafikverket, 2014b). This process is a typical stage-gate process (cf. Cooper, 1990) where different activities (stages) are followed by decision points (gates) where it is determined, based upon the information provided from the preceding stage, whether the process should proceed forward or be terminated (or whether some other relevant action should be taken). EN 50126 (CENELEC, 1999) describes the process to be followed to
Figure 12.1 The process followed according to TDOK 2014:0307 (Trafikverket, 2014b) with an indication of the corresponding phases in EN 50126 (CENELEC, 1999). Adapted from TDOK 2014:0307 (Trafikverket, 2014b) with an indication of the corresponding phases in EN 50126 (CENELEC, 1999). G = gate, i.e. a point where a decision is taken as to whether or not one should continue the process.
assure that all the aspects of an asset from “the cradle to the grave” are considered. However, different methods can be used to provide the information required by EN 50126 (CENELEC, 1999).

12.2 Combining STA routines and the OptiKrea process

During the OptiKrea project, two different situations were identified where STA deals with the conceptual design phase. The first situation concerns procurement and involves STA procuring a product based on a requirement specification that is used in a tender request and the actual development of new concepts takes place at the winning supplier. The other situation concerns research projects, where STA can take part in the concept generation. Based on these two different situations, a process based on the methods developed during the OptiKrea project was created. Figure 12.2 shows the overall process developed within the OptiKrea project, starting with the detection of a need to introduce a new product or modify an existing product, and ending with the selection of one concept or a handful of concepts with the greatest potential for addressing that need. The method stages dealt with in the present thesis have a blue background and are marked with a bold frame. For the case of research projects, the left-hand path of Figure 12.2 can be followed. For procurement, the right-hand path of Figure 12.2 can be merged into the process demanded by TDOK 2014:0307 (Trafikverket, 2014b). More specifically, it is the phase entitled “Make plans for a new or modified component in the infrastructure” in Figure 12.1 that corresponds to the right-hand path in Figure 12.2. The methods for task definition, i.e. topic clarification, goal specification and requirement specification, can be used in this phase to identify the information required in the phases of pre-study, overall system description and requirement specification. The concept selection method can be used to evaluate the tenders received from suppliers and select the best tender, and must in this case be stated in the tender request.

The methods can be useful tools for STA within their organization (e.g. during meetings involving participants from different divisions). However, the assumption of the OptiKrea project was that better products would be developed if different market actors were to collaborate and the goal of the project was to facilitate collaboration between different actors on the railway market. The OptiKrea Group has suggested that relevant actors should be involved already during the planning phase (see Figure 12.1) so that STA in a structured way can collect relevant information. One way of accomplishing this would be to invite relevant actors to a workshop where a common task definition would be performed from which STA could take information that could act as input to the pre-study, the overall system description and the requirement specification. This workshop would also be a good opportunity for STA to communicate to the potential suppliers the direction in which they are heading so that the suppliers could prepare for the tender request to come. At the same time, the suppliers
and other actors would have the possibility of influencing STA’s direction through the information which they could provide during the workshop.

There would naturally be barriers to overcome to hold such a workshop. During such collaboration, it must be remembered that each person has different aims for joining the group. The OptiKrea project was a research project, and no competing actors were present. In a procurement situation, all the possible suppliers must be invited to participate in such a workshop to comply with the competitive neutrality required from STA. At a workshop with competing suppliers, one could expect that the participants would not want to mention any information that might give their competitors an advantage. One could not expect the suppliers to share information about their latest inventions with their competitors. However, background information concerning the topic of the workshop should be less problematic. The workshop participants would at this stage have the possibility of influencing how STA will formulate the require-
ments by providing relevant background information that might be to their own advantage. While the process leader should keep the different aims of the participants in mind, such a workshop would be an opportunity for the community around the product to be procured to influence the development and provide STA with important information before the tender request is formalized. The suppliers would receive information on STA’s views on the product and future needs, which can provide them with a hint concerning the direction which they should choose for the development of their products to meet the requirements of the upcoming tender request. The suppliers would also have the opportunity, based on the information which STA would provide about the procurement, to contact STA and provide them with information after the workshop has taken place.

Another difficulty would be how to persuade the relevant actors to participate. The suppliers could probably be able to see a benefit for themselves. Other relevant actors such as maintenance contractors might be less interested unless they could see some benefit, despite the fact that they might possess relevant knowledge and experience regarding the topic. Different measures would be possible in this connection, one of which would be to include participation in such workshops as a condition in the maintenance contracts. However, the preferred measure would be to create an incentive that would make the workshop attractive to the maintenance contractors. One such incentive could be the benefit to be derived from participation in directing the design of products to be installed in the infrastructure. Such a workshop would be an opportunity to advocate new maintenance methods which the developed products could be adapted to. However, since it can take several years to develop new products and have them installed in the infrastructure, this might be hindered by short-term thinking. Serious actors, however, would hopefully be able to see the benefits of influencing the development of new railway products. Nevertheless, one cannot escape the fact that there is a need to improve the present procedure for awarding maintenance contracts to prevent contractors from benefiting from a less reliable infrastructure. Continuing to apply the present procedure would definitely be counterproductive. The methods developed in the research conducted for the present thesis would probably have their greatest effect in combination with a collaboration model that would give all the actors an incentive to work towards a common goal.

12.3 Adaption of the OptiKrea methods

The methods developed in the OptiKrea project did not take their starting point in the framework used at STA. Therefore, there are further possibilities of adapting them to the existing framework in order to facilitate for STA to execute the demands in EN 50126 (CENELEC, 1999) in an efficient way. Several such possibilities would be identified during a first test of applying the OptiKrea methods during an EN 50126 (CENELEC, 1999) process. A measure
that would be likely to facilitate the application is to adapt the questions provided in Table 8.8 for the problem clarification phase. Although these questions cover several of the aspects in EN 50126 (CENELEC, 1999), they could be adapted to facilitate the transfer of information and more questions could be added specifically to address the requirements in EN 50126. Similarly, it might be possible to perform a further adaption of the headings provided for the specification of requirements (see Table 2.2) to the requirements in EN 50126 (CENELEC, 1999).

12.4 Roadmap for implementation

The OptiKrea project provided a unique opportunity to test methods to be used during meetings between collaborating partners on the Swedish deregulated railway market. The participants were very satisfied with the project and were of the opinion that it would be a beneficial way of working in the future. One purpose of the project was to provide practical methods that could be implemented at STA after the end of the project. A possible road map for implementation is therefore provided in this section, in order to facilitate this step. It is suggested that the methods should be implemented first in a limited context, such as a certain division of STA for specific projects. The ideation method has been applied in three different research projects and therefore it would be very beneficial to try an application in an ordinary work situation for a normal procurement. It is proposed that this phase should be executed as an extended beta phase where the method would be adapted in several iterations in a limited context until its users were satisfied with it. At this point, it would be relevant to implement the method in a more comprehensive context, but preferably still in a step-wise manner, to be able to learn from its application. It is stressed here that there are many different contributions obtained in the OptiKrea project that may be beneficial in different situations. The entire process would not have to be applied in all situations, but certain methods or modules could be applied.

The concept selection method has only been applied once in the OptiKrea Group and it would probably be beneficial to carry out more tests before initiating an implementation. These tests could, however, be performed at STA.

Further, it is suggested that the implementation process should be followed by a specially assigned person. The methods themselves are designed in such a way that they are possible to implement without a facilitator, but it is believed that the chances of a successful implementation will increase if one person can collect the experience from its use and analyse how the methods can best be adapted to the work at STA. This person could be a researcher from academia or a specially assigned person at STA. An important issue is, of course, the provision of sufficient time and other resources to fulfil this task. After the implementation has been completed, the learning derived from monitoring the implementation itself can be formalized and used in the implementation of results from other research projects and pilot studies initiated by STA.
References


CEN (European Committee for Standardization), 2010. EN 13306: Maintenance terminology. Brussels: CEN.

CENELEC (European Committee for Electrotechnical Standardization), 1999. EN 50126: Railway applications: the specification and demonstration of reliability, availability, maintainability and safety (RAMS). Brussels: CENELEC.


Trafikverket, 2014b. *Införande av ny eller modifierad komponent i anläggningen* [Introduction of a new or modified component in the infrastructure]. (TDOK 2014:0307) [In Swedish]. Trafikverket.

Trafikverket, 2014c. *Utveckla och utforma system- och teknikkrav* [Develop and formulate system and technical requirements]. (TDOK 2014:0306) [In Swedish]. Trafikverket.


Part II: Overview of generated concepts
Introduction

This part of the thesis shows a popular scientific collection of concepts generated in the OptiKrea ideation workshops described in Chapter 5 and 6. The concepts are not a contribution by the author but an outcome of the OptiKrea project. It should be kept in mind that the concepts were developed by the participants under time pressure.

Only those concepts that were documented with written words and/or sketches by the participants are included in this overview. Of the concepts documented in this way, 23 concepts have been removed because some partner in the project wished to keep the entire or parts of these concepts confidential. A few concepts were removed because they did not make sense. The original drawings made by the participants have been worked over by a professional illustrator who has transferred them to a similar format for clarity. The original wording in Swedish has been translated to English (this translation has not been reviewed by a professional language reviewer and might contain errors). The content of this part of the thesis cannot be directly compared to the results reported in Part I of the thesis.
Ideation topic 1

How can deterioration of the track geometry in turnouts be prevented?

<table>
<thead>
<tr>
<th>No.</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image" alt="Diagram of altitude GPS with turnout settled X mm" /></td>
</tr>
<tr>
<td></td>
<td>When the turnout has settled X mm a signal is sent to a tamping machine.</td>
</tr>
<tr>
<td>2</td>
<td>The position of the turnout can be analyzed with built-in sensors such as accelerometers.</td>
</tr>
<tr>
<td>3</td>
<td>The position of the turnout can be determined by camera inspection and image processing.</td>
</tr>
<tr>
<td>4</td>
<td>Decrease the impact through suitable transition zones → less settlement</td>
</tr>
</tbody>
</table>

![Diagram showing the transition zones with rigidity](image)
5 New type of sleepers

Sight holes through the sleepers make it easy to determine if the sleepers have moved by using a laser or a laser and CCD cameras.

6 Injection casting

A hole in the sleeper for injection casting.

7 Maintenance strategy that is suited to the behavior of the turnout – perform the same measures as on the ordinary track.

8 Repeated tamping with stabilization until the track geometry is stable (3-4 times).
<table>
<thead>
<tr>
<th>9</th>
<th>&quot;SLAB-TRACK&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A slab cast in one piece under the turnout’s most vulnerable part.</td>
</tr>
<tr>
<td></td>
<td>Crossing or switch system (point machine).</td>
</tr>
<tr>
<td></td>
<td>Separate modules for the point driver might be required.</td>
</tr>
<tr>
<td></td>
<td>Resilience can be controlled.</td>
</tr>
</tbody>
</table>

![Diagram of slab track](image)

- The entire slab can be post-adjusted by injection casting. The slab is lifted to the right position with a jack or cranes.
- In combination with USP a relatively exact rigidity can be achieved.
- Under sleeper pads that are softer than today are mounted on rigid slabs → constant rigidity.

| 10 | Reduce the excitation of the turnout – attenuate vibrations from the crossing. |
| 11 | How do one get functioning transitions? For example step-by-step different pads between rail/sleeper. Gradually decreasing space intervals between sleepers. |
| 12 | Stoneblower technique → decrease settlement |
| 13 | Material in the substructure that draw off liquid. Drainage in the subballast that can stand the load from the turnout. |
| 14 | Ballast clean under the turnouts and drain. |
| 15 | Water must be removed from the ballast. Drainage system required, pipes? |
| 16 | Other material than pebbles. |
| 17 | Stoneblower material should be draining. There are also polyurethanes that can be used that are porous. |
By having a rail/rail fastening system/wheel that absorbs all movements one gets a sleeper that lies still which prevents deformation of the ballast.

Allowing weaker sleepers can be part of allowing more movement before the force reaches the ballast. Can be solved by softer under sleeper pads mounted on sleepers with large area.

Disadvantage: expensive solution.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Stiffer rails in the turnout.</td>
</tr>
<tr>
<td>24</td>
<td>Impregnate ballast with glue to decrease deformation in the beginning, or in problematic areas.</td>
</tr>
<tr>
<td>25</td>
<td>Wider and thinner sleepers</td>
</tr>
<tr>
<td><img src="image1" alt="Diagram of wider and thinner sleepers" /></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Prevent settlement</td>
</tr>
<tr>
<td><img src="image2" alt="Diagram of longer sleepers" /></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Decrease axle load.</td>
</tr>
<tr>
<td>28</td>
<td>Decrease the number of train passengers.</td>
</tr>
<tr>
<td>29</td>
<td>Build over all turnouts, for example in tunnels.</td>
</tr>
<tr>
<td>30</td>
<td>Good and large contact area between ballast and sleeper.</td>
</tr>
<tr>
<td>31</td>
<td>Change to ballast free tracks or track solutions that are considerably more stable.</td>
</tr>
<tr>
<td>32</td>
<td>Why is the ballast layer so thick? A thin layer on for example asphalt gives less possibility for the shape to change. Take drainage measures instead.</td>
</tr>
</tbody>
</table>
33 A constant area ratio along the entire turnout, for example wider sleepers in certain parts.
Modularized parts of the turnout that are mounted on slabs.

34 Injections beneath the ballast to strengthen the subballast

Polymer injections. Perhaps ballast glue works?

insulated joint

ballast

subballast

some appropriate material
Ideation topic 2

How can the transition zones between rail sections of different rigidity be designed to ensure a smooth transition?

<table>
<thead>
<tr>
<th>No.</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rail</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>rail</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>rail</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>rail</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
glued ballast

rigidity between segments

reinforced concrete

steel slab (beams)

slab track

VIEW FROM ABOVE
More solid rubber pads on the first part of the slab track.
Reinforce the weak layer, possibly excavate so that the load is low.

soften the rigid track with rail pads/under sleeper pads

rubber sleepers cast in concrete

Under sleeper pads on a number of sleepers.
Sleeper with built-in air cushion

Each side can be adjusted by means of air pressure.

Adjustable sleeper with excenter screw

Stiffer rubber pads on the first part of the slab track.

Some system to connect the modules without joints
- easy to lift

The number of transition modules I is determined by the speed and axle load in track.
Need of double beams to be investigated!

1+2 are steel beams with USP

\[ \ell_1, \ell_2 \]

\( \ell_n \) are adapted to the speed

Gradually decreasing area

same distance = standard distance
The track is positioned by adjusting screws or spacer plates.

The packed ballast is left untouched.
# Ideation topic 3

How can turnouts be protected from snow and ice?

<table>
<thead>
<tr>
<th>No.</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Coat the turnout with a slippery material (surface treatment). Use the whole of a tarpaulin (instead of two halves) to cover turnouts. Insulation underneath.</td>
</tr>
<tr>
<td>2</td>
<td>Coat the turnout with a slippery material (surface treatment). Use the whole of a tarpaulin (instead of two halves) to cover turnouts. Insulation underneath.</td>
</tr>
<tr>
<td>3</td>
<td>Spoiler wagon that directs the air flow downwards/sideways (so that snow is not sucked up and stuck at the end of the train).</td>
</tr>
<tr>
<td>4</td>
<td>Move the switch blades often – half of the snow per blade.</td>
</tr>
</tbody>
</table>
Wings in the turnout
- between the rails
- outside
(avoid lee-sides)

Limit the stroke (avoid large snow lumps)
The engine removes snow
- plough
- brush
- air
- mechanical
- laser (strong)
- IR
- glycol (focused jet) – build into turnout

Adaption of turnout.

**Aerodynamic interplay**

Lift up the turnout.
Ideation topic 4

How can good track geometry of a modular turnout be achieved and retained?

<table>
<thead>
<tr>
<th>No.</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><a href="#">Diagram: Adjustable screws towards the transition zone. “hydraulic cushions”</a></td>
</tr>
<tr>
<td>2</td>
<td><strong>Increase/control/change the rigidity</strong></td>
</tr>
<tr>
<td></td>
<td><a href="#">Diagram: Design a transition with changing sleeper space interval from the module to ordinary track.</a></td>
</tr>
<tr>
<td>3</td>
<td><a href="#">Diagram:</a></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>4</td>
<td>Adjust vertically and laterally by means of Desec or Ameca cranes.</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Cranes Diagram" /></td>
</tr>
<tr>
<td>5</td>
<td>Cradle under the rail that can adjust the angle of the rail.</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Cradle Diagram" /></td>
</tr>
<tr>
<td>6</td>
<td>Prohibit construction in curves, to decrease the risk for skewing.</td>
</tr>
<tr>
<td>7</td>
<td>Vertical and lateral adjustment by means of telescopic handling gantry &amp; lifting transportation trolley.</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Adjustment Diagram" /></td>
</tr>
</tbody>
</table>
A) Pipes with nozzles that can be attached to the stoneblower and distribute shingle to different points.

B) Add binder to the shingle so that it stays in position and do not sink down in the ballast.

Start out ordinary tamping from the crossing. Adjust to projected geometry (x y z).
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 11 | - Create the same settlement as on the ordinary track.  
    - Create the same surface pressure as on the ordinary track.  
    - Create different compression of the substructure.  

SIDÉ VIEW  
[before]  

![Diagram](image)  

This makes it possible to perform the same type of maintenance work in the turnout as on the ordinary track.  

More suitable pads with lower stiffness to compensate the increased rigidity in the turnout (a turnout is naturally more rigid than the ordinary track).  

**Use another elasticity underneath fixed crossing and underneath insulated joints, otherwise the same as on the ordinary track.**  

| 12 | Flexible fastening system where rail and pads can be adjusted vertically, both in module and transition zone (35 m from module).  

| 13 | Standardized way of working during installation of turnouts that implies a good, stable substructure (digging out and filling to frost-proof depth, requirement for drainage) that fulfills requirements for fixed track and the use of a crane during installation.  
    Recycling of worn out modules so that an “exchange system” can be arranged.  
    Document, educate, certify and procure based on this.  

| 14 | Hoist rings placed along each side so that the module can be lifted by a crane with a yoke. In connection with this a type of “stoneblower” to insert ballast under the module.  

| 15 | To use another material than ballast under the module. A material that is easier to inject and to remove, like sand. Requires a container around the module.  
    Automatic supply of material to retain an even position.  
    Can also be adjusted for each vehicle. Requires well defined modules.  

| 16 | The use of under sleeper pads decreases settlement by distributing the load. Make small holes in the module through which under sleeper pads substance can be pressed out when it is worn.  

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Use long transition zones between lower surface pressure in the turnout and higher surface pressure on the ordinary track.</td>
</tr>
<tr>
<td>18</td>
<td>Use pads between rail and steel beams to decrease settlement by distributing the load.</td>
</tr>
<tr>
<td>19</td>
<td>Install condition monitoring equipment for the track geometry (for example simple camera system for certain critical turnouts). Alarm when the track geometry exceeds the threshold values.</td>
</tr>
<tr>
<td>20</td>
<td>Rap wedge in place with a sledge hammer or fasten/adjust wedge with an adjustment screw bolt.</td>
</tr>
<tr>
<td>21</td>
<td>The concrete should not attach to the underside of the module so that the injected concrete can be removed during reinvestment and ballast cleaning.</td>
</tr>
<tr>
<td>22</td>
<td>Introduce a ballast suction apparatus that can decrease the amount of ballast in the crossing section.</td>
</tr>
</tbody>
</table>
Adjust the contact area so that it is as constant as possible along the turnout and do not change drastically anywhere (but should be a bit more area on one side).
The position of the module can be found (possible to calibrate) by means of a laser level that is placed on one point and “reading posts” in each corner.

Coarse adjustment with yoke (required after installation).
- 1 year
- 15 years
- 30 years

“Floating module”

Fine adjustment with “screw” ±30 mm, possibly an edge beam that the rail rests on. Can be done with 5 years interval, “self-adjusting screw”.

Fine adjustment can be done with an impact wrench, for example adjusting screw.

The position of the module can be found (possible to calibrate) by means of a laser level that is placed on one point and “reading posts” in each corner.
29 How can we guarantee a good ballast surface pressure after adjustment?
a) 10 mm superelevation
b) machine stabilizer

30 Fixed point for screw nut.

31

32

33 Introduce absolute coordinates with GPS pucks in each corner that are read regularly by a measurement wagon.
|   | Develop an accessory to today’s tamping machines, so that the module can be adjusted.  
<table>
<thead>
<tr>
<th>New way of vibrating where the “forks” are rotating, not bent inwards.</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>Camera system that monitors the track geometry for certain critical turnouts.</td>
</tr>
</tbody>
</table>
Ideation topic 5

How can satisfactory drainage of the track superstructure of turnouts be achieved?

<table>
<thead>
<tr>
<th>No.</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arrangement to not prevent water transportation from the track due to ducting, noise barrier etc.</td>
</tr>
<tr>
<td></td>
<td>The cable troughs today create a &quot;bath tub&quot; that shut in water.</td>
</tr>
<tr>
<td>2</td>
<td>Secondary ditches that remove the flow of water before it reaches a railway yard.</td>
</tr>
<tr>
<td>3</td>
<td>Build sleepers that have measurement sensors that can be read off through “balises/RFID technique”.</td>
</tr>
<tr>
<td>4</td>
<td>An alternative is to have drainage systems like pumps that start automatically when necessary (only applies to critical turnouts).</td>
</tr>
</tbody>
</table>

railway yard

Noise barrier

superstructure

substructure

primary drainage

secondary drainage

railway yard
Wind turbine powered by wind from passing trains and ordinary wind.

The train pushes the water sideways.

For turnouts – automatic (electric) bilge pump (like for boats) in the drainage wells.

Lower the ground water level through pumping

“Tunnel drill” Ø ½ m
8 Pressure plate (from engines and wagons) that gives pump impulse to the drainage pumps.

9 Especially the brace pits in the turnout must be drained! Can be done though sloping pipes with large enough diameter. Like in Kmb.

10 Equipment for cleaning of drainage pipes alternatively repairing of the pipes from the inside. Requires that the design allows this.

11 The drainage pipes that lead to the drainage wells should have a larger diameter. Periodic maintenance with flushing (a sewer cleaning truck probably has the right technique).

12 IMV100 can measure stiffness. Is there a relation between the stiffness and the water level → PhD project. Measure the change in dynamic response in the embankment to estimate the drainage function → IMV100 could measure this nationally.
Re-lining of old pipes where the re-lining still has drainage holes (How do you do it if they are supposed to be inflated?). Cleaning with “rotating spray brush”, like a sewer cleaning, adapted to a large diameter.

Develop measurement methods to find water levels and degree of pollution. Ground penetrating radar can find the present water level, but a continuous measurement method might be required, like a temperature sensor on a rod that is used to check the ground frost depth over time. Spot test with tube. Check of drainage well.

Method for testing the drainage function - measure the flow at certain points at a certain type/amount of precipitation. “Continuous” measurement technique with floats in wells. Or automatic bilge pump.
Leads to completely ballast-free solutions on bridge. Both snow and water free with good drainage.

No ballast required here! Mount sleepers and rail on a frame so that the water automatically disappears.

Drainage pipe with filter that prevents particles from entering the pipe.

The filter can be accessed at certain points and cleaned. Vacuum cleaners can be cleaned by reversing the flow.
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Introduce ballast free track.</td>
</tr>
<tr>
<td>25</td>
<td>Learn from other countries → inform → budget → plan</td>
</tr>
<tr>
<td>26</td>
<td>Introduce holistic thinking in the PIA group.</td>
</tr>
<tr>
<td>27</td>
<td>Employ more people that are passionate about the issue.</td>
</tr>
<tr>
<td>28</td>
<td>Simplicity during procurement.</td>
</tr>
<tr>
<td></td>
<td>Turn-key contracts without direction are dangerous.</td>
</tr>
<tr>
<td>29</td>
<td>STA should control how the drainage is carried out and apply the stick and the carrot. Inspections should never be performed by those who carry out the maintenance work that should be performed.</td>
</tr>
<tr>
<td>30</td>
<td>The stick and the carrot is a well-tried method to make the human (the employee) to work for the employer.</td>
</tr>
<tr>
<td>31</td>
<td>Combine &quot;basic staffing&quot; in contracts with long-term measures.</td>
</tr>
</tbody>
</table>

**Diagram:**
- **track/turnout**
- **ballast/macadam**
- **concrete trough**
- **pump**
- **drainage**

Place the pump (boat pump) in the drainage wells and spurt the water far out in the forest.

For extremely difficult conditions were natural out-flow cannot be realized.
<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>Long-term contract to create open ditches during track replacement with a guarantee of 500 km of new drainage per year for 20 years (STA the owner?). Also think of appropriate “penalty measures” if the contract is not followed.</td>
</tr>
<tr>
<td>33</td>
<td>Give line section managers earmarked money that must be spent within 3 years (carefully described what they should do and how often!).</td>
</tr>
<tr>
<td>34</td>
<td>Today, there is no consensus that drainage issues are important. Information material and LCC calculations should be developed. Arrange a special ”drainage meeting” with lecturer + decision meeting about drainage measures. Make a project where the effects of drainage are made clear.</td>
</tr>
<tr>
<td>35</td>
<td>Start pilot project, document and give an account.</td>
</tr>
</tbody>
</table>
| 36  | Initiate processes within STA that take care of incorrect constructions that have lost their function. For example:  
- ditches that have silted up.  
- pipes that have stopped up.  
Replace units that are not functioning with the latest technique. Through clear impact assessment and LCC data! |
<p>| 37  | Avoid building railway in waterlogged areas! |
| 38  | In troublesome areas, make a new section so that one has time to do it right. |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
| 40 | **Problem inventory!**  
Scan the environment to find places where water transport from the environment can give problem (contour map). |
| 41 | Collaborate with forest owners (farmers), that they get paid to keep the ditches open. |
| 42 | Give adjacent landowners the mission to drain in return for money, including performance based agreement. |
| 43 | STA obtains their own specialized machinery for embankment- and ditch cleaning or possibly a procurement of this. Make planes for the entire country and cycles of about 5 years. |
| 44 | Start a joint project between railway, road and municipalities for recycling of ballast masses and disposal of pollutants. Start production promptly. |
- Ditch cleaning as a national procurement gives enough volume for special machinery and simplifies planning.
- Ditch masses should be taken onboard for recycling. For example, ballast suppliers should be imposed to receive residual matter for recycling.
- During cycles of about 5 years the entire network should be covered by driving from south to north and back. Should also include well tried-out drainage cleaning, in the form of pipes.

Recycling of ballast to ballast supplier guarantees cost-efficiency and environmentally friendliness. Volume gives good prices. Coordination with road projects, dam constructions, and so on.

A fixed number nationally distributed for good logistics.
Masses (collected after the track replacement) or are recycled at the ballast suppliers. Municipalities and road maintainers can also take their masses there. Masses are thrown out into the forest when possible.

Several pipes with filling like hawser suck water over ridge. Thirsty trees = birch or particularly thirsty bushes are planted out.

The sun is a good pump 200-400 l per day. Connect hoses to thirsty trees. Let the tree do the job.

Talk about automatic pump!
| 50 | ![Plant thirsty trees](image) |
| 51 | Introduce animals that like water. |
| 52 | More efficient snow removal to avoid problems with melting away of snow (during Spring). Maintenance coupled to the drainage problem areas. Snow removal or drainage? Or maintenance? |
| 53 | Don’t melt so much snow in the turnout, the melted snow run downwards (shut off the turnout heating, or use other techniques). |
| 54 | Ballast cleaning in the traditional way (ballast cleaning machine). Avoid “bath tub effect”. |
| 55 | Enlarge ballast cleaning to the entire width of the superstructure, to prevent a “bath tub”  
Deep ducting (today 3 dm, tomorrow 12 dm, cables can be placed on 3 dm). |
| 56 | Place wires in the ground/substructure to be able to measure the drainage function. Double wires are used in the isolation of district heating pipes. When the conductivity increases there is a leakage. |
| 57 | More drainage wells at the turnout. Pump when necessary. |
58 Place submersible automatic electric pump where water accumulates.

59 Use the technique of packed soil = concrete. To prevent water from entering.

60 Layer for better outflow from superstructure

Textile, plastic, asphalt, etc (or metal sheet such as steel plate).

61 Layer for better outflow from superstructure
62. **Pack a first layer of ballast so that the stone surface gets an outward angle.**

   ![Diagram of ballast pack](image)

   **Drainage mats with 1 dm space interval between.**

   The material must be resistant so that it's not cut to pieces by the ballast.

63. **Dig up ditches and pack the material hard together on the embankment so that the track is raised.**

   ![Diagram of embankment with ditches](image)

   Distribute the material to both sides to decrease the building height (the overhead contact wire can be adjusted 5 dm).

64. **This area should be kept free from plants and small particles. Machine that can clean this can be adapted if the inclination is small (today's embankment might be too steep). Then no time is required in the track.**

   ![Diagram of cleaned area](image)

   Collect bushes for chipping, the chip price today is 300 SEK/m³.

   **Supplement with drainage mats (closely situated, that is several layers).**
Bush clearance to avoid roots in the drainage. Today, there are machines that apply biocides directly on the new stumps. This implies the least possible spreading of biocides and prevents a stump from becoming 10 new shoots.

Allow heating of ballast (kills roots). Combine with environmentally friendly biocides.

Introduce inspection requirements on existing drainage system connected to action plans. Investigate all possibilities.
Impregnate the ballast with environmentally friendly biocides.

Effective control of thick roots.
Exposure to radioactivity is effective but dangerous perhaps?

Black ballast absorbs more heat, gives better evaporation.

Darker ballast and embankments → warmer embankment → evaporation
Suitable distance so that the clearing machine can access.

For turnouts: drainage pipes for the brace pit with large enough diameter should be standard.

* inclined sheet iron -> automatic outflow
* corrugated profile (like roofing sheet)
* the angle α not too large.
* advantage: No moving parts (like pumps) that can cause trouble

Water drinking vegetation on a suitable distance.

+ Check how the problem is solved in other countries that have waterlogged areas (like Holland).
+ Agreement with farmers that get the task to control the drainage and to fix it when necessary. The result is controlled by STA.

"The railyard ditch" (and cuttings or where outflow can be difficult to construct along the track).
### 80
Moisture meter in the ballast of critical turnouts. When the moisture exceeds a certain threshold value measures are taken to decrease the moisture. Install sumps in these turnouts in which the pump starts automatically when the water level is too high.

### 81
**National perspective:**
Finance a new approach to the issue by means of the PIA project.
1) Study projects that have been carried out
   - cost
   - design – is standard good enough?
   - long-term results
   - what influence do turn key contracts have?
2) Make LCC calculations of conventional methods compared to a new approach
3) Present the results in a larger forum and discuss, carry out procurement where expensive machines are allowed to be competitive (LCC calculation part of the procurement).
4) Rationalize the handling of masses by not only procuring ballast supply but also recycling.
5) Pay attention to constructions that have lost their function (analysis of condition).
Part III: Papers
Paper 1
Abstract
The deregulation of the Swedish railway system started in 1988 with the vertical separation of train operations from the railway infrastructure, shortly afterwards followed by public procurement of train operations. Contracting of maintenance followed in 2001 upon the split-up of the infrastructure manager into a client and contractor. The purpose of introducing contracting was to increase efficiency. The present study investigates, by collecting primary data at the actor level, how the maintenance of the Swedish railway system has been affected by contracting. The results suggest that the current design of railway maintenance contracts motivates a short-term maintenance perspective and a preference on the part of contractors for corrective maintenance. Further consequences are an insecure work situation for railway workers, the replacement of specialized machines with standard machinery and a loss of competence. The consequences are leading to a deteriorating railway infrastructure. To improve the situation, the infrastructure manager needs systematically to try out alternatives, of which a non-exhaustive list has been presented in this paper, to find empirical evidence for a successful concept.

Keywords: railway system, maintenance, deregulation, performance-based contracts, contracting, competitive tendering, Sweden

1 Introduction
The deregulation of the Swedish railway system started in 1988 with the vertical separation of train operations from the railway infrastructure. The deregulation was driven by the need to find new possibilities of financing the railway sector and increasing efficiency [1]. Railway infrastructure ownership, investments and maintenance responsibility were transferred to a national authority, the Swedish Rail Administration. Public procurement of train operation on railway lines began shortly afterwards. In 1998, new policies led to the split-up of the Rail Administration into a
client and contractor. Competition in the area of railway maintenance started in 2001 with contracts in the low-traffic parts of the network. Again, the purpose was to increase efficiency [2]. By 2012, 95% of the maintenance carried out on the State-owned railway infrastructure was being procured publicly [3]. In 2010, the Rail Administration and several other Swedish authorities were merged to form the Swedish Transport Administration (STA). STA is responsible for the strategic planning of roads, railways, ship transport and aviation, as well as investment in and the operation and maintenance of roads and railways in Sweden. These responsibilities include the procurement of railway maintenance. The maintenance contracts are performance-based and include performance requirements. The contractor determines what maintenance measures are to be taken in order to fulfil these requirements. The choice of maintenance methods will influence the cost, quality and time consumption. Such contract types are expected to yield higher productivity and innovation. The deregulation of the Swedish railway infrastructure is still an on-going process and the final outcome has yet to crystallize. Continuous monitoring of the process is necessary to understand its impact. Research on the outcomes of Swedish railway maintenance procurement is limited, probably because it is a rather new phenomenon. In 2006, Stenbeck [5] studied how performance-based incentives in a railway maintenance contract influenced contractor performance. Lingegård [6] made some conclusions about the current state of Swedish maintenance contracting, but focused on the benefits and challenges of replacing construction and maintenance contracts with integrated product-service offering (IPSO) contracts. The development of maintenance procedures has important implications for the long-term sustainability of the railway network. The purpose of this study is to understand, by gathering information at the actor level, how the maintenance of the Swedish railway system has been affected by the deregulation process and, in particular, the procurement and tendering process. Primary data gathered by interviewing experienced people regarding the situation in the maintenance sector will narrow the gap in our present knowledge of the effects originating from the deregulation process. Descriptive empirical research offers an insight into the circumstances that lie behind statistics and numbers.

2 Method

The data collection for this research study was performed using in-depth semi-structured interviews. One person from STA, one from a contractor company and one from a supplier company were interviewed. These respondents were consciously chosen since they had worked at the Rail Administration before the deregulation started, have an engineering background and have been working within their respective segments of the railway industry for 20-30 years. The interviewees have worked in different positions and possess a broad knowledge of the overall situation within their respective organizations. Therefore, the validity of their answers is high regarding the situation which they can describe. The respondents are detailed in Table 1. The starting points for the interviews were open-ended questions about what consequences the deregulation has had on the maintenance and what the collaboration between the contractors, infrastructure manager and suppliers look
like. The interviews were recorded and transcribed. In addition, data were collected during field studies and meetings where the opportunity emerged to pose questions to and to discuss related issues with field workers and consultants, for example. These observations were recorded through making notes. The study used an inductive approach. The gathered data were carefully categorized based upon the emerging themes. The thematic analysis linked the data gathered from each respondent and aimed to see the bigger picture not accessible by each individual. The results of the analysis, presented in Section 3, were sent to and confirmed by all the interviewees. The validity of the results is increased by the fact that the respondents have corresponding views, although they work in different segments of the railway sector and therefore have different perspectives on the situation.

<table>
<thead>
<tr>
<th>#</th>
<th>Organization</th>
<th>Current position</th>
<th>Railway experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>STA</td>
<td>Specialist (former team leader and maintenance manager)</td>
<td>24 years</td>
</tr>
<tr>
<td>2</td>
<td>Supplier</td>
<td>Development manager</td>
<td>31 years</td>
</tr>
<tr>
<td>3</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>27 years</td>
</tr>
</tbody>
</table>

Table 1: Details of the main respondents

The study is based on a small number of interviews and, therefore, even though the respondents work in different sectors, an obvious limitation of the study is the generalizability of the results. The approach selected is justified by the detailed information which could be accessed, facilitating the design of further studies of the subject. Another limitation of the study is the fact that Swedish maintenance contracts currently have different content, and different contract specifications might lead to different consequences. Future research can show if the conclusions from the present study represent the general situation within the field.

3 Results

Section 3.1-3.4 summarize the primary data gathered from the respondents. The data were in accordance and the results have been validated by all the respondents. The number after each quote shows which respondent has been quoted according to the numbering in Table 1.

3.1 The challenges of procuring maintenance contracts

Before the deregulation, maintenance was centrally controlled. Maintenance was planned and performed with the aim of profiting in the long-term future. When visiting a site to perform maintenance, a railway worker would simultaneously look for other faults and attend to them. Replaced parts were installed at a site where they would have the right performance or were taken care of and used as spare parts later. Today the focus is being transferred to short-term profits:
“Earlier, you thought more preventively […] that [part] is in rather a bad condition, we’ll change it now that we’re here anyway […] Whereas now […] you go there and do [the job] and absolutely nothing else because you’re not being paid for that.” (3)

“When I started […] everyone’s reflex action, when you were out doing a job, was to throw [the replaced parts] onto vehicles and then you drove somewhere and took care of everything that was worth taking care of.” (1)

“When you have to go out and repair and fix things that you shouldn’t have to go out and [fix] at all, [this] also disturbs the train traffic, which in turn costs money. Our experience is that this is also much worse after this deregulation. Extremely short-term economic profit is governing a lot of things.” (2).

The reason for this development is how maintenance contracts are procured. The purchase of operation and maintenance for the Swedish railways is governed by the Swedish Public Procurement Act. STA acts as the purchaser of services with the desired performance through public procurement. STA specifies the criteria for a procurement in a tender request and contractors show their interest in the contract through tenders that specify the cost at which they will perform the work. The tenders are then evaluated by STA based upon the criteria in the tender request.

The competition for contracts is very tough. The basic services, typically including operation and daily maintenance, safety inspection and emergency duty, are included in the basic agreement. The contractors bid a fixed price for the basic services during a contract period of five years, whereas several of the emergency operations and special operations are charged for per hour or operation. Charging for services in addition to the basic services is a way for the contractor to increase their income:

“You try, of course, to bid as low as possible, otherwise you don’t get the contract, and then you need to chase those additional costs all the time […] and everyone does that of course.” (3)

Obtaining a contract for provision of the basic services is a way to be in the right position when other jobs appear or are announced. In this way money can be earned. One consequence is that it is advantageous for the contractors if the railway infrastructure is in a bad condition, since this means more emergency work that can be charged for outside the basic service agreement:

“So you bid extremely low and then you hope there will as much additional [work] as possible. And the worse the infrastructure is, the more money the maintenance contractor will earn.” (3)

In other words, the contractor earns more money on corrective maintenance than on preventive maintenance. The basic agreement of the contract is complied with using as little resources as possible, just so that the requirements are fulfilled.

One difficulty in procuring a maintenance contract is the lack of knowledge of the events taking place and the actions required in the future:
“Purchasing a normal contract, e.g. [for] replacing a turnout, that’s a completely different matter. [The contract says,] ’Do this and do that,’ and you can easily do an inspection. Procuring a maintenance contract is something completely different.” (3)

The contractors have difficulties interpreting the tender requests and foreseeing what will have to be done to fulfil the performance requirements. Examples of typical considerations are how much snow removal will be necessary (depending on the weather conditions), the length of time during which the maintenance staff will be able to access the track to carry out their work (influencing how much time the railway workers will have to spend waiting to complete a job), and the condition of the track section in question (determining, for example, how many times the track will have to be aligned to sustain the agreed quality). The contractor is not completely responsible for the track section and there is a possibility that STA, during the time of validity of the contract, will decide to take actions that might benefit the contractor (e.g. upgrade the railway foundation) or might not benefit the contractor (e.g. increase the traffic). It is very important that the tender requests cannot be interpreted in different ways and this is an issue of great current interest. The tender writers try to find the loopholes in the request that open up the possibility of cutting costs in the tender and having STA paying for additional work outside the basic part of the contract later on. Deviations from the contract specifications, as interpreted by the infrastructure owner or the contractor, have occurred and have led to legal cases, because the specifications of what should happen in the event of deviations have not been detailed enough in the contract. Eventually it comes down to winning the contract by bidding lower than one’s competitors, even though it is clear from the start that one’s bid is unrealistic:

“This money doesn’t at all cover the work we must do.” (3)

“It’s better to calculate so that we lose 10%, because if we don’t get [the contract] we lose 20%.” (3)

If the contract is won, the quest to cover the deficit in other ways starts. How large the profit or loss will be will not be known until the contract period has ended.

Another difficulty is to define the responsibilities. Many actors are involved and responsibility easily falls between two stools. Naturally, the procurement staff at STA tries to improve the tender request and contracts by revising their content. The people responsible for calculating the tenders must therefore adapt to every new request. There is a desire for a standardized contract so that the calculations can be made in a routine manner and fewer resources will be wasted.

3.2 The risk of losing competence

The railway infrastructure has a long lifespan. Consequently, a long-term perspective is beneficial and STA wants to take the right decisions about the maintenance of the installations. To be able to achieve this, STA requires information and knowledge regarding the usage, maintenance and degradation rate of every part of the infrastructure and the right competence to judge the effects of
these parameters. Resources can then be guided to where they have the greatest effect.

The railway-specific knowledge possessed by STA is decreasing, resulting in difficulties to perform adequate procurements. STA must not only know what to ask for, but also be able to judge if the contracts are fulfilled in an honest way. As the contractors decide what measures to take in order to maintain the track, STA does not have full information about how the track is being maintained, making it harder to correlate performance to executed actions.

A natural part of procuring maintenance contracts is changing the contractor due to a more competitive bid. The turnover of contractors is high. A possible reason for this is that, once a contractor has gained experience of a certain railway section, they become familiar with the track and understand what price tag is relevant:

“You get to know a railway infrastructure, you get to know a track section, you know what it costs and then you win [the contract] one [time] and then the next [time] you’re kicked out, because then you know how to set the price and then there’s always someone who [bids] lower.” (3)

However, knowledge and experience of a track section are essential in maintenance work. Knowledge about where on the track problems normally arise makes both preventive and corrective maintenance more efficient:

“In the long run you develop such things as good competence, good knowledge about your infrastructure and thereby fast repair work.” (3)

Changing the contractor maintaining a railway section has several consequences. The organization of the contractor leaving the area in question is dismantled. Typically, the personnel living in the area are hired by the new contractor. This gives the possibility of some transfer of track knowledge to the new organization that has to be established when a new contractor wins a bid. However, knowledge is lost in every change of contractor, as well as the trust of the workers in the different employers. In addition, the workers experience an insecure work situation, as they need to be rehired when the contractor is replaced.

3.3 Specialized machinery is replaced by standard machines

High railway quality is dependent on specialized machines that are typically expensive to acquire. When the Rail Administration was in charge of the operation and maintenance nationally, the machines could be used within the entire railway infrastructure. Today there are no requirements concerning machines in the contracts since the specifications are performance-based. As the contractors are smaller organizations, the utilization of the machines in each company is not high enough to motivate the purchase of specialized machinery, due both to the reduced length of the infrastructure to maintain and to the fact that the contract period is not long enough to make it possible to write off the depreciation of the machines. In addition to the problem of the length of the contract period, one has the typically long
delivery time for advanced machines, which can be 1-2 years. Certain machines would give the possibility of performing services cheaper, which would result in lower tenders, if the contract length could motivate a purchase:

“And then you can understand this phenomenon, that if every contractor is to own their machine and run it 200-300 hours a year, then it will be expensive. But if one owns the machine and succeeds in [hiring it out], then it will be very cheap instead.”

The contractors do not take the risk of purchasing specialized machinery in case they might not win another contract. Instead, excavators with different accessories are typically used. Although excavators can be used for several different purposes, they cannot give the same quality as specialized machines. Hence, the railway is wearing unnecessarily fast. However, multi-purpose excavators are a cheap solution, at least in the short term:

“Good machines give good track quality, there is no doubt about that. But what you find now is all types of cheapest possible solutions.”

While contractors are avoiding buying new machines, the machinery dating from the time before the deregulation and still in operation is growing older and less reliable. By using specialized machines, many operations could be performed in a shorter time, meaning that large operations that are currently being performed during weekends could be carried out during night breaks. In this way, the utilization of machines and personnel would be more efficient.

One possible direction to follow in the future is that different companies would specialize in certain areas, such as tamping, to achieve the necessary volume to motivate the purchase of special equipment. Local contractors would then attend to basic operation and maintenance. Another alternative, which has been used in some instances, is to rent machines from abroad. Guarantees of work from STA tied to the purchase of specialized machinery would also be a possibility.

The problem of utilization levels is not solely a problem of machinery. The organization as a whole must continuously be adapted when contracts are lost and won and the work load subsequently changes. This is especially tough when projects are procured late and forecasts from STA are changed.

3.4 Following up contractor performance

There are three different sets of rules and regulations describing how the operation and maintenance should be executed, governing safety inspection, maintenance inspection and periodical replacement. While safety inspection is included in the basic service agreement of the maintenance contracts, maintenance inspection and periodical replacement are typically recommended but optional. This is connected to the fact that the contracts are performance-based and, therefore, STA is not supposed to direct how the contractors maintain the railway infrastructure, as long as it meets the agreed performance criteria. However, the recommended rules and
regulations are not followed, either because of a loss of competence or because no one bothers to implement them, which affects the results in a negative way:

“Previously there used to be a set of rules and regulations for how the [infrastructure] should be looked after out along the track. And it’s our experience that one is losing competence or isn’t bothering about [applying] them fully. One isn’t lifting things in the proper way, one isn’t doing the jobs one should be doing and this of course affects the result negatively.” (2)

As the contractors carry out the maintenance inspections themselves, they can adjust the inspection comments to suit their current workload. At the end of a contract period, an inspector is to examine the railway infrastructure to assure that the condition has been maintained at the agreed level. The inspector examines the signal, track and electricity systems. It is difficult to be an expert in all these areas and, consequently, the contractor has an advantage, knowing each area in more detail and being familiar with the railway section concerned. Moreover, there have been instances when the final check at the end of the contract period has not been performed. Each track section has a track manager who follows up the condition of the infrastructure. Meetings are held regularly where all the parties concerned discuss different issues concerning the section:

“It turns into an investigation of what’s been done and what hasn’t been done, more than focusing on the actual point at issue, the problem out on the track.” (3)

The contractors, trying not to exceed their budgets, try to find someone else who could be responsible or to delay work:

“So the more brazen you can be, the more you can charge. And the more you can bear your client telling you off, the cheaper it will be for you, the more you will earn from it [...]. You earn money by delaying things.” (3)

Should the quality of the infrastructure be found to be unsatisfactory by the inspector at the end of the contract period, this will not result in any consequences in the next procurement, since the contractor with the lowest bid will win regardless of earlier performance. One means of control that STA has is the option of rewarding a contractor by extending their contract by two years. It is up to STA to decide if it wants to exploit this option. Earlier it was quite common for STA to do so, but it is not as common any more:

“Perhaps because STA is not satisfied with the service and thinks that the grass is greener on the other side.” (3)

4 Discussion

From the results, the following key consequences of the deregulation and contracting of Swedish railway maintenance have crystallized:

- The maintenance perspective has shifted from the long term to the short term.
Corrective maintenance is to some extent replacing preventive maintenance.
The follow-up assessment of the contractors’ work is not satisfactory.
Railway workers are experiencing an insecure work situation.
The technical competence at STA is decreasing.
Specialized machines are being replaced with standard machinery.
The railway system is deteriorating.

The following measures for remedying the consequences have emerged from the results:
- Standardize the tender requests.
- Evaluate maintenance procurement from a life-cycle-cost (LCC) perspective.
- Value quality and earlier performance to a greater extent.
- Use contract extension options tied to performance.
- Introduce longer contract periods.
- Establish a machinery pool from which contractors can rent specialized machines.
- Give contract guarantees tied to the purchase of specialized machinery.
- Initiate national or Nordic procurement of specialized services.
- Maintain the technical competence at STA.
- Establish an adequate inspection function.

Figure 1 illustrates how the consequences and proposed measures are related. A non-exhaustive number of additional measures, drawn from the literature, have been added. The consequences and measures are discussed in more detail in Sections 4.1-4.6.

The literature on maintenance procurement for the Swedish railways is limited, probably due to the rather recent introduction of such procurement. However, competition among train operators and road maintenance contractors was initiated at an earlier stage, and therefore we can make some comparisons with the literature published in this area, as well as the literature on maintenance contracting in other countries.

4.1 Short-term profit and corrective maintenance

The current design of the maintenance contracts motivates the contractors to work in an unsustainable manner to maximize their short-term profit, shifting the operative maintenance perspective from the long term to the short term. As STA has an explicit policy of proactive maintenance [7], it is alarming that the results suggest that the contractors prefer corrective maintenance.

4.1.1 Low tender bids

There exists a common understanding that replacing a public monopoly or a publicly guaranteed private monopoly with competition between companies results in
improved efficiency [1]. However, a certain level of competence is required to design tender requests and evaluate bids to avoid the negative consequences of obtaining unrealistic or strategic bids [8]; otherwise the possible gains may be lost. Sweden, Germany and Great Britain have experienced unrealistic bids which have resulted in train operating companies being unable to survive on the amount of subsidy in the bids [9]. For example, the Swedish State in 2003 had to transfer a large amount of money to Swedish State Railways (SJ) to avoid bankruptcy, partly because several of SJ’s contracts for passenger services were very unprofitable, since SJ had won tendered contracts with excessively low bids [1]. Maintenance contractors make unrealistically low bids to win the contract and then chase
additional work, e.g. corrective maintenance, which must cover the loss generated from the basic part of the contract. In a review of STA’s procurement of roads and railroads in general, the Swedish National Audit Office [4] pointed out that contractors often have the possibility of compensating for a low bid with additional reimbursements. Some mechanism for time and material reimbursement is difficult to avoid in contracts, since there will always be conditions which the contractor and procurer cannot foresee.

Just as in the case of train operating companies, the survival of maintenance companies is uncertain due to repeated tendering. Both Swedish and British train operating companies have had problems winning contracts for the same franchise twice in a row [1]. High turnover may be unavoidable when the market learns and adapts to new conditions. One possible reason for a company placing a low bid is that the company actually is operating in a cost-efficient manner. However, the present study points to other explanations in the case of maintenance contracting. The true cost becomes obvious to the contractor once they have been maintaining the line for a period of time. Nash and Smith [10] note that it was expected in the British market that the incumbent would start with an advantage in terms of knowledge of costs and markets. However, they found instead that thus far only a small fraction of the franchises obtained through refranchising had been won by the main incumbent. Yet many incumbents went on to win other franchises in different parts of the country. This rather suggests that overoptimistic bids might be the result of poor information, leading to the “winner’s curse” [10], if not merely the result of a strategic decision taken by the contractor. In contradiction to this result, Lingegård [6] found, when interviewing contractors and STA staff, that a contractor in charge of maintaining a track section was likely to win the next procurement, since their organization was already established in the area and could therefore offer a lower bid than their competitors. As the procurement of maintenance started in 2001 and was subsequently introduced at a slow pace, not that many reprocurements have actually taken place. Further, the turnover of contractors may be different depending on the geographical area and respondents may therefore have different impressions. Such geographical variation is beyond the scope of this study, but would be an interesting topic for future research based on comparisons with actual figures from STA. Predatory bids could also be a reason for low bids, as has been found among train operators [8], although this has not been found in the present study.

### 4.1.2 Valuing soft parameters

The quality of maintenance performance seems, among other things, to correlate to the fact that STA cannot take earlier contractor performance into account during tender evaluation. This reduces the contractors’ incentive to aim for a good reputation through overachievement [4]. Companies with higher standards will either be outperformed by bolder companies or have to adopt a similar behaviour to survive on the market.

Evaluating maintenance procurement from a life-cycle-cost perspective, as well as valuing quality and earlier performance to a greater extent, is in the present study
suggested to shift the focus from the contract price only and to counteract a short-term perspective and preference for corrective maintenance. According to Stenbeck [5], so-called soft parameters (i.e. other parameters than price) were initially given a 30% weight during the bid selection at the Road Administration, which was the former counterpart of the Rail Administration in the area of road transport. Due to complaints from the contractors that the bid selection had become too arbitrary and subjective, the weight was gradually reduced to 3.5% in 2004. The former Rail Administration assigned a 70% weight to price in their bid selection in 2004 and experienced similar complaints to those received by the Road Administration. The soft parameters in both agencies were, however, based on documentation submitted with the offer and not actual performance in the field [5]. Hence, it is possible that another approach could be beneficial which would include an objective way of assessing the soft parameters.

During the first refranchising in Great Britain, the franchising body opened up the possibility of submitting a wide range of proposals rather than being prescriptive concerning what investments and improvements in service the offer should contain. During tender evaluation, issues such as realism and past delivery had to be weighed against ambitious plans for the future, which is a much more difficult task than simply comparing the subsidy bids for a stipulated set of services [10].

It is clear that it is difficult to find the right balance between the different contents in a tender request and a systematic approach to developing an efficient tender process is necessary to reach empirical conclusions. The use of contract extension options appears to be an undervalued means of motivating the contractors to achieve higher performance. By specifying in detail what must be fulfilled for an extension option to be exploited, one could increase the competitive pressure during the contract period. There are examples of successful bonus and fee systems leading to a decreased number of faults and improved quality, see e.g. [11].

4.1.3 Longer contract periods

Contractors try to maximize their profit during the contract period, setting the long-term perspective aside. In her research work, Lingegård [6] proposes the introduction of IPSO contracts, where construction and maintenance are simultaneously procured, to mitigate short-term thinking. She investigated the possible effects of long-term contracts and found that they could provide long-term planning, making it easier to make investments due to the longer payback time. Excessively long contracts would, however, make it difficult to estimate the cost and the contractors would have to charge extra for the risk, possibly making it more expensive for the infrastructure owner. It is interesting to note that some British train operation franchises were let for up to 20 years, this period being conditional on the operator’s performance and on them implementing ambitious investment plans, to encourage greater investment on lines that required major investments for their operation [10]. Railway infrastructure has a long lifetime and it is probably not reasonable to sign maintenance contracts for the complete lifetime of a track section. Therefore, even if contract lengths are increased, other measures also need to be
taken to assure a long-term perspective. A balance between extending the contracts and keeping the competitive pressure is required.

### 4.1.4 Managing uncertainty

It has been found that one difficulty in procuring performance-based maintenance contracts is the lack of knowledge of what will happen in the future and, consequently, what maintenance measures will be required to fulfill the contract. Stenbeck [5] described three success factors for performance-based specifications: how well the owner can describe and define the contracts, how compliance is measured and how deviations are handled, i.e. how the contractor is penalized for non-fulfilment or rewarded for excess delivery. As noted by Nash [9], in a contract involving parties whose knowledge of future circumstances cannot be perfect, situations are likely to arise (e.g. increased traffic) which require details to be renegotiated, entailing the risk of opportunistic behaviour on the part of either party. Therefore, to the list of success factors must be added a specification of the circumstances in which contracts can be renegotiated and how they are to be renegotiated. None of these success factors appear to be sufficiently fulfilled during maintenance contracting.

### 4.2 The technical competence at STA is decreasing

The results suggest that there is a risk that STA will have difficulties formulating a successful contract if they lose technical knowledge. An issue that is perhaps even more crucial is that the information asymmetry between the infrastructure manager and the contractor can be expected to increase if the technical knowledge of the STA staff, especially the inspectors, decreases. This will give the contractor the possibility of concealing technical errors and omitting to fulfill contract requirements. Further, technical competence is crucial for making the right decisions regarding the railway infrastructure. Therefore, there are several incentives to maintain technical competence at STA.

### 4.3 Follow-up assessment of the contractor’s work is inadequate

The findings suggest that the follow-up assessment of the contractors’ maintenance work is unsatisfactory, entailing the risk of unfair competition. The Swedish National Audit Office [4] noted in its review of the procurement of roads and railways that the contractors mentioned experiencing insufficient monitoring of the fulfilment of contract conditions. This issue was also raised by STA’s internal audit. According to Stenbeck [5], inspections are even more important in performance-based contracts than in ordinary contracts, because, without inspections, the contractors have an incentive not to meet demands and thereby save money in both the short and the long term. It seems quite naive that Swedish contractors should be operating without regular inspections from either the infrastructure owner or an independent third party. The contractors are themselves performing the asset
inspections, giving them an incentive to adjust the results to suit their current work load or to fulfil the contract requirements.

Deviations from the contract specifications, as interpreted by the infrastructure owner or the contractor, have occurred and have led to legal cases because the specifications of what is to happen in the case of deviations have not been detailed enough to let the parties come to the same conclusion. Stenbeck [5] found that technical acceptance limits were not written explicitly. The contract owners hinted, during interviews, that technical acceptance limits was part of the contract code, but it was embedded in the code and only understandable by experts. This definitely creates a risk of diverging interpretations, opportunistic behaviour from the contractor and legal cases.

An adequate inspection function is required and to facilitate this function, detailed and clear contract specifications must be in use. By standardizing the contracts, the different parties would not risk arguing about new specifications. In addition, standardizing tender requests has been proposed as a way to simplify bidding and thereby reduce the barriers for smaller companies [4]. It seems reasonable to assume that this would also save costs for larger companies, as well as simplify the evaluation procedure at STA.

4.4 Railway workers experience an insecure work situation

Not only the maintenance companies, but also their employees experience an insecure situation because of repeated tendering. When a new company wins a contract, it typically rehires the staff of the previous company, creating a situation where the new employer can dictate the conditions. There is a risk that railway workers will leave the business for other opportunities, impairing the organization’s competence and local track knowledge. This was avoided in the British reform by organizing the passenger transport business in train operating companies franchised by private operators [12], although this was probably not the main aim of the arrangement. Whoever won the bidding took over that company, including its staff and assets, for the franchise period [10]. However, as pointed out by Nash and Smith [10], this arrangement imposes less pressure on labour costs. Longer contract periods would, even if they do not solve the problem of insecurity, at least improve the situation.

4.5 Specialized machines are replaced with standard machinery

After the British separation of the infrastructure from the train operations, rolling stock was leased from separate rolling stock leasing companies [10]. Yvrande-Billon and Ménard [12] found that the short train operating contracts imposed by the reform translated into a reluctance on the part of the rolling stock companies to invest in new rolling stock due to the uncertainty of recovering a long-term investment. In the end, a standardization of rolling stock resulted, since standardized rolling stock could be rented to any train operating company. In Sweden, the length
of the maintenance contracts has led to contractors using modified excavators and aging machines dating from the time before the deregulation, instead of specialized modern machines that could enhance the quality of the infrastructure. The contractors are not confident enough about winning another contract to dare to purchase high-cost machines. Similarly to the present findings, Stenbeck [5] found that cost reductions had primarily been achieved through staff cuts and using less expensive and less advanced standardized machinery, rather than process and product innovations, as a result of maintenance contracting for the Swedish road and railway infrastructure.

To ensure the use of more specialized machines, the present study proposes that a machinery pool should be established from which contractors could rent specialized machines, or that STA should connect the guarantee of a contract to the purchase of specialized machinery. Train operators can already lease rolling stock, so the establishment of a machinery pool for maintenance contractors seems to be a surmountable challenge. However, if contractors are to be induced to use the machines, they must see an economic benefit. If their own organization is not using all of its capacity, they will probably prefer to use their own resources instead of renting machines. One alternative would be for STA to demand the use of certain machines for certain services, but this would not be aligned with STA’s aim of using performance-based contracts.

Increasing the volume of services could be another way to motivate investments in machinery. This can be accomplished by longer contract periods and procuring certain services, e.g. tamping, on a Nordic or national basis. Local contractors would perform basic maintenance. Lingegård [6], however, found that contractors were concerned that too many larger contracts would lead to a locking of the market, excluding some contractors for years. Further, she ascertained a concern that smaller contractors would not be able to bid for large contracts. Staff from STA, however, mentioned that international contractors would probably be interested in large-size projects, thereby increasing the number of competitors [6]. Lowering the entry barriers through introducing a machine pool could allow smaller contractors to bid for large maintenance contracts. The Swedish National Audit Office [4] found in their review of the procurement of roads and railroads that large contracts in general had not led to higher productivity than small contracts and identified the causes to be larger risks, less competition, and the practical difficulties involved in formulating and following up performance-based specifications. The opposite strategy was tried by the former Road Administration in 2004-2005 [5]. Maintenance areas were split up into sub-areas according to both space and scope, with the aim of increasing competition by making it easier for small and specialized companies to participate. Bidders could also combine sub-areas that were adjacent with regard to space or scope and offer a volume discount per combination. Some financial savings were achieved with the method, but the administrative costs were higher than intended, since the computer-based evaluation had to be supplemented with manual methods [5]. Coordination between different maintenance actors performing different activities must not be neglected. Just as in the case of contract lengths, the contract volume is a trade-off between different parameters.
4.6 The railway is deteriorating

The results and the effects described in Section 4.1-4.5 suggest that the quality of the railway infrastructure is deteriorating. In contradiction, the Swedish National Road and Transport Research Institute found that maintenance contracting in Sweden had resulted in 14% lower costs, with no effect on the failure rate [7]. However, typical indicators of failure rates, such as broken rails, indicate failures of a maintenance regime and do not provide an adequate indicator of the quality of that regime [13]. The track does not react at once to inadequate maintenance; the consequences are delayed because it takes time for initiated defects to propagate to the extent that they become a danger. Moreover, when the deterioration has progressed too far, it is not possible to regain sufficient quality through maintenance and the component affected must be replaced. There is a risk that poor maintenance in the long term will lead to safety concerns [13] and high costs, as the railway infrastructure will have to be replaced prematurely. How to define quality is a challenge and analyzing in detail how quality should be measured is beyond the scope of the current study. An obvious concern regarding cost savings is whether they have been achieved at the expense of quality and life cycle costs; without clarity regarding quality effects, economic savings are difficult to value [5]. Due to the resistance of the railway infrastructure to poor maintenance, it is possible that deteriorating quality may not be observed in the quality indicators utilized until failure occurs.

5 Concluding remarks

Yvrande-Billon and Ménard [12] suggest that merely changing the ownership does not provide the incentives and outcomes typically desired when performing deregulation. In reforming public utilities, regulators must take into account how the conditions which they impose influence choices [12]. The authors of the present study agree with the assertion that the deregulation of railway maintenance in itself must not be of evil, but that it was realized with quite naive expectations concerning the outcome and insufficient insight into what is required from the procuring agent to avoid pitfalls. Unfortunately, many of the problems experienced after the initiation of competition in the case of train operation and road maintenance have reappeared after the introduction of competition in the area of railway maintenance.

The Swedish National Audit Office [4] is worried that the low number of bids is decreasing the competitive pressure. The results of the present study suggest that, in the case of railway maintenance, increasing the competition in the current situation might be detrimental to the long-term performance of the railway infrastructure. Decreasing the prices even more as the contractors are already struggling might result in further neglect of the maintenance required to keep the railway infrastructure in shape. Some contractors might also go bankrupt, which would, at least temporarily, decrease the number of bidders. Instead, the first step is to find out how maintenance can be procured in a satisfactory way that takes the long-term perspective into account, includes appropriate incentives and follow-up assessments,
and keeps the competitive pressure throughout the contract period. STA must facilitate the contractors’ work by providing adequate prognoses and must announce procurements well in advance so that the contractors may have a chance to adapt their organization without unnecessary costs. Finding the right contract type and quality indicators is not easy and requires a systematic approach that takes into account the needs of both the infrastructure manager and the contractor. Once the necessary groundwork has been carried out, the focus can again be turned to enhancing the competition.

Investment in railways is problematic in that the time horizons are long and the risks high, making it relatively unattractive to the private sector [10]. Alexandersson and Rigas [8] noted that the European railway industry so far appears to have responded to external factors. To alter this tendency, they propose that more new actors should be attracted from other sectors. The present findings suggest that an important factor in accomplishing this in the maintenance sector is addressing the machinery issue. Facilitating access to specialized machines would increase the number of companies that would be able to compete in public procurement of maintenance contracts.

A critical point is that STA is actually the only actor that possesses comprehensive knowledge of the entire Swedish railway infrastructure and is therefore in the best position to optimize it with a holistic perspective. A key question is how maintenance contracts can be designed to avoid suboptimization and give incentives for the provision of cost-efficient maintenance that will preserve the technical level of the railway infrastructure in the long term.

From the findings of the present study we draw the conclusion that the current design of railway maintenance contracts motivates unsustainable contractor performance and, consequently, accelerated wear of the railway. To improve the situation, STA needs systematically to try out alternatives, of which a non-exhaustive list has been presented, to find empirical evidence for a successful concept.

### 6 Future work

To generalize the presented results the inclusion of several more respondents, especially representatives from different contractors, is suggested in order to obtain complementary perspectives on this research topic. Further studies of documents and reports from STA, as well as governmental reports, would clarify in more detail the results expected from the deregulation process and the outcomes that have materialized so far, in addition to providing data triangulation. By performing a comprehensive mapping of the contract designs that have been tried and the circumstances in which they have been tried, as well as studying a sample of maintenance contracts in more detail through a multiple case study, one would obtain an indication of which directions are most fruitful to follow.
Acknowledgements

The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged.

References

Paper II
Ideation methods applied in a cross-functional inter-organizational group: an exploratory case study from the railway sector

Anna Malou Petersson1 • Jan Lundberg1 • Matti Rantatalo1

Received: 29 May 2015 / Revised: 25 May 2016 / Accepted: 12 September 2016 / Published online: 27 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The conceptual design phase is a critical step, since it influences the subsequent steps during product development with regard to cost, quality and performance. Previous research has focused on cross-functional teams within an organization. However, many product development projects benefit from the participation of members from different organizations, not least during the conceptual design phase of technical products, where it is essential to consider different aspects of the product-to-be. Therefore, we conducted an in-depth case study of a cross-functional inter-organizational group testing ideation methods in a real-life setting within a development project in the railway sector. The group comprised participants from an infrastructure manager, a supplier, a maintenance contractor and research bodies. The tested ideation methods were Method 635, the gallery method and the SIL method. The participants found working in a cross-functional inter-organizational group to be beneficial both during the group-analysis of the topics and during the generation of ideas on how to address the ideation topic. Applying the ideation methods to the ideation topics facilitated the sharing of information between participants, and the diversity of the group manifested itself in several ways during ideation. Overall, the gallery method was most popular, and the SIL method was least popular among the participants.

Keywords New product development • Concept generation • Cross-functional groups • Inter-organizational groups • Ideation methods • Method 635 • Gallery method • SIL method

1 Introduction

The conceptual design phase is a critical step, since it influences the subsequent steps during product development with regard to cost, quality and performance (e.g. Rubenstein 1994; Okudan and Tauhid 2008). To focus attention on different aspects of the product-to-be as early as possible, in order to avoid becoming locked in an unsuitable solution, working in cross-functional teams has become increasingly popular. Further, as functional diversity is introduced in work groups, more creative and innovative ideas and solutions may be procured (van Knippenberg et al. 2004). Heterogeneous groups are expected to outperform homogeneous groups or non-interacting individuals in fulfilling creative tasks, because heterogeneous groups have a greater range of skills and resources to draw from (Milliken et al. 2003). However, the presence of a diverse group and its potential range of skills and abilities do not guarantee the effective utilization of those skills and abilities. An expert in one domain will often not know what information has value for an expert in another domain (Straus et al. 2011), and information sharing is essential to make use of a group’s available informational resources collectively. For various reasons, the group’s members may be less willing to contribute to the group work, or other processes might inhibit information sharing and therefore decrease group performance and productivity (e.g. Pelled and Adler 1994; Jassawall and Sashittal 1999; van Knippenberg et al. 2004).
When a group’s composition is diverse, the use of structured ideation strategies may be able to bolster cognitive activities associated with the creative process by forcing the consideration of many alternatives (Milliken et al. 2003). Several researchers and practitioners have focused on developing practically useful ideation methods that aim to encourage the ideating participants to propose a large amount of ideas on a particular issue during a limited time frame. Shah et al. (2000) classified ideation methods into two categories, intuitive and logical methods. Intuitive methods stimulate the unconscious thought processes of the human mind, whereas logical methods involve systematic decomposition and analysis of the problem (Shah et al. 2000). In the present study, we focus on intuitive methods. The social psychology literature on brainstorming (Osborn 1957), the most well-known intuitive ideation method, is extensive, and the claim by Osborn (1957) that brainstorming results in enhanced efficiency has been refuted by a large body of research. Although it has been found that people feel more productive working in a group than working on their own (Stroebe et al. 1992; Paulus et al. 1993), the combined output of individuals ideating on their own (nominal groups) outperforms that of verbally interactive groups in terms of non-redundant ideas and quality of ideas (e.g. Mullen et al. 1991; Byron 2012). Since verbal interaction contributes to process losses (Mullen et al. 1991), several so-called brainwriting methods have been suggested. The common feature of these methods is that, instead of communicating ideas verbally, ideas are documented by each participant in silence by writing. Several researchers have highlighted the importance of sketching during engineering design (e.g. Shah et al. 2001; Yang 2009) and so-called brainsketching techniques communicate ideas through sketches rather than words. Hybrid ideation methods combine brainwriting and/or brainsketching with verbal interaction. Compared to brainstorming, other ideation methods have received comparably less attention, and many methods have been suggested, but not studied from a scientific point of view.

Studies performed on ideation methods have mainly involved inexperienced users, predominantly students, working under experimental conditions. Previous research has focused on cross-functional teams within an organization. However, many technical medium- to large-scale product development projects benefit from the participation of members from different organizations in the product development team. As outsourcing is becoming increasingly popular and deregulation has opened up previously closed markets in many countries, different functions are being spread out over several actors specializing in different functions related to a product. There are examples of tests of ideation methods with multidisciplinary teams (Chulvi et al. 2012, 2013; Seidel and Fixson 2013). Rxfelt et al. (2011) tested an approach for cross-company team-work involving the customer in service innovation, and this approach included the application of ideation methods. However, despite the fact that previous research has shown examples of both the customer’s (Hannola et al. 2009) and the subcontractor’s (Liker et al. 1998) importance for the introduction of improvements, no in-depth “real-life” tests of intuitive ideation methods with teams consisting of members from different organizations have, to the best of the authors’ knowledge, been conducted. Such teams differ in several ways from those in the majority of previous studies. Firstly, members from different organizations are professionals representing several different market actors. According to Straus et al. (2011), little is known about what happens in groups whose members bring different organizational cultures and political agendas. Secondly, such groups typically consider complex engineering topics that require domain-specific knowledge. According to Mesmer-Magnus and DeChurch (2009), highly complex task domains typically require specialists with different expertise and backgrounds to integrate information in order to achieve a high-quality solution. Kavadias and Sommer (2009) found analytically that the nature of the problem matters and that experimental evidence in the brainstorming literature might have been influenced by simple ideation topics that did not represent real situations. Thirdly, such groups themselves define and own the topics which are to be the subject of ideation. Bolin and Neuman (2006) suggested that differences concerning the ideation topic, e.g. in the abstractness of the topic or in the strength of the incentive for high performance, would affect the outcome. In addition, previous research on idea generation (Sutton and Hargadon 1996; Jackson and Poole 2003; Gish and Hansen 2013) suggests that in order to understand how ideation can be successful in reality, the context must be taken into account. According to Shah et al. (2000), case studies regarding the use of ideation methods in industry would be very valuable for evaluation of the usefulness of these methods in real life. However, it is difficult to gain access to case studies in real-world settings and to conduct them (Shah 1998). The present study provides a rare opportunity to test ideation methods in a real project.

The main objective of the study is to test and evaluate the gallery method, a variant of Method 635, and the SIL method in a cross-functional inter-organizational group in a real-world setting through an in-depth case study. By combining an account of the ideation outcomes with a presentation of the participants’ views on the methods and an analysis of their behaviour during ideation, the study aims to provide a rich picture of how the methods work in the given context. A secondary objective is to learn about other effects that arise due to the composition of the group.
Several ideation methods are relevant to the present study. In the original Method 635 (Rohrbach 1969), a brainwriting technique, six participants write down three solutions to the problem and then pass the sheet of paper to their neighbour, who reads through the ideas and contributes three further ideas or developments. When the sheet has passed between all the participants, the ideation stops (Rohrbach 1969).

Alternative brainwriting techniques that have features similar to those of Method 635 are the brainwriting pool (Geschka et al. 1973) and the pin card technique (VanGundy 1984). The basic brainsketching technique involves each group member individually sketching their ideas on sheets of paper and passing the sheets on, after a few minutes, to another member, who continues to sketch ideas, using the already generated ideas as a source of inspiration for new ideas (van der Lugt 2002). Shah et al. (2001) suggested an alternative brainsketching technique called C-sketch. In C-skech, the participants work independently on developing a sketch (no textual description is allowed) of their proposed solution to the ideation topic, and when the cycle-time ends, the sketch is passed on to their neighbour, who adds, modifies or deletes aspects of the solution during the next cycle. The solutions pass between all the participants, so that when the session is completed, each solution will have been worked upon by every participant, and the number of solutions equals the number of participants (Shah et al. 2001).

Examples of hybrid ideation methods are the gallery method (Pahl et al. 2007) and the SIL method (VanGundy 1984). The gallery method starts with an idea generation step of 15 min where the participants use sketches, supported by annotations as necessary, to describe ideas, followed by a step where the sketches are hung on a wall so that all the group members can see and discuss them for 15 min. During an additional 15 min ideation step, the participants use the inspiration from the discussion step to further develop ideas (Pahl et al. 2007). The SIL method also starts with an individual ideation step (VanGundy 1984). Thereafter, two members read aloud one idea each, and all the group members try to combine these two ideas verbally into one idea. During the next step, a third group member reads their idea aloud, and the group then tries to integrate this idea with the previous one, and the process continues until an integrated solution is found (VanGundy 1984).

2.2 Related scientific studies

van der Lugt (2002) compared a version of brainsketching that allowed the participants to explain their ideas briefly after each round of idea sketching to brainstorming. It was found that significantly more ideas were generated during brainstorming than during brainsketching, but more connections with previous ideas were made during brainsketching (van der Lugt 2002). Comparing C-skech to Method 635 (allowing only textual description) and the gallery method in an experiment involving mechanical engineering students and practising designers, Shah et al. (2001) found that C-skech outperformed Method 635 in the three measured areas of quality, novelty and variety of designs generated and was better than the gallery method with regard to novelty and variety. The gallery method was better than Method 635 concerning these three parameters, suggesting that sketches are a useful means of communication in idea generation in the field of engineering design (Shah et al. 2001). It should be noted that the number of ideas was not compared between the methods, since that quantity was fixed in the ideation method instructions.

Linsey et al. (2011) conducted a study on how two key factors of brainsketching, C-skech, Method 635 and the individual phases of the gallery method influenced the quantity, quality, novelty and variety of ideas. The two key factors were as follows: (a) how the ideas are displayed to other group members (“rotational view” or “gallery view”) and (b) how the ideas are communicated between group members (written words only, sketches only or a combination of words and sketches), resulting in six different group ideation conditions. The experiment involved mechanical engineering students working on the design of a device to shell peanuts. They found that most ideas were generated when the groups used rotational viewing and communication by means of a combination of words and sketches. A larger number of high-quality concepts were generated with gallery viewing in combination with sketches only than were generated using any other condition. There were no significant differences between the results for the different conditions with regard to novelty or variety (Linsey et al. 2011). Linsey and Becker (2010) conducted a complementary study by letting nominal groups ideate on the peanut sheller problem given the same amount of time. They found that real teams using rotational view conditions developed a larger number of ideas than equivalent nominal groups. Gallery viewing was better than nominal groups when words and sketches were used to represent the ideas (Linsey and Becker 2010).

Rexfelt et al. (2011) tested Method 635 and brainstorming in two different cross-functional groups, each consisting of three persons from a manufacturing firm and three persons from a customer firm, as one step in a service development process. The results suggested that the affiliation of the participants did not affect their productivity, whereas their experience of development work and comprehension of the specific method used did have an effect.
on their productivity (Rexfelt et al. 2011). The group using Method 635 struggled and generated fewer ideas that were suitable for the context compared to the brainstorming group. Rexfelt et al. (2011) explained this finding by the fact that Method 635 was carried out in silence, and thus the participants could not be directed in the right direction neither by the coaches nor the other participants in the group, and that the group applying Method 635 did not have access to an inspirational poster that illustrated a vast range of service categories, and which was present when the brainstorming group generated ideas.

2.3 Expectations

A number of expectations were present as to what outcome the methods would yield in the present study. The gallery method gives the participants the opportunity to work thoroughly on their ideas and revise them after a presentation and discussion step, where they can obtain feedback from other participants who possess complementary knowledge and experience. It was therefore expected that the gallery method would generate more complete and detailed concepts of higher quality, in the form of sketches, than the other methods. Method 635 was expected to generate more ideas than the gallery method and the SIL method, and this expectation was based on the instructions for Method 635, the time it provided for individual ideation and anecdotal reports concerning its use among students in an engineering design course that had taken place for several years at Luleå University of Technology (LTU). Findings made by Linsey et al. (2011), obtained under different conditions, partly supported our expectations from Method 635 and the gallery method. To the authors’ knowledge, there is no previous study of the SIL method published in the literature. Based on the instructions for the SIL method, it was believed that it could help the participants overcome mental barriers by forcing them to try to combine ideas. It should be noted that all the methods were performed under the condition that the participants could describe ideas by both sketches and words according to their own preferences.

2.4 Context

The case study was performed in the Swedish railway sector and focused on turnout development. The turnout is a vital part of the railway infrastructure, and a failure of a turnout, especially in a critical location, can cause significant delays and societal costs. In 2012, turnout-related failures were among the top ten causes of hours of disruption in Sweden (Trafikverket 2013). Therefore, the OptiKrea project, run by Luleå Railway Research Center at LTU, was initiated with the intention of promoting the technical development of turnouts. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining and utilizing the railway, as well as performing railway research. The idea behind the OptiKrea project was that by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions would be found. A parallel goal of the project was to develop working methods facilitating innovation that are tailor-made for the railway sector and can be used in the future. At the heart of the project is the so-called creative team, which consists of representatives from each collaborating actor. This team is cross-functional in the sense that it represents different functions that are important when developing a turnout, i.e. research, design, manufacturing, management, maintenance and disposal.

3 Methods

The empirical findings in the present study were obtained using an exploratory case study. A case study is suited to situations where it is impossible to separate the studied phenomenon’s variables from their context and offers an in-depth description and analysis of a bounded system (Merriam 2009).

3.1 The case and participants

The case in question concerned concept generation in an inter-organizational cross-functional team aiming to innovate within the field of railway turnouts (as a part of the OptiKrea project). The team had six participants, four railway professionals and two academics. The participants were carefully chosen based upon their experience from the railway sector and knowledge of their respective organizations. Hence, they were able to give representative views on what would be useful for the actors which they represent. Each of the railway professionals has 20–30 years of experience from the railway sector. Two professionals work with turnout-related issues at the infrastructure manager, the Swedish Transport Administration (STA), one professional works at an international company manufacturing turnouts and one at a contractor performing maintenance. The academics work with railway-related projects in close cooperation with STA and industry. In Sweden, the infrastructure manager typically runs its research and development projects in cooperation with or through research institutes and universities, and therefore, academic researchers are relevant participants in the group. Details of the participants are given in Table 1. All of the participants were males. The case study is unique in that it takes place in a real-world setting and brings together members from four different organizations. STA is the only
national buyer of services and products from the contractor
and supplier, resulting in an asymmetric power relation
between the organizations which the participants represent.
Similarly, academia is dependent on funding from STA in
many projects, and several different universities and uni-
versity divisions compete for such funding. None of the
participants had tried any of the ideation methods before,
except one of the academics, who had previously tried
Method 635 and the gallery method.

3.2 Procedure

The group met four times over a 4-month period. Between
the meetings, the participants worked with their normal
tasks and did not use the ideation methods. At the first
meeting, the group was introduced to ideation methods and
had a large number of established ideation methods pre-
sented to them (see Table 2). In a discussion with the
researchers, they chose three of the presented established
methods to be tested: Method 635 (Rohrbach 1969; Pahl
et al. 2007), the gallery method (Pahl et al. 2007) and the
SIL method (VanGundy 1984). The basic criterion for the
choice was that the method should be rather simple to
perform. The gallery method was chosen because the group
believed that it would give well-devised concepts, the SIL
method because it was believed that it could help the group
to overcome mental barriers and Method 635 because it
was believed to be capable of providing many ideas and
because it was different from the others. It should be noted
that the participants wanted to be able to use both sketches
and words when describing their ideas, and therefore the
version of Method 635 used in this study is similar to the
modified version found by Linsey et al. (2011) to be the
most effective in terms of the number of ideas generated.
One reason for letting the participants choose the ideation
methods to be evaluated was to get the participants
engaged and motivated to remain involved in the study,
which naturally took time from their normal work load.
Another reason was that their choice would indicate what
type of method their organizations would be prepared to
use in the future, which would increase the probability of
such methods being implemented during future product
development in collaboration between different actors. The
participants were asked not to make a choice based mainly
on their personal preferences, but on what they thought
could be useful to the organization which they represented.

<table>
<thead>
<tr>
<th>No.</th>
<th>Organization</th>
<th>Position</th>
<th>Field of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>Product and method development</td>
</tr>
<tr>
<td>B</td>
<td>Supplier</td>
<td>Development manager</td>
<td>Product development of turnouts</td>
</tr>
<tr>
<td>C</td>
<td>STA</td>
<td>Specialist</td>
<td>Maintenance management of turnouts</td>
</tr>
<tr>
<td>D</td>
<td>STA</td>
<td>Specialist</td>
<td>Track and turnout</td>
</tr>
<tr>
<td>E</td>
<td>Academia</td>
<td>Professor</td>
<td>Mechanical engineering and product development</td>
</tr>
<tr>
<td>F</td>
<td>Academia</td>
<td>Postdoctoral research fellow</td>
<td>Applied acoustics and signal processing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brainstorming</td>
<td>Osborn (1957)</td>
</tr>
<tr>
<td>Electronic brainstorming</td>
<td>Gallupe et al. (1991)</td>
</tr>
<tr>
<td>Individual brainstorming</td>
<td>Finke et al. (1992)</td>
</tr>
<tr>
<td>Method 635</td>
<td>Rohrbach (1969); Pahl et al. (2007)</td>
</tr>
<tr>
<td>Brainwriting pool</td>
<td>Geschka et al. (1973)</td>
</tr>
<tr>
<td>ICR grid</td>
<td>Wodehouse and Ion (2012)</td>
</tr>
<tr>
<td>Gallery method</td>
<td>Pahl et al. (2007)</td>
</tr>
<tr>
<td>Collective notebook</td>
<td>VanGundy (1984)</td>
</tr>
<tr>
<td>SIL method</td>
<td>VanGundy (1984)</td>
</tr>
<tr>
<td>Lead user method</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Delphi method</td>
<td>Dalkey and Helmer (1963)</td>
</tr>
<tr>
<td>Quick-Delphi method</td>
<td>Developed by the last author of the present paper</td>
</tr>
<tr>
<td>TRIZ</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Synectics</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
<tr>
<td>Bionics</td>
<td>Herstatt and Kalogerakis (2005)</td>
</tr>
</tbody>
</table>
During each ideation session, one person was attending via video conferencing due to travel restrictions imposed by his organization. It is a typical desire from employers and employees that telework should be possible. Due to delays caused by technical issues concerning this person’s participation via video conferencing, all the methods could not be carried out exactly as planned. This had an impact on the ideation times. This was the only undesired outcome of the technical issues; in all other respects the participant in question contributed on the same level as the other participants. The actual execution of each method was presented in Table 3, which also shows how the other participants. The actual execution of each method in Table 3, which also shows how much time the participants spent working individually (in the case of Method 635 with inspiration from the other participants’ ideas) and as a verbally interactive group, respectively. Different ideation topics were chosen for each ideation method, since using the same topic for each method may result in the solutions from the previous ideation session(s) being reused in the subsequent session(s) and comparison of the ideation outcome between the methods would not be meaningful. Instead, we aimed to find equivalent problems that were based on actual needs, were domain-specific and open-ended, and had a large amount of possible solutions. Firstly, the participants were asked to come up with topics which they wanted to work on within the OptiKrea project and which, based on their experience, they thought were equivalent. Thereafter, the authors of the paper made a review of the suggestions to pick out the topics that best fulfilled the equivalence criteria. The participants were informed of this selection and agreed upon the ideation topic about 1 week before each ideation session. The chosen ideation topics are presented in Table 4. The issues on which the ideation was performed required at least a basic technical understanding in general and railway-specific knowledge in particular. At the start of each session, one of the participants presented the ideation topic, after which a common group-analysis of the topic took place in order to identify causes of problems and known solutions to problems. Thereafter, one of the methods was tried. The first author presented the instructions for the method concerned to the group by

<table>
<thead>
<tr>
<th>Method</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual/group/total ideation time (min)</strong></td>
<td></td>
</tr>
<tr>
<td>635</td>
<td>1. Each participant works individually for 5 min and comes up with at least three suggestions on how to address the ideation topic. The ideas are sketched and/or written down on a sheet of paper.</td>
</tr>
<tr>
<td></td>
<td>2. When the time is up, each participant gives the sheet containing the suggestions to the neighbour on their left.</td>
</tr>
<tr>
<td></td>
<td>3. Each participant reads through the suggestions on the sheet of paper which they have received from the neighbour on their right and adds improvements/comments to suggestions, combines suggestions to form new suggestions, and/or uses suggestions as a source of inspiration to come up with new ideas, during a total time of 5 min. One is allowed to ask the neighbour on one’s right what is meant by a suggestion that has been received.</td>
</tr>
<tr>
<td></td>
<td>4. When the time is up, each participant gives the sheet of paper which they have most recently been working on and which now contains the work of two participants to the neighbour on their left, and step 3 is repeated.</td>
</tr>
<tr>
<td></td>
<td>5. The process continues until each sheet of paper has been passed between all the participants, i.e. when it has been returned to the person who started working on it on a blank sheet of paper.</td>
</tr>
<tr>
<td>Gallery</td>
<td>1. The participants individually sketch and/or write down suggestions on how to address the ideation topic on a sheet of paper for 15 min.</td>
</tr>
<tr>
<td></td>
<td>2. The sheets of paper are attached to a wall. The group gathers around one of the sheets hanging on the wall and the creator of the suggestions explains them to the other participants, who give constructive feedback. The group then moves on to the next sheet, which is explained, after which feedback is given, and this process is continued until all the participants have received feedback on their suggestions. Each participant can use approximately 5 min for presenting their ideas and receiving feedback.</td>
</tr>
<tr>
<td></td>
<td>3. Each participant takes down their sheet of paper and works individually on that sheet for 5 min to develop their suggestions or come up with new ideas using the feedback which they have received from the other participants and using the other participants’ suggestions as a source of inspiration.</td>
</tr>
<tr>
<td>SIL</td>
<td>1. The participants individually sketch and/or write down their suggestions on how to address the ideation topic during 10 min.</td>
</tr>
<tr>
<td></td>
<td>2. Two participants are randomly selected and each presents one of their suggestions to the rest of the group.</td>
</tr>
<tr>
<td></td>
<td>3. All the participants try to combine these suggestions into one concept by interacting verbally and sketching/writing on a whiteboard.</td>
</tr>
<tr>
<td></td>
<td>4. When the group is finished with the first two suggestions, a third group member presents another suggestion. The group then tries to combine this suggestion with the suggestion which resulted from step 3. Alternatively, a fourth suggestion is presented by any participant and an attempt is made to combine this fourth suggestion with the third suggestion.</td>
</tr>
<tr>
<td></td>
<td>5. The process of presenting suggestions and trying to combine them with each other continues until all the ideas have been presented or the time is up.</td>
</tr>
</tbody>
</table>
means of a projector, ran a timer to keep track of the time, announced when the group should move on to the next stage of the method, and answered questions about the instructions during the session. During the test of the SIL method, the participants became so confused about what they were supposed to do or reluctant to do it that the first author had to take an unplanned active role in leading the session.

3.3 The views of the participants

After each ideation session, a short group interview regarding the participants’ experience of the applied method took place to capture the participants’ immediate reactions to the ideation method which they had tried; individual interviews about the ideation methods took place after all the methods had been tried. After all the methods had been tried, each participant answered a questionnaire presenting different statements about each method. The participants answered on a continuous scale from “Do not agree at all” to “Agree completely,” by making a mark on a line which was about 100 millimetres in length and where 0 represented “Do not agree at all.” The position of the mark was measured in millimetres from 0 with a ruler and divided by the total length of the line. We calculated the average value for each statement and, despite the small number of participants, the average values for the subgroups academics and railway professionals. The participants were also asked to rank the methods. The questionnaire formed the basis of an individual semi-structured interview. The interviews were audio-recorded and the transcribed recordings, along with the questionnaires, served as the basis for analysing the participants’ views on the methods.

3.4 Outcome of ideation sessions

3.4.1 Number of concepts and ideas

The number of concepts suggested by the group as possible solutions to each ideation issue was counted, as well as the number of concepts that were sketched. For the purpose of this study, a concept is regarded as one idea or a combination of two or more ideas which serves as a solution to the problem on its own merits. A concept may address only one subpart of the problem, and it may be possible to combine a certain concept with other concepts. As a starting point, concepts were defined as sketches and text that clearly belonged together and were marked off from other sketches and text on the sheets of paper. It was a straightforward task to identify the concepts on the sheets of paper. The concepts generated using the SIL method were distinguished in a different way from the concepts generated with Method 635 and the gallery method, because there were no sheets of paper with solutions after the ideation session in the SIL method. The concepts had instead to be identified by means of photographs of the whiteboard where the participants had made sketches of and written their ideas and on the basis of the transcript records. The concepts were constructed based on how the participants talked about them. Typically, a certain time interval would concern the same concept and then the group would switch over to a discussion of something new that was clearly separate from the preceding concept. The concepts turned out to be a very broad collection of solutions, ranging from a single abstract phrase to technical solutions which several participants had worked on. In all the examples of concepts and ideas presented throughout this paper, the text or parts of the text have been translated from Swedish into English. Examples of abstract concepts generated by a single person are shown in Table 5. These concepts are abstract since they describe something that should be carried out, but do not state how it should be accomplished. An example of a concept which several participants had worked on is Concept 4 in Table 6. Four different people have added ideas to the concept, consisting mainly of words and two sketches. Concept 5 in Table 6 shows an example from the gallery method of a sketch with annotations which one person had worked on. The concepts contained not only physical solutions to the problem, but also organizational or operational and maintenance-based solutions. These concepts clearly also represented possible solutions to the ideation topic.

### Table 4 Ideation topics used for each ideation session

<table>
<thead>
<tr>
<th>Method</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>How can track geometry deterioration be prevented in turnouts?</td>
</tr>
<tr>
<td>Gallery</td>
<td>How can the transition zones between rail sections of different rigidity be designed to ensure a smooth transition?</td>
</tr>
<tr>
<td>SIL</td>
<td>How can turnouts be protected from snow and ice?</td>
</tr>
</tbody>
</table>

### Table 5 Three examples of abstract concepts and the ideas contained in these concepts

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>635</td>
<td>Recurrent tamping with stabilization until the track geometry is stable (3–4×)</td>
<td>Tamping Stabilization Recurrent 3–4×</td>
</tr>
<tr>
<td>2</td>
<td>635</td>
<td>Decrease axle load</td>
<td>Decrease axle load</td>
</tr>
<tr>
<td>3</td>
<td>635</td>
<td>Ballast clean under the turnouts and drain</td>
<td>Ballast clean Drain</td>
</tr>
</tbody>
</table>
### Table 6  Examples of concepts and the ideas contained in these concepts

<table>
<thead>
<tr>
<th>No.</th>
<th>Method Sketch and text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>635 Slab cast in one piece under turnout’s vulnerable parts Crossing Switch system Injection casting Lifted to right position Jack Cranes Under sleeper pads Softer than today</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Gallery</td>
<td>Steel slab Angle Bolt fixation</td>
</tr>
</tbody>
</table>
Several concepts which were worked on by several participants included suggestions on how to accomplish some function or detail that was incompatible with another suggestion, either physically, because the suggestions could not be contained in the same concept, or logically, because it made no sense to use both alternatives at the same time or to solve the same function. When a larger part of the concept (more than half of its inherent ideas) remained the same upon changing between two different variants, it would be counted as one concept and the variants would be captured in the number of ideas, as is the case with the concept shown in Concept 4 in Table 6. This concept consists of 9 different ideas. Of these ideas only two would not be compatible, i.e. using jacks and cranes to lift the slab, which cannot be done simultaneously. Since this incompatibility involved less than half of the ideas, this concept was not split into two. If more than half of the ideas of a concept were incompatible, it was split into two or more different concepts.

When applying the SIL method, several times the group discussed a new concept and then made a suggestion on how this concept could be combined with a previously discussed concept. In these cases, the new concept was incorporated in the previously suggested concept if, as in Concept 6 in Table 6, less than half of the ideas in the combination of concepts were incompatible. In Concept 6 in Table 6, the first suggestion was to protect the turnout by means of a shutter. Another participant then saw the possibility of using the shutter as a part of a fence which he had been thinking of. In connection with a discussion about how to use locomotives to remove snow, e.g. with a plough or brush, it was suggested that locomotives could increase the air current through the fence by putting a big plate in front of it.

The total number of non-redundant ideas was counted for each method. Because the concepts were of different types, it was challenging to define an “idea” as a unit for counting purposes. We used the procedure developed by Linsey et al. (2005, 2011) as a starting point. In this procedure, the basic definition of an idea is something that solves one or more functions in the functional basis (Linsey et al. 2011), with the functional basis defined as a standardized set of function-related terminology that allows repeatable and meaningful function representation (Hirtz et al. 2002). However, as a result of the ideation topics, several generated concepts were not meaningful to try to fit into a functional basis and other generated concepts required completely different functional representations. For the purpose of the present study, we instead defined ideas as the units which each concept could be systematically decomposed into, and which could be expressed as a key phrase consisting of a verb phrase containing a maximum of one verb, or a noun phrase. In this way, it was possible to capture ideas from concepts of different types.

### Table 6 continued

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Sketch and text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>SIL</td>
<td><img src="image" alt="Sketch" /></td>
<td><img src="image" alt="Ideas" /></td>
</tr>
</tbody>
</table>

- **No.**: 6
- **Method**: SIL
- **Sketch and text**: ![Sketch](image)
- **Ideas**: ![Ideas](image)

Several concepts which were worked on by several participants included suggestions on how to accomplish some function or detail that was incompatible with another suggestion, either physically, because the suggestions could not be contained in the same concept, or logically, because it made no sense to use both alternatives at the same time or to solve the same function. When a larger part of the concept (more than half of its inherent ideas) remained the same upon changing between two different variants, it would be counted as one concept and the variants would be captured in the number of ideas, as is the case with the concept shown in Concept 4 in Table 6. This concept consists of 9 different ideas. Of these ideas only two would not be compatible, i.e. using jacks and cranes to lift the slab, which cannot be done simultaneously. Since this incompatibility involved less than half of the ideas, this concept was not split into two. If more than half of the ideas of a concept were incompatible, it was split into two or more different concepts. When applying the SIL method, several times the group discussed a new concept and then made a suggestion on how this concept could be combined with a previously discussed concept. In these cases, the new concept was incorporated in the previously suggested concept if, as in Concept 6 in Table 6, less than half of the ideas in the combination of concepts were incompatible. In Concept 6 in Table 6, the first suggestion was to protect the turnout by means of a shutter. Another participant then saw the possibility of using the shutter as a part of a fence which he had been thinking of. In connection with a discussion about how to use locomotives to remove snow, e.g. with a plough or brush, it was suggested that locomotives could increase the air current through the fence by putting a big plate in front of it.

The total number of non-redundant ideas was counted for each method. Because the concepts were of different types, it was challenging to define an “idea” as a unit for counting purposes. We used the procedure developed by Linsey et al. (2005, 2011) as a starting point. In this procedure, the basic definition of an idea is something that solves one or more functions in the functional basis (Linsey et al. 2011), with the functional basis defined as a standardized set of function-related terminology that allows repeatable and meaningful function representation (Hirtz et al. 2002). However, as a result of the ideation topics, several generated concepts were not meaningful to try to fit into a functional basis and other generated concepts required completely different functional representations. For the purpose of the present study, we instead defined ideas as the units which each concept could be systematically decomposed into, and which could be expressed as a key phrase consisting of a verb phrase containing a maximum of one verb, or a noun phrase. In this way, it was possible to capture ideas from concepts of different types.

### Table 6 continued

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Sketch and text</th>
<th>Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>SIL</td>
<td><img src="image" alt="Sketch" /></td>
<td><img src="image" alt="Ideas" /></td>
</tr>
</tbody>
</table>

- **No.**: 6
- **Method**: SIL
- **Sketch and text**: ![Sketch](image)
- **Ideas**: ![Ideas](image)
that were expressed by different means. The counting rules are shown in Table 7. Examples of how ideas were counted are shown in Tables 5 and 6. In Concept 4 in Table 6, the sketches illustrate what is mentioned in the text and do not add any additional key phrases. Concept 5 in Table 6 involves attaching a steel slab onto the concrete slab at an angle with bolts to soften the transition between track areas of different rigidity. The transition from ordinary rail to slab track was derived from the ideation topic and does not count as an idea. This way of defining ideas reflected the scope of a concept (how much work and thought had been put into it) and made it easy to find redundant ideas that had been used in several concepts. However, it does not cover all the information contained in sketches, e.g. the positioning of items in relation to each other. Judging the redundancy turned out to be a straightforward task, as ideas were either identical (e.g. the use of sleeper pads was a common suggestion when testing Method 635) or clearly different. A valid question is whether information contained in the audio recording but not written down or sketched by any participant should be included in the analysis of concepts and ideas, since, without the recording, this information would either have been lost or possibly stored only in the mind of the participants. We decided to use the recorded information to show the potential of the ideation methods. The extent to which concepts were actually documented during the ideation sessions is captured through the elaboration of the concept descriptions. A concept that was only described by spoken words would not have been captured through the inherent documentation of the methods.

### 3.4.2 Elaboration of concept description

The elaboration of a concept description reflects how much work the participant(s) have put into communicating the concept and how much discussion took place regarding the concept during the ideation session. A high degree of elaboration does not necessarily mean that the concept in question is of high quality or involves a high number of ideas. The elaboration of each concept description was defined as the number of sketched details, the number of written words and the number of spoken words used during verbal interaction. Sketched details were defined as follows: (a) physical objects, (b) intentional holes through a physical object, (c) indication of dimensions, (d) indication of quantity and (e) arrows indicating movement. The counting rules are shown in Table 8. Table 9 shows examples of how sketched details and written words were counted. Concept 7 in Table 9 shows a straightforward example containing two rails and six sleepers, counted as eight sketched details. Concept 8 in Table 9 shows another example with more and different sketched details. A hole created because it was surrounded by different other objects, like the hole in the area between the rail, two sleepers and the ballast body in Concept 8, was not included. In contrast, the holes in the sleepers in Concept 9 in Table 9 were counted. Each length or angle indicated is counted as a separate detail. An example of an indication of dimension is shown in Concept 9 in Table 9.

### 3.4.3 Viability and ability

The quality of an idea was described in terms of its viability and ability, which reflect its usefulness. The viability of an idea describes the possibility of practically implementing the solution with respect to technology, cost and safety. The current rules and restrictions were not considered, since these can be changed. The ability of an idea describes its capacity to solve the problems of the ideation topic in question. While the other parameters were analysed by the first author, the viability and ability were judged by the participants individually about 3 months after the last ideation session. The participants received textual instructions concerning what to consider when judging the viability and ability, and a description of each concept (copies of sketches and written words from the ideation sessions and a summary of spoken key phrases, when relevant). They were asked to make a mark on a line between 0 and 1, indicating how viable or able, respectively, they thought each concept was. Zero corresponded to “not at all viable/able” and one corresponded to
completely viable/able." The value was calculated using the same procedure as for the questionnaire. Some of the concepts were combinations of suggestions that could have been a concept on their own merits. The participants were asked to judge the different inherent concepts in a sample of such concepts. The participants could choose themselves how much time they put into the judgment process and the participant who did it most carefully used about 3 h. They did not use external information, but relied on their knowledge and prior experience. We checked if there were any obvious differences between how the academics and the railway professionals, respectively, judged the ideas. The parameters used to measure the ideation outcome are summarized in Table 10.

3.5 Behaviour of the participants

Observations of the ideation sessions, together with the transcribed audio-recordings of the ideation sessions, served as a basis for analysing the behaviour of the participants. To analyse the content of the ideation workshops, we used a variant of the coding system developed by Jackson and Poole (2003). Instead of the time used for the different activities, we counted the number of words used in the different activities. We used the main activities specified by Jackson and Poole (2003), i.e. idea statement, elaboration, criticism, direction and going off at a tangent. A description of each activity is to be found in Table 11. The distribution of the words spoken by the participants during the ideation session was analysed.

Material from the workshops was also analysed qualitatively to identify patterns in the behaviour of the group and its participants and how the diversity of the group manifested itself.

4 Results and discussion

4.1 The participants’ experience

Section 4.1 presents the results regarding the participants’ experience of the group composition and the ideation methods and relies on data from the interviews, questionnaires and observations.

4.1.1 Group composition

All the participants were convinced that the composition of the group, with members expressing different views on the ideation topics, was a winning concept. They all thought that the group possessed a large amount of competence and experience and that this competence and experience facilitated the finding of good solutions and helped them avoid continuing to work on solutions that had previously proven to be bad.

All the participants thought that the presentation and discussion of the ideation topic were important due to the diversity of the group. It was important to achieve a common understanding of the issue and survey the known causes before starting ideation; otherwise one had

<table>
<thead>
<tr>
<th>Table 8 Summary of the rules for counting the elaboration of concept description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The elaboration of a concept description is defined as the number of sketched details, written words, and spoken words related to a concept in its description</td>
</tr>
<tr>
<td>2. The number of written or spoken words is counted, without distinguishing between the different parts of speech, on the sheets containing solutions or the transcript records, respectively</td>
</tr>
<tr>
<td>(a) An abbreviation of one word is counted as one word</td>
</tr>
<tr>
<td>(b) An abbreviation of several words is counted as the number of words included in the abbreviation</td>
</tr>
<tr>
<td>(c) Sketched arrows indicating &quot;gives&quot;/&quot;results in&quot; are counted as one word</td>
</tr>
<tr>
<td>(d) A number is counted as one word</td>
</tr>
<tr>
<td>3. Sketched details are defined as</td>
</tr>
<tr>
<td>(a) Physical objects</td>
</tr>
<tr>
<td>(b) Intentional holes through a physical object</td>
</tr>
<tr>
<td>(c) Indication of dimension (only lengths and angles given in the present set of concepts)</td>
</tr>
<tr>
<td>(d) Indication of quantity</td>
</tr>
<tr>
<td>(e) Arrows indicating movement</td>
</tr>
<tr>
<td>4. A physical object is a discrete object that must be possible to distinguish in the sketch to be counted as a detail. The shape of the object is not considered</td>
</tr>
<tr>
<td>5. Holes are counted as objects if they are added intentionally and go through a physical object</td>
</tr>
<tr>
<td>6. Each length or angle indicated is counted as a separate indication of dimension</td>
</tr>
<tr>
<td>7. Each quantity indicated is counted as a separate indication of quantity. In a diagram, different values of dependent quantity are counted as separate indications of quantity. A changing quantity of constant slope is only counted once</td>
</tr>
<tr>
<td>No.</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
different starting points and understanding of the topic. In addition, the participants thought that it was beneficial to use the stated causes of problems as inspiration when generating ideas. They mentioned that when running out of ideas, they reviewed the causes to find new angles of approach.

All the participants thought that working in the group was really enjoyable, inspiring, stimulating and motivating and that its activities could improve turnouts. Generating ideas was, however, not the only objective of participation in the group. Using it as a means of interfacing with the other group members was also important. Naturally, the industrial participants wanted to promote their company and show the infrastructure manager why they should be their choice among competing companies. Both the industrial participants mentioned that if the competitors of their companies had participated, they would not have been able to speak freely and would have kept good ideas to themselves. This effect has been observed in another related project run by the last author, where several actors from the same actor segment participate.

### 4.1.2 Method 635

Method 635 was perceived as stressful by the majority of the participants. They thought it was tough to describe one’s suggestions and ideas in 5 min without talking, especially if one is not skilled in drawing. On the other hand, half of the participants stated that they experienced the stress factor and the regular input of ideas from the other participants as exciting and enjoyable. They considered the method to be very dynamic because new ideas were continuously generated by the input of ideas from the other participants. According to the participants, it was inspiring to think about other participants’ suggestions and, after exhausting one’s own ideas, the input from others helped in generating new ideas. The input from others made them forget ideas they had worked on during the previous round. The participants thought that the time restriction prevented the apprehension of their own ideas. Some participants described how they were pushed out of their normal thinking paths. The participants thought that ideas that might normally have been presented as one’s

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Measurement method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>The number of</td>
<td>Counted by the first author according to the rules in Table 7</td>
</tr>
<tr>
<td></td>
<td>Non-redundant ideas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-redundant concepts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sketched concepts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-sketched concepts</td>
<td></td>
</tr>
<tr>
<td>Elaboration</td>
<td>The number of</td>
<td>Counted by the first author for each concept according to the rules in Table 8</td>
</tr>
<tr>
<td></td>
<td>Details in sketches</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Written words</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verbal words used in the statement and elaboration of the concept</td>
<td></td>
</tr>
<tr>
<td>Viability</td>
<td>The possibility of practically implementing an idea with respect to technology, cost, and safety</td>
<td>Judged by the participants individually on a continuous scale between 0 and 1 (0 = not at all possible, 1 = completely possible)</td>
</tr>
<tr>
<td>Ability</td>
<td>The ability of the idea to solve problems dealt with in the ideation topic in question</td>
<td>Judged by the participants on a continuous scale between 0 and 1 (0 = not at all able, 1 = completely able)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 11 Activities used to code the ideation sessions, based on Jackson and Poole (2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
</tr>
<tr>
<td>Idea statement</td>
</tr>
<tr>
<td>Elaboration</td>
</tr>
<tr>
<td>Criticism</td>
</tr>
<tr>
<td>Direction</td>
</tr>
<tr>
<td>Going off at a tangent</td>
</tr>
</tbody>
</table>
own ideas entirely could end up in a suggestion from
someone else. As the sheets of paper with ideas were cir-
culated, the participants thought that the suggestions were
elaborated more and more, making it harder to add
something.

The participants thought that the high productivity might
mean that the quality of the ideas was worse and the
applicability lower compared to the other methods. An
element of scrutiny was desired by the participants, as
some suggestions were perceived as unrealistic.

4.1.3 Gallery method

The participants appreciated the fact that, when applying
the gallery method, they had more time available than
when they worked with Method 635. There was time for
describing and understanding the suggestions.

The method was described as being a comfort-
able method to use. One’s own suggestions could be
developed in peace and quiet, and then the benefits and
drawbacks could be figured out. The participants’ prese-
tations and the feedback from the other participants were
highly appreciated since they gave insights that the receiver
of the feedback had not thought of himself. The partici-
pants valued the fact that, afterwards, time was available
for revising the suggestions based on the feedback and by
using ideas from other participants. The participants
thought that the productivity was lower and that the con-
cepts would be better devised and have higher applicability
compared to those generated during Method 635.

Half of the participants, all of them railway profes-
sionals, selected the gallery method as their preferred
method, as shown in Table 12, which shows how each
participant ranked the three methods. The remaining rail-
way professional, who in the end chose the SIL method,
said that he too would have chosen the gallery method as
his preferred method if the choice had been based on what
method he felt was the nicest. Still, he finally chose the SIL
method as his number one method because he thought that
it pushed him out of his comfort zone and forced him to
think in new ways. One of the academics described the
gallery method as being rather ordinary, resembling a
normal meeting. Both academics chose Method 635 as
their preferred method.

4.1.4 SIL method

The participants thought that combining ideas was a way to
come up with more suggestions and that by combining
ideas they were forced to think differently. However, not
all the participants were impressed by the method. One
participant found it abstract and forced and could not see
the point of it. The majority of the participants were
reluctant to try to combine ideas that initially appeared
impossible to combine, and during the ideation session,
they had to be convinced by the first author that they should
actually continue trying, and then they discovered that they
were able to create new suggestions.

During the general discussions on how to combine ideas,
there was naturally an element of scrutiny present, since
the participants had to consider if and how ideas could be
combined and what needed to be added in order to make a
suggestion work. This element of scrutiny was appreciated
by all the participants, since they were anxious not to work
on suggestions that were completely unrealistic and eager
to find ways of improving a good suggestion.

4.1.5 Questionnaire

Figure 1 presents how the participants on average rated
different aspects of each method in the questionnaire. The
statements in Fig. 1 were also analysed separately for the
academics and railway professionals. It turned out that the
variation between participants was greater than that
between the two subgroups, and therefore, no conclusions
about possible differences between the two subgroups
could be established. The only tendency that could be
determined was that the railway professionals found the
gallery method more useful than the academics, who pre-
ferred Method 635 in this respect, similarly to what was
found based on the ranking of methods in Table 12.

The gallery method was popular and received the
highest average scores from the participants for several
parameters, such as “I think the method is useful” and “I
felt engaged during the workshop”, as shown in Fig. 1.
When Shah et al. (2001) compared C-sketch to the gallery
method and Method 635 (using only textual description),
the gallery method was also found to be popular and scored
highest with regard to the participants’ views on the cre-
ative outcome and the promotion of creative cognitive
processes (Shah et al. 2001). The gallery method allowed
the participants to talk and discuss, which they expressed
was very exciting and important to them. However, the SIL
method allowed even more time for group work, but was
still not as popular. Most of the participants actually

Table 12 Ranking of the methods by participant A–F, with “1”
given to the method liked the least and “3” to that liked best

<table>
<thead>
<tr>
<th></th>
<th>Railway professionals</th>
<th>Academics</th>
<th>Sum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Gallery</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>SIL</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
preferred the silent Method 635 to the SIL method when ranking the methods (see Table 12). This implies that working in a verbally interactive group is not sufficient to create an enjoyable session. Possible reasons for the SIL method being the least popular method may be that the participants felt uncomfortable trying to combine ideas which they did not consider to fit together and that they perceived the method as less structured than Method 635 and the gallery method. It was expected that the SIL method would challenge the participants, but not to such a degree that they would become reluctant to use it.

It is interesting to note that all the railway professionals preferred the gallery method (except for one, who said that he would have chosen it if he had only based his choice on what method felt most comfortable), whereas the academics preferred Method 635. This might reflect the way in which the participants are used to working in their own organization. An interesting finding was that the professionals working at STA consistently awarded the lowest marks or marks that were among the lower marks for “Many ideas were new to me” compared to the other participants. An explanation could be that, as employees of the infrastructure manager, they are exposed to more ideas from others compared to the other members of the team.

4.2 Ideation outcomes

The concept and idea generation rates are presented in Fig. 2. Figure 3 shows the number of ideas for each combination of viability and ability for each method. Since the methods had different total times for ideation, the histograms are normalized to 60 min. The viability/ability of each idea is the average of the participants’ ratings.

In terms of concept and idea generation rates, Method 635 outperformed the other methods. Since verbal interaction has been found to contribute to process losses (Mullen et al. 1991), this result is expected as Method 635 provided more time for individual ideation than the gallery and SIL methods. However, other factors might also have played a role in this case. The results might be partly explained by expectations. The participants knew that the “3” in Method 635 stands for coming up with three ideas in every round. Although they were informed that fewer or more ideas would also be acceptable, they probably felt that they were expected to generate several ideas before giving the sheet of paper to the next person. During the other workshops, no such specific goal was specified. Specifying goals has been found to increase performance during group work (Wegge and Haslam 2005) and even eliminate the productivity gap between nominal and interactive groups (Paulus et al. 1993). The stress factor probably also exerted an influence, since the time constraint prevented the participants from evaluating their own ideas before writing them down. In addition, previous research has shown that individuals under time pressure work at a faster rate and that time limits are inversely related to the amount of task focus shown by groups (Karau and Kelly 1992).

The gallery method invited the participants to sketch their suggestions, and this method had the highest fraction of sketched concepts (see Fig. 2), as expected. Method 635, however, had almost as high a number of concepts including a sketch, generated over a shorter period of time. Table 13 shows the distribution of the number of sketched details, written words and verbal words describing each concept. Not all the concepts were described by both sketches and written words, and therefore, the average
number of sketched details and words has been calculated with respect to both the total number of concepts and the number of concepts that had at least one sketched detail or one written word, respectively. The methods exhibited different characteristics with regard to the distribution of sketched details, written words and verbal words in the concept descriptions. This is interesting, because only sketches and written words are saved for post-ideation processing if no additional documentation is performed, and sketches typically contain information that is difficult

| Table 13 Distribution of the sketched details and written and verbal words in the descriptions of concepts |
|-------------------------------------------------|-----------------|-----------------|-----------------|
|                                                   | 635 Gallery SIL | 635 Gallery SIL |
| Number of sketched details                       |                 |                 |
| Average for all the concepts                     | 5.9             | 18.2            | 3.2             |
| Average for concepts with at least one sketched detail | 12.8           | 20.4            | 9.1             |
| Number of written words                           |                 |                 |
| Average for all the concepts                      | 15.4            | 11.2            | 4.9             |
| Average for concepts with at least one written word | 15.4           | 12.0            | 6.2             |
| Number of verbal words                            |                 |                 |
| Average for all the concepts                      | 0               | 105             | 274             |

Fig. 2 Concept and idea generation rates

Fig. 3 Histograms, normalized to 60 min, of concepts as a function of their average ability and average viability as judged by the group members
to capture by textual description. Not surprisingly, the number of spoken words per concept increased and the number of sketched and written details decreased as the time for verbal group ideation increased (see Tables 3 and 13).

The gallery method had a considerably lower concept and idea generation rate than Method 635. This is in accordance with the participants’ perception that the number of suggestions was smaller with the gallery method than with Method 635, although this was not reflected in the questionnaire. The participants suggested that the concepts would be better devised and that the quality, as well as the applicability, might be higher with the gallery method. Table 13 shows that the number of sketched details in the concepts from the gallery method was higher compared to the corresponding number for the other methods, but on average the gallery method did not result in more ideas per concept than Method 635, despite the fact that the average available time for working on each concept was considerably higher with the gallery method (see Table 14). This indicates, although some information in the sketches might have been lost in the idea counting, that the participants sketched their concepts more carefully when working with the gallery method, and the resulting sketches might have been easier to comprehend, but the participants did not use the time to add ideas as to how to solve problems connected with different features of their concepts or to create new concepts. For example, adding more sleepers and holes, as in Concept 9 in Table 9, adds more sketched details that can clarify the concept to others, but does not add any new ideas as to how to address the ideation topic. According to Table 15, neither the average viability nor the average ability differs significantly between the methods. However, Fig. 3 shows that the average ratings of the ability and viability of the ideas generated when applying the gallery method are less scattered than the corresponding ratings of the ideas generated with the other methods, especially Method 635, which might explain the participants’ perception that concepts generated using the gallery method would be of higher quality.

As shown in Fig. 1, Method 635 scored lower than the other methods with respect to the participants’ perception of the possibility of using the ideas presented in practice. One explanation might be that the very high number of generated ideas, including more low-rated ideas, compared to the other methods. These low-rated ideas might have dominated the participants’ perception of the method, although it actually generated a higher number of ideas with high ability than the other methods, as shown in Fig. 3. In the study by Shah et al. (2001), C-sketch and the gallery method had higher mean quality scores than Method 635 (using only textual description). However, since they did not mention how many ideas from each method had a certain quality (and variety and novelty) score (only average values were given), it is still possible that Method 635 had a similar or higher absolute number of ideas with a high-quality score, if many more ideas were generated. This would be interesting to know, as this was found to be the case in the present study.

Linsey et al. (2011) found, contrary to this study, that when participants were allowed to use both sketches and words to describe solutions, gallery viewing generated more product solutions (which are analogous to concepts in the present study) than rotational viewing. Rotational viewing, however, produced an overall greater number of ideas per function than gallery viewing. A major difference between the execution of the gallery method in the present study and how the method was executed in Linsey et al.’s study (2011) was that no presentation and discussion step was included in the gallery viewing condition in the latter study. Therefore, one possible explanation of the different findings may be that the time used for presentation and discussion in the present study resulted in significantly fewer suggestions being generated in this step than in the ideation steps and that this was not compensated for by an increased number of ideas being generated in the following short ideation step. The role of presentation and discussion steps should be investigated in future research.

| Table 14 Average number of ideas and average available time for each concept |
|-----------------|----------|----------|
|                 | 635      | Gallery  | SIL      |
| Average number of ideas per concept | 3.4      | 3.5      | 5.2      |
| Average time available for each concept (min) | 3.9      | 9.3      | 15.7     |

| Table 15 Viability with respect to technology, cost and safety and the ability of the idea to solve the issue |
|-----------------|----------|----------|----------|
|                  | 635      | Gallery  | SIL      |
| Viability        |          |          |          |
| Average          | 0.62     | 0.56     | 0.54     |
| Average range    | 0.78     | 0.61     | 0.67     |
| Average standard deviation | 0.29     | 0.22     | 0.24     |
| Ability          |          |          |          |
| Average          | 0.54     | 0.66     | 0.64     |
| Average range    | 0.59     | 0.51     | 0.57     |
| Average standard deviation | 0.21     | 0.18     | 0.21     |
In conclusion, the participants had on average a realistic view of the concept quantity, but misjudged the concept quality after the ideation sessions. This is intriguing, since they themselves judged the ability and viability of the concepts, which means that, when they themselves had to think specifically about the usefulness of each concept, they reached conclusions that were different from what they spontaneously felt about the collection of concepts after the ideation sessions.

The participants were also asked to judge a sample of concepts at different stages of the ideation session, to find out if they thought that the concepts had been improved by the elaboration of others (in the case of Method 635) or by the performed combinations (with the SIL method). It turned out that these samples included concepts that had improved and concepts that scored approximately the same or worse after receiving additions or being combined with another concept or idea. An example of a concept that became worse after it had received an addition from another participant is Concept 4 in Table 6. In this example, the initial suggestion was to have a slab cast in one piece under the turnout’s most vulnerable parts and to post-adjust it by means of injection casting while lifting the slab to the right position with a jack or a crane. This suggestion received an average viability score of 0.74 and an average ability score of 0.60. When one participant added the idea of a having a separate slab for the point driver, the average viability score did not change considerably but the average ability decreased to 0.40. As a third participant added the idea of under sleeper pads to attain an exact rigidity, the concept received an average viability score of 0.92 and average ability score of 0.77, a considerable increase. However, as a fourth participant suggested making the under sleeper pads softer than those used today, the scores decreased to 0.75 and 0.45, respectively. Another example was the suggestion to let an engine remove snow by means of a brush or other mechanical special devices, receiving an average viability score of 0.63 and average ability score of 0.67. However, as another participant added the idea to adjust the turnout to the engine’s special devices the average ability score stayed the same but the ability score decreased to 0.33. The overall score given to these concepts concerned the entire concept including the “bad” modification. It is not within the scope of this study to investigate this in detail, but rather to conclude that modifications of concepts can deteriorate the original concept. Linsey et al. (2011) noted that the quality of a product solution frequently changes as team members add their ideas. They found that embellished product solutions tend to be higher quality product solutions and give examples of how an overall product solution can drastically improve as individuals add ideas (Linsey et al. 2011). They did not, however, investigate how product solutions changed for the worse as they received additions and how common this was. Shah et al. (2001) found that when the gallery method was used, the first ideation step provided suggestions with high scores for variety, but low to medium scores for quality. After the second ideation step, the scores for variety decreased, but the quality of the ideas improved. Shah et al. (2001) suggested that the participants picked up ideas from others that improved their own concepts, but also made them less varied. It appears that there is a knowledge gap concerning how often individual concepts deteriorate or improve as they are modified. This is important to know when the group moves on to concept selection. If a concept receives an overall low score, there might still be high-quality features embedded in that concept that can be useful when developing the end product.

From Table 15 it can be observed that the average range (the average difference between the minimum and the maximum score) and the average standard deviation of the viability and ability are quite high for all the methods. This reflects the fact that the participants in general answered differently as to how viable and able the ideas were. In some cases, the range for a specific idea was 1, meaning that at least two participants totally disagreed on its viability or ability. No difference could be found between the railway professionals’ and the academics’ judgment of ideas. This is in contrast to previous studies using raters, where the inter-rater agreement concerning quality-related parameters is typically quite high. This is believed to be a result of the domain-specific topics and the ability of the highly experienced expert participants to generate advanced solutions. These factors made it much more difficult in the present study to judge the quality of ideas compared with previous studies without domain-specific topics and with participants who did not possess expert knowledge or experience of those topics. It is an important finding that, although all the participants have substantial knowledge of and experience from the railway sector in general and turnouts in particular, at this stage, they had different opinions on what ideas were useful. An additional reason for the participants’ different opinions might be their different perspectives on the ideas and what they actually know about the possibility of implementing them. It is interesting to note that Chulvi et al. (2012) used expert raters with at least 8 years of professional experience in the domain-specific area concerned to rate solutions generated by multidisciplinary teams of PhD students or professional designers working on design problems; they found, similarly to this study, that the experts’ responses had a significant dispersion, especially concerning the degree of usefulness. Chulvi et al. (2012) suggested that this might “indicate that in the absence of available data, even experts find it difficult to assess and compare the potential usefulness of a product.” This finding implies that it is
inadvisable to discard ideas at this stage, especially based on the opinion of only one person, since all the possibilities of the solution might not be understood. On the other hand, some participants might not have understood certain drawbacks of certain ideas, and these drawbacks might change their judgment. In conclusion, this underlines the importance of using a concept selection method to guide the choice of ideas to be developed further.

Two of the participants expressed concern that not many ideas were novel, although all the methods were rated fairly highly concerning the participants' apprehension of the number of ideas that were new to them (see Fig. 1). An idea may be new to the individual who conceived it, to humanity or to some subset of humanity greater than one (Nickerson 1999), e.g. the ideating group. Some of the suggested ideas had been thought of before the ideation session by the participants or had been noticed in another context. Such ideas had been encouraged by the fact that the participants had got to know the ideation topic beforehand. The participants had already been working with some of the ideas that were presented. The proportion of these categories of ideas is probably very different depending on whether the ideation topic is a new or old issue. Old issues have already been the subject of much deliberation, and ideas have been conceived as to how one can solve the problems which they consist of. However, if an old issue has still not been solved, an ideation session is a good opportunity to share ideas and discover a new angle on one's own ideas. If this is the case, it might be especially appropriate to use methods that push the participants outside of their normal thinking paths and involve people with different perspectives in order to find new ideas. The remarks stating that not many new ideas had been presented might also highlight the fact that the “old” ideas had not yet been discussed and scrutinized thoroughly and therefore would reappear until they had finally been discarded or accepted for further development. We suggest that this has been a neglected topic in the research on ideation methods. Sarkar and Chakrabarti (2009, 2014) offer a framework for exploring this topic through their definition of twelve different types of searches for ideas. According to their definition, a new search in the solution phase is activated when a designer comes up with an idea that was not previously known to the designer, i.e. was not derived from any knowledge base available or known to the designer. In an experiment letting novice and experienced designers work on two different problems individually, without the use of any design methods, it was found that only 15 out of 814 searches in the solution phase were categorized as new (Sarkar and Chakrabarti 2014). In another experiment (Sarkar and Chakrabarti 2009), groups of three members worked on two different problems using brainstorming, functional analysis, ideal design or an innovation situation questionnaire. In this experiment, it was found that with brainstorming, about 10 % of the total number of search solutions were new solutions, whereas, when the other methods were used, on average 1 % or less of the total number of solutions were new. According to these findings, the generation of new ideas is very rare, but can be enhanced by an appropriate ideation method. The distribution of known and novel ideas has typically not been taken into account when evaluating ideation outcomes, possibly due to the complexity of the issue. There is quite a difference between listing already known ideas and aiming to encourage creativity in such a way that novel and potentially radical ideas arise.

4.3 Behaviour of participants

Table 16 shows the distribution of spoken words between the participants during the verbally interactive group time of the gallery and SIL workshops. The speaking time was unevenly distributed, since during both sessions almost half of the spoken words were uttered by two participants. According to Gibson (2010), researchers within the field of small-group dynamics assume that some individuals speak infrequently on the grounds that others are thought to have more to contribute to a task that everyone wants to see performed well. Other reasons might be that participants might choose not to share their ideas for different reasons, or they might not get anything said because of group communication structures and personal low “verbal latency” (i.e. the ability to “jump in” as an opportunity appears in a discussion [see e.g. Burke (1974)]. During interviews, the participants mentioned the problem of low verbal latency as the reason for the unequal distribution of words. One negative consequence of such behaviour is that it can hinder access to the full range of skills and resources, and the content of a meeting may be governed by those who speak the most. Unequal participation might also have consequences for post-encounter behaviour, e.g. compliance with decisions (Gibson 2010). When the members of a team possess diverse sets of knowledge, less information is shared within it, and therefore, it is particularly important to implement procedures that structure discussion, to

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gallery</th>
<th>SIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>E</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>F</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Leader</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>
enhance team sharing and consequently performance (Mesmer-Magnus and DeChurch 2009). We therefore believe that an unequal distribution of spoken words is a drawback and should be mitigated to release the full potential of the different knowledge and experience represented in a cross-functional group.

The distribution of spoken words between the activities during the verbally interactive time of the gallery method and the SIL method is shown in Table 17. It turns out that most of the spoken words were dedicated to the elaboration of ideas, whereas only a minor part of them was dedicated to stating an idea. In the gallery method, it was in the verbally interactive stage that the participants, according to instructions, presented their suggestions from the individual ideation and gave feedback concerning other participants’ suggestions, and therefore, it is not surprising that not very many of the words were used for idea statement. However, during the SIL method, the suggestions were generated during the verbal interaction time and still only 5% of the words spoken were used to state an idea. Jackson and Poole (2003) suggested that group members are not creating a list of ideas as much as they are constructing shared meaning about each idea and recording it publicly. As shown in Table 17, criticism was rare. Accordingly, the participants reported that they thought that the group was not criticizing ideas and that they could express whatever came into their mind.

The qualitative analysis of the transcript records showed that the elaboration was performed partly to make sure that all the participants understood the idea, which is a finding that might be typical of inter-organizational groups dealing with technical and other complex problems and their solutions. Through relevant discussions, the participants were directed towards thinking of variations of their ideas and in this way concepts evolved. Four different ways in which the diversity of the group manifested itself were found in the qualitative analysis of transcript records: (a) objections to other participants’ ideas, (b) associations inspired by other participants’ ideas, (c) former experience of similar solutions and (d) immediate feedback to questions. Table 18 shows excerpts from the transcript records where participants made objections to other participants’ suggestions. These objections in several cases led the group to explore how objections could be addressed by altering suggestions.

Elaborating ideas can lead to new ideas that would otherwise not have been obtained, in accordance with the research by Seidel and Fixson (2013) and Pelled et al. (1999), suggesting that debating ideas results in a more thorough exploration of the solution space. The elaboration of ideas may also save the group time at a later stage of the project. Table 19 shows examples of how participants were inspired by other participants’ suggestions and either elaborate them or create other suggestions.

The diverse experience of the participants was also evident in comments where participants talked about previously proposed similar solutions, as exemplified in Table 20, which the other participants did not know of. This experience was in several cases exploited for the benefit of the current group interaction, e.g. by adapting solutions or by discarding suggestions as they had already been proven not to work.

The diverse knowledge and different backgrounds of the group members made it possible for a participant to obtain a quick answer to anything he was wondering about in relation to a suggestion; examples of this are shown in Table 21. It is not a straightforward task to measure quantitatively the degree to which the different backgrounds of the participants helped the evolution of ideas, but it is obvious that the more knowledge and experience there is gathered in the group, the higher is the probability that someone will know something that can help improve another participant’s suggestion.

The participants of the group represented different organizations, and it could be observed, in agreement with the participants’ own views, that they sometimes took the opportunity—consciously or unconsciously—to promote things that their organization provided or considered developing. This concerned not only the industrial actors, but also the academics, e.g. in talking about suggestions for work that academia could perform or prototypes that could be built and would benefit their projects. These attempts made by participants to promote their own organizations did not dominate the ideation sessions and often there was the potential of a win–win situation, but this is something that must be taken into consideration when ideating in real-life settings.

Another way in which the different backgrounds of the participants played a role was in the group-analysis of the topics and finding out the causes of problems. It is beyond the scope of the present paper to study this process in detail, but it was observed that the participants contributed different perspectives on the topics, which resulted in quite a comprehensive group-analysis of the topics in a short time. The identified causes were used as inspiration during
the ideation sessions and in this way resemble the way the inspirational poster presented to the brainstorming but not the Method 635 group in the study by Rexfelt et al. (2011) was used. Rexfelt et al. (2011) reported that the Method 635 group in their study struggled coming up with suitable ideas, and it is interesting to note that we did not make similar observations during our test of Method 635. An interesting area for future research is how the scope of the group-analysis of topics is related to the outcome of the ideation.

The version of Method 635 used in the present study allowed the participants to ask the neighbour sending them his sheet of paper to explain anything they did not understand. This is not allowed in C-sketch, for example, in order to encourage misunderstandings that can lead to novel ideas (Shah et al., 2001). However, in the present study, the participants never asked their neighbour for an explanation when performing Method 635.

One of the reasons for choosing the three methods tested in the present paper was that they were expected to be rather simple to perform. In this respect, it turned out that Method 635 and the gallery method were easy for the participants to understand, and they found the instructions for these methods easy to follow. In contrast, the participants had problems grasping the SIL method. They did not understand the instructions, had many questions about what to do, and seemed quite confused, as is exemplified in the excerpts from the transcript records in Table 22. Moreover, the original instructions by VanGundy (1984) about integrating all the generated ideas into the same solution had to be abandoned, and instead the participants tried to combine two suggestions at a time, and when a new suggestion was presented, the participants reviewed previous suggestions to check if they could find something appropriate to combine it with. They also reviewed the remaining suggestions that they had come up with individually to see if something could fit, rather than trying to combine any two suggestions. In the end, this worked quite well, but one cannot maintain that the method was easy to perform at the first attempt.

### 5 Limitations

Although our study has revealed interesting findings and insights concerning the use of ideation methods in real-life settings, and although these findings and insights can be of value to academics and practitioners, the case study method has inherent limitations when it comes to the generalization of results. Our empirical findings add to the cumulative body of knowledge on ideation methods and concept generation in different types of groups, and future research should further explore these limitations.

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SIL</td>
<td>F: […] those solutions, what is attractive about them is that they don’t cost anything other than a little programming work. D: […] the latter yes, the first is about safety. A: Then you have to remake all the point machines.</td>
</tr>
<tr>
<td>2</td>
<td>SIL</td>
<td>B: And I imagine a few wings and things within the turnout precisely to make that interplay with the vehicle […] and try to arrange for there not to be any lee-sides inside the turnout, because the trains actually come from both directions, so that you in some way can lead the snow out through wind power and let it carry the snow out. D: Now, there is a difference between single track and double track, if you think about double track, then you have trains coming from the same direction, is that really an advantage? F: It would be good if it could be constructed in such a way that it worked in all directions, so that it got cheaper in some way maybe.</td>
</tr>
<tr>
<td>3</td>
<td>SIL</td>
<td>F: Yes. Or that (the turnout) is high enough like that, then you can manage to plough properly with a deeper plough or something that is put down or, or something. […] D: I don’t know if I should be critical but […] what you should think about is that we pass both a crossing and a check rail, so that the height of your fixed installation on the train must be about a decimetre up, it is only those trains where you can lower and raise the plough with a driver that you dare to drive through a turnout at full speed. If you are to trust some automatic function that raises it before the check rail, then I think we will smash things to pieces. F: Yes, but it’s a bit the same with that shutter there, you have to be able to rely on it opening. D: Yes, exactly. I believe more in compressed air, in thinking along such lines rather than [a] pure mechanical [solution] F: Or, [a] mechanical [solution], then one could have, so to speak, things that are elastic, like these rubber flaps that we talked about, things that endure contact with the turnout.</td>
</tr>
</tbody>
</table>
research can clarify to what extent the findings can be
generalized to other instances of cross-functional inter-organizational groups and determine whether similar findings can be made in other real-life settings. The findings and conclusions of the paper are based on the results for this particular group. Using another group might have resulted in different findings due to different group characteristics exerting different influences. Possible influencing factors in this connection are the personalities of the group members, their experience, backgrounds, gender, age and other personal characteristics. Further, the group acted in the railway sector and studying a cross-functional inter-organizational group from another sector could result in other findings.

Because only a single group was considered, we were forced to change the ideation topic between ideation...
sessions. Although the different topics were chosen with care so that they would be equivalent and although they were found to lead to similar engagement among the participants (see Fig. 1), it cannot be excluded that their impact influenced the ideation outcomes differently. The number of non-redundant ideas and concepts was carefully counted according to a specified procedure to make the count as fair as possible between the methods. However, subjectivity will to some extent influence the counting, and in the present case the different nature of the concepts made the identification of single ideas challenging. We therefore believe that there is a need to develop more comprehensive procedures for counting these types of ideas.

The length of the ideation sessions differed due to technical difficulties. Therefore, the number of concepts and ideas were given as rates in Fig. 2, and the histograms

<table>
<thead>
<tr>
<th>No.</th>
<th>Method</th>
<th>Excerpt</th>
</tr>
</thead>
</table>
| 1   | Gallery | E: For if you have a joint between each part, then you kind of get a bump over it  
A: Yes, exactly  
E: A hinge there  
C: In a sense there are beams like that in the City Tunnel. The problem is that they are probably not long enough. They are too short. So you only move that really stiff transition […] |
| 2   | SIL    | E: […] I am a bit fascinated about this simple idea of a shed. That is to say a cold space […] with a, now I’m not taking the overhead line into account. But in principle, I mean, that you have such a cold space here and here you come in and on the outside then the turnout has done its job. Yes. And so at least there’s no snow coming in. No heating. No maintenance costs. It opens up for a few, that we can add a few air fans and whatever. That can’t be so terribly expensive  
D: [A mine operator] built their own terminal and to have as little snow as possible in the ore they unload, they have built a giant tent with, it’s only made of tarpaulins. I mean, a steel frame and tarpaulins, not like we do it with wooden walls and a steel roof and so on. So there has to be a cheaper solution to building one of those and then you can take it away during the summer, you can take down the tarpaulins again  
E: Exactly, and then the overhead line can run under it […]  
D: They have built it over the overhead line […] |
| 3   | SIL    | C: And I have seen a railway built through a desert, [and there] you have the same problem, of course, but with sand  
F: [And …] the next thing I thought about in connection with the train was, it was […] this plough which could be adjusted and then create a gust of air, but, that it also has such a shape, now it is positioned, I don’t know how high above the rail the plough is positioned, does anyone know? […]  
A: The fixed ploughs on the engines?  
F: Yes, a bit up?  
A: Yes  
D: Almost a decimetre on the engines, but then the engine [designed for snow removal is closer to the] rail, it’s only a few centimetres, there’s a piece cut out for the rail heads and suchlike, but it’s special for snow clearance, of course  
E: How long a section does one say needs to be covered, I mean, how long is it?  
D: The length of the installation is 10–15 m in these longer turnouts  
F: [There should be] some cleats further back too |
in Fig. 3 were normalized to 60 min. However, this could be misleading, since it has been shown that the number of ideas can decline over time during brainstorming (Paulus and Dzindolet 1993), and it is conceivable that this could also be the case for other ideation methods. This would imply an overestimation of the concept and idea generation rate of Method 635. However, taking the number of concepts generated by Method 635 in 30 min and dividing that number by 60 min (the length of the SIL session), one obtains a concept generation rate of 0.77, which is still higher than that of both the SIL (0.38) and the gallery method (0.48 over a time of 60 min). Similarly, the idea generation rate of Method 635 was also higher when calculated for a 45 min period.

In the present study, the participants were asked to judge the viability and ability of the concepts rather than employing external raters. This introduces the risk of a bias, as the participants might rate their own concepts highly to gain personal benefits. Since the gallery method is the method in which it is easiest to find out the originator of each idea, all the concepts generated using this method were checked to see if there appeared to be any bias in the ratings. Bias was considered to have occurred if the originator had rated his idea considerably higher than all the other raters. Concerning the ability of concepts, no such cases were found. With regard to the viability of concepts, there were four concepts whose originator had given them a rating which was equal to or more than 0.3 units higher than the ratings awarded by the other raters. However, when comparing the ratings for the ability of the same concept, a similar pattern was not detected. We therefore think that it is more likely that the high ratings given for viability by the originators in these cases were awarded because the originators actually thought them to be good ideas. However, if any bias exists which we are unaware of, the use of six raters implies that, even if one of them rates his own idea very highly, the idea will still receive a low average score if the other five raters think that it is a bad idea.

### Table 22

Excerpts from the transcript records showing how participants were confused about using or reluctant to use the SIL method. L denotes the first author.

<table>
<thead>
<tr>
<th>No.</th>
<th>Excerpt</th>
</tr>
</thead>
</table>
| 1   | L: Then I would like you first of all to present an idea<br>  E: Okay<br>  L: And then F presents an idea. Then together [you] try to combine these two ideas. You are welcome to use the whiteboard<br>  E: Shall I start?<br>  F: Aha, okay, so then one should see if it can be combined with…<br>  L: [E’s] idea<br>  E: How long shall I talk about my idea? Or my two [ideas]<br>  L: It’s mostly, it’s more on a conceptual level<br>  E: Shall I present one or two, I had two<br>  L: You should start with one […]<br>  E: [Presents his first idea.]<br>  L: Okay, that can be added to<br>  E: Then I only had one [more] idea<br>  L: Okay, there will be another round<br>  E: Aha, only one idea at a time, okay<br>  2  | L: Do we see any contradiction? Or is it possible to combine those two?<br>  E: And then [F and I] are quiet?<br>  L: No, everyone can discuss. Everyone’s free to speak<br>  E: But they’re like two separate ideas in their own right<br>  A: They’re two different things<br>  L: Yes, so they can actually be combined, but the point is that they are to solve the same thing<br>  E: But maybe they can’t be put together<br>  3  | F: May I just ask, this method, is it a question of two [participants] presenting their ideas and then the others giving feedback, or […]?<br>  A: It’s supposed to be based on […]<br>  L: It’s a question of jointly, as it were, it’s not so much a question of feedback and criticism, but rather of trying jointly to combine [the ideas] to something that can be useful
idea, and in this way the effect of any existing bias will be low.

6 Future research

The findings in the present study have indicated several areas that we believe would benefit from future research. Firstly, how the group-analysis of topics, the presence of verbal interaction steps and the distribution of spoken words between participants influence the ideation outcome is important to understand in order to optimize ideation methods. Secondly, by understanding how the quality of concepts changes as they receive additions or are combined, one might avoid missing valuable contributions “hidden” in a low-quality concept and facilitate the design of concept selection methods. Thirdly, studying how the distribution of old and new ideas differs between different methods can deepen one’s understanding of how to increase the probability of truly new ideas being created during ideation.

We will use the findings from the present study to develop ideation methods adapted to cross-functional inter-organizational groups and test and validate these methods in real-life settings. A great number of concepts and ideas are generated when using ideation methods, and it is necessary to screen out the most promising ideas to make further use of them. A decision support method will be developed to facilitate this process.

7 Conclusions

It has been found that the participants in the present study enjoyed working in a cross-functional inter-organizational group. They found the different and complementary knowledge about the topics that the different participants possessed with regard to the ideation topics very valuable both during the group-analysis of the topic that took place before ideation and during the generation of ideas on how to address the ideation topic. The participants found the group-analysis of the topic to be very important, since they had different perspective on the issue and needed to achieve a common understanding of it before starting ideating. The participants contributed different perspectives on the topics. During ideation, the participants reported that they used the causes of the issues identified during the group-analysis as inspiration. During the meetings, the participants in the group sometimes took the opportunity to promote the organizations which they represent.

Applying the ideation methods to the ideation topics facilitated the sharing of information between the participants. The speaking distribution was, however, unevenly distributed between the participants during the interactive ideation. Most of the spoken words were dedicated to the elaboration of ideas, and this elaboration was performed partly to make sure that all the participants understood the idea. Criticism was rare. During these verbally interactive steps, four different ways were found in which the diversity of the group manifested itself, and which could lead to improved ideas and concepts, as well as additional ideas and concepts: (a) objections to other participants’ ideas, (b) associations inspired by other participants’ ideas, (c) former experience of similar solutions and (d) immediate feedback to questions. Concepts improved, scored approximately the same or deteriorated after receiving additions or being combined with another concept or idea. Hence, if a concept receives an overall low score, there might still be high-quality features embedded in that concept that can be useful when developing the end product.

The academics preferred using Method 635, while the gallery method on average was most popular among the railway professionals in the group. Overall, the SIL method was least popular among the participants. Method 635 and the gallery method were easy for the participants to understand and follow. In contrast, the participants struggled to grasp the SIL method and were reluctant towards how it was supposed to be performed. In the end, the original instructions of the SIL method had to be abandoned for a variant that worked for the participants.

The participants on average had a realistic view of the concept and idea quantities, as they thought that Method 635 had generated the highest number of suggestions. Method 635 was indeed found to have a considerable higher idea and concept generation rate than the gallery and SIL methods. The participants, however, misjudged the concept quality after the ideation sessions, as they thought that the gallery method had generated better devised concepts of higher quality and applicability than Method 635. The quality of the generated ideas, as rated by the participants in terms of their average ability and average viability, did, however, not vary significantly between the methods. The participants in general answered differently as to how viable and able the ideas were. In other words, the participants had different opinions on what ideas were useful in this stage.

Acknowledgments The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged. The OptiKrea partners participating in the project, the Swedish Transport Administration, Vossloh Nordic Switch Systems and Infranord are gratefully acknowledged for their contributions to the project. We would like to thank the editor-in-chief and three anonymous reviewers for their comments on earlier versions of this paper which have resulted in a significant improvement.
Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Paper III
Developing an ideation method to be used in cross-functional inter-organizational teams by means of action design research

Anna Malou Petersson* and Jan Lundberg
Division of Operation, Maintenance and Acoustics, Department of Civil, Environmental and Natural Resources Engineering, Luleå University of Technology, SE-971 87 Luleå, Sweden

*Corresponding author: E-mail: anna.malou.petersson@ltu.se; tel.: +46(0)7022419179; fax: +46(0)920-492818.

ABSTRACT

Using ideation methods in an optimal way has a great potential to increase the number of ideas which a team can contribute during the conceptual phase of product development in industry. Previous studies on ideation methods have been mainly statistical studies in a laboratory setting. In the present study, however, the aim was to develop tailor-made ideation methods in a specific context, among actors on a deregulated railway market, through close interaction between researchers and engineers in a real-world context. Considering previous relevant research on ideation methods, associative memory models and the findings from tests of established ideation methods performed in the same group, a number of preliminary design principles were formulated and implemented in an ideation method that combines individual phases of rotational and gallery viewing with phases of verbal group interaction. The method was tested and refined in a cross-functional inter-organizational group comprising participants from different actors in the railway sector. Besides its provision of qualitative and quantitative test results, the present study has considered the opinions of the participants in detail, which can give important insights into the factors determining whether or not such a method will be implemented in industry. The participants found the method to be more useful and to generate more ideas that could be used in practice compared to the established ideation methods which they had tried. The learning derived from the specific case was formalized into a number of design principles for ideation methods to be used in cross-functional inter-organizational groups.

Keywords: product development, concept generation, cross-functional groups, inter-organizational groups, ideation methods, action design research

1 INTRODUCTION

The creation of new ideas is of the utmost importance to engineers during the conceptual phase of product development. A great number of ideation methods, proposed by practitioners as well as researchers, are available to assist the engineer in the process of inventing and creating new products. However, little empirical data exist to assist the engineer in using these methods or selecting the optimal method. The most famous method, brainstorming (Osborn 1957), has been extensively studied in the social psychology literature, but, although much can be learnt from these studies, it is not possible to transfer the results directly to the area of engineering design for at least three different reasons. Firstly, these studies let the participants use words to express ideas, whereas designers use sketches to communicate. Secondly, the typical psychology study involves topics that do not require domain-specific knowledge to solve problems, whereas engineers engaged in product development deal with problems that do require such knowledge. Thirdly, whereas these studies have typically involved students, engineers engaged in product development typically work in cross-functional teams involving participants with complementary skills and knowledge. These differences imply that ideation methods have to be studied and developed in the context of engineering design to extract and provide guidance as to how engineers can make the best use of these methods to generate concepts. Linsey et al. (2011) examined how two key factors influenced the outcome when mechanical engineering students worked with a design problem. These key factors were the way in which a group’s ideas are displayed and the form of communication between group members. They found that over a 40 minute ideation session, 50% more ideas were generated with the most favourable combination of key factors than with the least favourable combination. Linsey and Becker (2010) made a complementary study letting nominal groups (groups of non-interactive individuals who ideate
on their own and whose non-redundant ideas are pooled and compared to the outcome of an interactive group) ideate on the same problem given the same amount of time. They found that real teams in several of the ideation conditions generated a larger number of ideas than equivalent nominal groups (Linsey and Becker 2010), which is a very interesting finding, since the extensive research on brainstorming in the field of psychology unambiguously shows that verbally interactive group work decreases the amount of ideas compared to the amount obtained through nominal groups (e.g. Mullen et al. 1991; Byron 2012). Using ideation methods in an optimal way clearly has a great potential to increase the number of ideas which a team can contribute during concept generation. This is especially important since past research on brainstorming in the psychology field has found that the quantity and quality of ideas are highly correlated (Dugosh et al. 2000; Paulus et al. 2015), and some supporting evidence for this finding has been found in the area of engineering design (Yang 2009; Linsey et al. 2011).

Previous studies on ideation methods have been mainly statistical studies exploring how a method or some parameter of a method influences the number of ideas generated, and in some cases the quality or creativity of ideas. A few studies of professionals using ideation methods in a natural context exist (e.g. Sutton and Hargadon 1996; Jackson and Poole 2003). The findings of these descriptive studies are available for the engineer to interpret how best to apply them in a specific context. However, no studies have, to the best of the authors' knowledge, tried to generate prescriptive knowledge regarding the design of ideation methods through the interaction between researchers and practitioners (in the present case engineers) in a real-world context. In the present study, the aim has been to develop and implement tailor-made ideation methods in a specific context, among actors on a deregulated railway market. To this end, action design research (ADR) (Sein et al. 2011) was applied within the field of engineering design. Through the guidance provided by ADR, the learning derived from the specific case was formalized into a number of design principles for ideation methods which were generalized for the class of field problems of ideation in cross-functional inter-organizational groups. The findings add to the body of knowledge that can guide engineers as to how ideation methods can be used most effectively in different contexts.

2 BACKGROUND

2.1 Background of the study

The study was performed in the context of the Swedish railway sector, with a focus on turnout development. The turnout is a vital part of the railway infrastructure and a failure of a turnout, especially in a critical location, can cause significant delays and societal costs. In 2012, turnout-related failures were among the top ten causes of hours of disruption in Sweden (Trafikverket 2013). Turnouts cause at least 13% of the maintenance costs associated with the Swedish railway (Nissen 2009). Therefore, the OptiKrea project, run by Luleå Railway Research Center at Luleå University of Technology, was initiated with the intention of promoting the technical development of turnouts, especially from the point of view of maintenance and life cycle cost. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining, utilizing and researching the railway. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions would be found. A parallel goal of the project was to develop working methods, including ideation methods, which would facilitate innovation, would be tailor-made for the railway sector and could be used in the future. At the heart of the project is the so-called “creative team”, which consists of representatives from each collaborating actor. The team is cross-functional in the sense that it represents the different functions that are important when developing a turnout, i.e. research, design, manufacturing, management, maintenance, and disposal.

2.2 Associative memory models in the context of cross-functional inter-organizational groups

Associative memory models (AMMs) used in cognitive psychology (e.g. Collins and Loftus 1975; Anderson 1983) offer a framework for understanding the benefits of ideation involving participants with different functional backgrounds and knowledge, and are used to explain and justify theoretically
the method developed in the present paper. In these models, information in the long-term memory is treated as a network of interconnected nodes. The inter-relatedness of two distinct nodes is determined by their associative strength; i.e. closely related pieces of information are connected to each other more strongly than less related pieces of information. A stimulus idea activates (and hence makes more accessible) other nodes that are connected to it, and primarily those with a strong associative connection. Nijstad et al. (2002) argued that “idea generation differs from retrieval, because (new) ideas cannot be directly retrieved from memory”. However, stored knowledge must be used to generate ideas, and therefore idea generation necessarily involves retrieval processes (Nijstad et al. 2002). Therefore, Nijstad et al. (2002) suggest that idea generation is a two-stage process in which a cue activates a set of localized strongly interconnected and semantically related features. The idea production stage follows, where these sets are used to generate (new) ideas by combining knowledge, forming new associations or applying knowledge to a new domain. These ideas are added to the search cue to activate more sets in memory, leading to additional idea generation (Nijstad et al. 2002). Since semantically related images have strong mutual ties, successively activated images will often have strong associative connections which lead to a “train of thought” exploring connected ideas (Nijstad et al. 2002). AMMs explain why productivity loss occurs in group work and provide insight into how to avoid such loss (Linsey et al. 2011). In a group context, the first idea presented in the group might activate a common set of information in all the participants which leads to quick convergence concerning a small set of ideas (Linsey et al. 2011). To achieve a performance gain through idea exchange, an ideating subject must receive ideas belonging to categories which would have a low probability of being surveyed if the subject were to work alone (Perttula et al. 2006). Brown and Paulus (2002) represented a brainstormer’s knowledge of a given problem as a matrix of category transition probabilities. Simulations of the so-called associative memory matrix model showed that, if a brainstormer is presented with ideas belonging to categories of low accessibility, the number of ideas belonging to that category increases and also the total number of ideas overall; i.e. this makes the individual a more productive brainstormer (Brown and Paulus 2002). Hence, it would appear that a group of participants with different functional and organizational backgrounds within a relevant field would have an advantage compared to a functionally homogeneous group. A functionally homogeneous group may possess similar information and connections in the part of their associative memory network that applies to the ideation topic at hand. Then it could be expected that the ideas which they would come up with as a reaction to the ideation topic would be similar, as this topic would activate similar nodes of information. Exchanging ideas with each other to stimulate further activation would not be very helpful, since these ideas would primarily continue to activate ideas similar to those which the individual would generate on their own. On the other hand, if the associative memory networks of the participants were to have different nodes or a different configuration of associative strength between the nodes, the ideation topic would activate different information depending on the participant. If a participant were to be exposed to ideas from other group members, these ideas might activate parts of their network that would not initially be activated by exposure to the ideation topic, eventually leading to a group set of both more ideas and ideas belonging to more different categories than would be expected from a group with the same number of people with homogeneous backgrounds with respect to the ideation topic at hand. In support of this, Brown and Paulus (2002) found through simulations that interactive brainwriting (i.e. group members ideating in silence and exchanging ideas in written form) is not universally superior to individual brainstorming, but most effective when the members of the group possess different knowledge of the ideation topic.

2.3 Related work on ideation

Linsey et al. (2011) compared rotational viewing (as applied, for example, in Method 635) to gallery viewing (as applied in the gallery method) in an experiment involving mechanical engineering students working on the design of a device to shell peanuts. Whereas the participants are exposed to all the generated ideas simultaneously during gallery viewing, they are only exposed to one subset of ideas at a time during rotational viewing. Linsey et al. (2011) found that rotational viewing generated the highest number of ideas if both sketches and texts were used for describing ideas. Tests of established ideation methods in the same group as that participating in the present study (Petersson et
al. 2016) showed that Method 635, using rotational viewing, generated most ideas in the shortest time compared to the gallery method and the SIL method. Paulus and Yang (2000) have shown that a failure to improve production in interactive groups compared to that in nominal groups through cognitive stimulation may be caused by participants simply not attending to stimulus ideas. Both rotational viewing and gallery viewing encourage the participants to consider the ideas from the other group members. The processing of stimulus ideas is, however, performed within the limited capacity of short-term memory, and therefore it is reasonable to assume that a person can only attend to a limited set of complex ideas simultaneously (Perttula et al. 2006). Consequently, rotational viewing may be more beneficial than gallery viewing because it helps the participant to consider only a limited set of ideas at a time. On the other hand, if the attention paid to stimulus ideas becomes excessive, the performance decreases, as has been shown experimentally by Paulus and Yang (2000) and through simulations by Brown and Paulus (2002).

Linsey et al. (2011) found that gallery viewing produced more global product solutions (corresponding to concepts in the present study) and more high-quality product solutions than rotational viewing, while the average diversity of the product solutions was greater for the rotational viewing condition. One explanation may be that participants in the gallery viewing condition focus on a limited set of ideas from the entire pool during most of the ideation session and develop similar solutions with increased quality, while during rotational viewing they are forced to consider a new set of ideas in each successive round, which might enhance the variety by activating other parts of their memory network.

Perttula et al. (2006) found that individuals generated more ideas when they discussed ideas with other individuals. Although there were some doubts about why this effect arose, there is other research (Pelled et al. 1999; Seidel and Fixson 2013) showing that teams debating ideas come up with more novel innovations, indicating a more thorough exploration of the solution space. One explanation for this might be that, while discussing, new aspects of ideas appear which participants associate with parts of their memory network.

2.4 Action design research

The action design research (ADR) method was proposed by Sein et al. (2011) in an effort to blend design research with action research. Design science is the study of artefacts in their context (Hevner 2004), whereas action research, taken generally, is intervention in a social situation in order to both improve this situation and learn from it (Susman 1978). The purpose of ADR is to generate prescriptive design knowledge through learning from the intervention of building and evaluating an IT artefact in an organizational setting to address a problem (Sein et al. 2011). According to Rogerson (2014), it is uncommon that a design-based intervention will turn out as planned at the first attempt, especially when addressing a social situation with practitioners involved. Therefore, ADR is problem-driven and aims to build design principles based on iterative cycles in the same context (Wieringa 2012).

ADR was initially proposed within the field of information systems and has been applied in several different contexts, e.g. addressing the learning needs of the deaf community (Golding 2013), managing projects funded by a third party (Gröger 2014) and greenhouse gas emission reporting in the meat industry (Hilpert 2013). An IT artefact is a specific bundle of hardware and software that is assembled to fulfil information needs. Therefore, the present authors considered it reasonable to assume that the development of non-IT-based tools for information handling could also profit from ADR. As the current study aimed to develop an ideation method suitable in a certain context through interaction between researchers and engineers with different functional knowledge and from different organizations, it was deemed that ADR would be useful in guiding the study, as its strict and explicit principles and iterative cycles allow the researcher to deliver a practical outcome for the involved organization whilst simultaneously meeting academic standards.

By developing a method in close cooperation with practitioners, we expected that the user satisfaction would be higher compared to that achieved through the established methods tried by the group. It was further expected that the method would be easy to apply in new situations with participants from different organizations and different functional backgrounds, although it was not within the scope of the present study to test this expectation, but rather to test it in a future field test.
By implementing in the method being developed the findings of other researchers concerning the optimization of idea generation, we expected that the quantity of ideas generated would be greater compared to that generated using established ideation methods (Petersson et al. 2016).

3 METHOD

3.1 Procedure

ADR consists of four stages (Sein et al. 2011): 1) problem formulation; 2) building, intervention (i.e. test of the artefact in the target environment), and evaluation (BIE); 3) reflection and learning; and 4) formalization of learning. Each stage is based on certain principles and involves the execution of certain tasks. An overview of the method is shown in Figure 1. A full explanation of the principles and tasks and an account of how ADR was applied in the development of the novel ideation method, as well as reflections on the feasibility of using ADR for developing methods in the area of engineering design, are to be found in Petersson and Lundberg (2016). How each stage was carried out in the present study is briefly summarized in this section.

![Fig. 1 The stages of ADR and their adherent principles and tasks. Adapted from Sein et al. (2011)](image-url)
3.1.1 Problem formulation

The trigger for the present study was the insight that, as the different functions involved during the life cycle of railway products were spread out over different actors as a result of deregulation, developing new products would possibly become more difficult and the risk of sub-optimization would increase. As the function-specific knowledge belonged to different organizations, it was suspected that special efforts were needed to achieve a holistic picture of a product when developing it. Previous research has shown examples of both the customer’s (Hannola et al. 2009) and the subcontractor’s (Liker et al. 1998) importance for the introduction of improvements. By involving representatives from several actors when ideating, more viewpoints on the product to be developed would be shared and thus a better product would eventually be designed. The initial research question formulated was therefore how ideation should be performed to capture the benefits of collaboration during product and process development in the given context, and, specifically, how an ideation method should be designed. The participants in the project have been the infrastructure manager of the Swedish railway (the Swedish Transport Administration (STA)), a turnout manufacturer, a maintenance contractor, and researchers from Luleå University of Technology. Representatives from each participating organization formed the so-called “creative team”, in which issues concerning turnouts were proposed and addressed by means of the strategies proposed by the researchers managing the project. In Sweden, the infrastructure manager typically runs its development and research projects in cooperation with or through research institutes and universities, and therefore academic researchers were relevant participants in the group. After each test, the participants in the creative team gave their views on the method. The researchers analysed the data from the test and proposed changes to the method, and the participants of the creative team commented on the changes before the next test.

To the best of the authors’ knowledge, no ideation method has been developed specifically for groups with members possessing different functional knowledge and representing different organizations. According to Straus et al. (2011), there is no knowledge of what happens in inter-organizational groups to which different actors bring different cultures and agendas. Therefore, it was concluded that design principles for ideation methods for cross-functional inter-organizational groups would be interesting for other fields as well, as deregulation and outsourcing have become increasingly common. As a result, the research task was framed as addressing the following class of field problems: ideation in cross-functional inter-organizational groups. A literature review was performed on the subject of ideation methods and an exploratory case study was performed of a cross-functional inter-organizational group (the same group as that participating in the current study) testing three different established ideation methods (Petersson et al. 2016), in order to both validate previous research results and find empirical evidence that could be used to make an initial design of the ideation method to be developed.

3.1.2 Building, intervention and evaluation

The BIE cycles in the development of the method were divided into an alpha phase comprising iterations in the creative team (described in the present paper) and a beta phase involving future field tests. The first step of the alpha phase involved identifying the basic requirements for the method to be developed. These basic requirements, together with the results and conclusions obtained from testing established ideation methods (Petersson et al. 2016) and from previous research literature, served as a basis for working out a proposal for a novel method. The suggested method was discussed in the team and a few changes were introduced before the test of the first version of the method, referred to as OptiKrea method version 1 (OKMv1) throughout this paper. The method was tested, evaluated and changed in two iteration steps. After two iterations, the participants were satisfied with what became the second version of the method (OKMv2) and did not want to make further modifications.

3.1.3 Reflection and learning

The ideation method developed in the present study was adapted as the study progressed, according to the evaluations and analysis that took place, to reflect the increasing understanding of how the method
performed in reality. The increased understanding of the method characteristics clarified what broader
class of problems the method may be applied to.

3.1.4 Formalization of learning

Sein et al. (2011) suggest viewing the outcome as a solution that addresses a problem, and performing a
generalization by making a conceptual move on three levels: 1) generalization of the problem
instance, 2) generalization of the solution instance, and 3) derivation of design principles, i.e.
recommendations on how solutions addressing the same class of problems should be designed, on the
basis of the design research outcome. The third level requires a reconceptualization of the learning
from the specific solution instance into design principles for a class of solutions as identified in level
2. In the present study, the learning from the specific instance was translated into design principles for
ideation methods to be used in cross-functional inter-organizational teams.

3.2 The creative team

The creative team was a group of six participants comprising four railway professionals and two
academics. The professionals each had 20-30 years of experience from the railway sector. Two
worked with turnout-related issues at STA, one at an international company manufacturing turnouts
and one at a contractor performing maintenance. The academics had experience from railway-related
projects performed in close cooperation with STA and industry. Details of the participants are
provided in Table 1. Before the OptiKrea project, only participant E had experience of ideation
methods.

<table>
<thead>
<tr>
<th>#</th>
<th>Organization</th>
<th>Position</th>
<th>Field of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>Product and method development</td>
</tr>
<tr>
<td>B</td>
<td>Supplier</td>
<td>Development manager</td>
<td>Product development of turnouts, turnout engineering</td>
</tr>
<tr>
<td>C</td>
<td>STA</td>
<td>Specialist</td>
<td>Maintenance management of turnouts</td>
</tr>
<tr>
<td>D</td>
<td>STA</td>
<td>Specialist</td>
<td>Track and turnouts</td>
</tr>
<tr>
<td>E</td>
<td>Academia</td>
<td>Professor</td>
<td>Mechanical engineering design, maintenance,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>product development</td>
</tr>
<tr>
<td>F</td>
<td>Academia</td>
<td>Postdoc. research fellow</td>
<td>Applied acoustics and signal processing</td>
</tr>
</tbody>
</table>

3.3 Test of ideation methods

The two versions of the developed method were tested by the group with an interval of four weeks
between the tests. The established methods in the exploratory case study were tested over a three-
month period that commenced about one year before the tests of the developed method were initiated.
Between the two tests of the developed method, the participants worked with their normal tasks and
did not use any ideation methods.

A different ideation topic was chosen for each ideation method, since using the same topic for all
the methods would mean that solutions from the first ideation session could be reused in the
subsequent sessions and comparison of the ideation outcomes between the methods would not be
meaningful. Instead our aim was to find equivalent problems that were based on actual needs, were
domain-specific and open-ended, and had a large amount of possible solutions. Firstly, the participants
were asked to come up with topics which they wanted to work on within the OptiKrea project and
which, based on their experience, they thought were equivalent. Thereafter, the authors of the paper
made a review of the suggestions to pick out the topics that best fulfilled the equivalence criteria. The
participants were informed of this selection and agreed upon the ideation topics about one week before
each ideation session. The chosen ideation topics are presented in Table 2. The issues on which the
ideation was performed required at least a basic technical understanding in general and railway-
specific knowledge in particular.
Table 2 Ideation topics used for each ideation method test

<table>
<thead>
<tr>
<th>Method</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>How can track geometry deterioration be prevented in turnouts?</td>
</tr>
<tr>
<td>Gallery</td>
<td>How can transition zones between rail parts of different rigidity be designed to ensure a smooth transition?</td>
</tr>
<tr>
<td>SIL</td>
<td>How can turnouts be protected from snow and ice?</td>
</tr>
<tr>
<td>OKMv1</td>
<td>How can the position of a modular turnout be adjusted?</td>
</tr>
<tr>
<td>OKMv2</td>
<td>How can satisfactory drainage of the track superstructure be achieved?</td>
</tr>
</tbody>
</table>

At the start of each session one of the participants presented the ideation topic, followed by a common group-analysis of the topic to identify causes of and known solutions to the problem involved (30-60 minutes long), before the test of one of the ideation methods started. Before a test started, the participants were instructed to try to create many ideas and not to criticize each other negatively during the application of the ideation method. They were instructed that thorough scrutiny (both negative and positive) would be possible in a subsequent step of the project. The execution of the established methods which were tested in the exploratory case study (Petersson et al. 2016) and to which the developed methods will be compared in terms of outcome is presented in Table 3.

A detailed description of the tests of the established ideation methods is to be found in Petersson et al. (2016). The first author presented the instructions for the method concerned to the group by means of a projector, ran a timer to keep track of the time, announced when the group should move on to the next stage of the method and answered questions about the instructions during the session. In this capacity the first author is referred to as the “supervisor” throughout the paper.

3.4 The views of the participants

Both after the test of version 1 of the developed method and after the test of version 2, the participants answered a questionnaire and a short group interview regarding the participants’ experience of the applied method took place to capture the participants’ immediate reactions to the ideation method that they had tried. The questionnaire presented different statements about each method, and the participants assessed their level of agreement with these statements, answering on a continuous scale from “Do not agree at all” to “Agree completely”, by making a mark on a line which was about 100 millimetres in length and where 0 represented “Do not agree at all”. The position of the mark was measured in millimetres from 0 with a ruler and divided by the total length of the line. The average value for each statement was calculated and, despite the small number of participants, the average values for the subgroups “academics” and “railway professionals” to see if there were any differences in their views. The group interviews were audio-recorded and the transcribed recordings, along with the questionnaires, served as the basis for analysing the participants’ views on the methods. A linear regression was performed between each pair of statements in the questionnaires to find out if any correlation existed. All the five methods tested by the creative team during the project were included, resulting in 30 data points.

3.5 Number of concepts and ideas

Each method was divided into individual ideation phases (IPs) and group review phases (GPs) and an analysis was made of how the number of non-redundant concepts and ideas varied over the different phases and between the methods. Further, the number of ideas reused in a later phase of the versions of the developed method was counted.

Each set of sheets of paper containing the suggestions from each ideation session and each recording from the group phases were examined carefully. As a starting point, concepts were defined as sketches and text that clearly belonged together and were marked off from other sketches and text on the sheets of paper. The contents of all the concepts were hierarchically broken down into key phrases consisting of a verb phrase containing a maximum of one verb, or a noun phrase. These key phrases were defined as ideas. All the ideas from an ideation session were organized systematically and the key phrases that implied physically doing the same thing were defined as being the same idea, although different words were used to describe that idea. When all the ideas had been organized and
compared, the total number of ideas from each phase could be counted, as well as the number of ideas contained in each concept. Some concepts contained variants that were not compatible. If half or more of the ideas were incompatible, the concept was split and counted as two different concepts that shared some ideas. Similarly, if more than half of the ideas of each of two concepts were shared by those concepts, the two concepts were counted as one concept containing variants. A summary of the rules for counting concepts is to be found in Table 4, while a corresponding summary for ideas can be found in Table 5.

Table 3 The execution of the established ideation methods that were tested (Petersson et al. 2016).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP</td>
<td>GP</td>
</tr>
</tbody>
</table>

**Table 3**

**Method 635**

**IP1** 1) Each participant works individually during five minutes and comes up with at least three suggestions on how to address the ideation topic. The ideas are sketched and/or written down on a sheet of paper.

2) When the time is up, each participant gives the sheet containing the ideas to the neighbour on their left.

3) Each participant reads through the suggestions on the sheet of paper which they have received from the neighbour on their right and adds improvements/comments to suggestions, combines suggestions to form new suggestions, and/or uses suggestions as a source of inspiration to come up with new ideas, during a total time of five minutes. One is allowed to ask the neighbour on one’s right what is meant by a suggestion that has been received.

4) When the time is up, each participant gives the sheet of paper before them, which now contains the work of two participants, to the neighbour on their left, and step 3 is repeated.

5) The process continues until each sheet of paper has passed between all the participants, i.e. when it has been returned to the person who started working on it as a blank sheet of paper.

**Gallery method**

**IP1** 1) The participants individually sketch and/or write down suggestions on how to address the ideation topic on a sheet of paper for 15 min.

**GP1** 2) The sheets of paper are attached to a wall. The group gathers around one of the sheets hanging on the wall and the creator of the suggestions explains them to the other participants, who give constructive feedback. The group then moves on to the next sheet, which is explained and receives feedback, and this process is continued until all the participants have received feedback on their suggestions. Each participant can use approximately five minutes for presentation and receiving feedback.

**IP2** 3) Each participant takes down their sheet of paper and works individually on that sheet for five minutes to develop their ideas or come up with new ideas using the feedback which they have received from the other participants and using the other participants’ suggestions as a source of inspiration.

**SIL method**

**IP1** 1) The participants individually sketch and/or write down their suggestions as to how to address the ideation topic during ten minutes.

**GP1** 2) Two participants are randomly selected and each of them presents one suggestion to the rest of the group. The participants themselves choose what suggestion to present.

3) All the participants try to combine these suggestions into one concept by interacting verbally and sketching/writing on a whiteboard.

4) When the group is finished with the first two suggestions, a third group member presents another suggestion. The group then tries to combine this suggestion with the suggestion which resulted from step 3. Alternatively, a fourth suggestion is presented by any participant and an attempt is made to combine this fourth suggestion with the third suggestion.

5) The process of presenting suggestions and trying to combine them with each other continues until all the ideas have been presented or the time is up (60 minutes).
Table 4 Summary of the rules for counting concepts

1. A concept consists of one idea or a combination of ideas, and stands as a solution to the problem on its own merit. The concept in question may address only one subpart of the problem and it may be possible to combine the concept with other concepts.

2. If a concept contains incompatible ideas and half or more of the ideas are incompatible, the concept is split and counted as two different concepts that share some ideas.

3. If more than half of the ideas of each of two concepts are shared by those concepts, the two concepts are counted as one concept containing variants.

4. If a concept is a subpart of another concept, they are counted as one concept.

5. Concepts that reframe the problem and do not specifically address the problem as described, but meet higher-level needs are counted.

Table 5 Summary of the rules for counting ideas. Adapted from Linsey et al. (2005, 2011) for the current study

1. Ideas are units which a concept is systematically broken down into and an idea can be expressed as a key phrase consisting of a verb phrase containing a maximum of one verb, or a noun phrase.

2. If a certain idea is used multiple times in a concept, this idea is counted once for that concept.

3. New combinations of already-counted ideas are counted as separate ideas.

4. Ideas count even if they are unnecessary or deteriorate the concept.

5. Categories of ideas only count as ideas when no subordinates are given.

6. Ideas must be shown and not just implied to be counted.

7. Ideas in different concepts are counted as the same idea if they imply physically doing the same thing, although different words are used to describe that idea.

3.6 Behaviour of the participants

Observations made during the ideation sessions, together with the transcribed audio recordings of the sessions, served as a basis for analysing the behaviour of the participants. A variant of the coding system developed by Jackson and Poole (2003) was used to analyse the content of the ideation workshops. The number of words used in the different activities was counted. The main activities specified by Jackson and Poole (2003) were used: idea statement, elaboration, criticism, direction, and going off at a tangent. A description of each activity is to be found in Table 6.

Table 6 Activities used to code the ideation sessions, drawing upon Jackson and Poole (2003)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea statement</td>
<td>Spoken contribution of an item to be recorded as a possible solution to the problem.</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Non-critical clarification (explaining), rephrasing or discussion of ideas.</td>
</tr>
<tr>
<td>Criticism</td>
<td>Negative statements or judgments about proffered ideas.</td>
</tr>
<tr>
<td>Direction</td>
<td>Guiding or structuring the idea-generation activity.</td>
</tr>
<tr>
<td>Going off at a tangent</td>
<td>Interaction that is off-topic and breaks the “singlemindedness” of the idea generation, whether or not it relates to other group work tasks, e.g. jokes, discussion of the group’s task in general or discussion of given directions.</td>
</tr>
</tbody>
</table>

The distribution of the words spoken by the participants during the ideation session was analysed. The distribution of the speaking frequency, i.e. how many times each participant spoke during a session, was also calculated. During this analysis, agreeing expressions consisting of a maximum of three words that were not an answer to a direct question were neglected (e.g. yes, h’m). This was partly because such expressions did not add anything to the discussion and partly because they were difficult to hear on the recordings (especially when uttered by those participants sitting furthest away from the microphone) and counting would therefore not give a fair picture of the speaking distribution. However, when a speaker was interrupted by such a comment and then had to “speak up” again to continue, this was counted as speaking twice, since speaking up again after being interrupted takes an effort and is important in understanding the speaking distribution. The same rule applied when a speaker said something that made the others laugh and then had to start speaking again. It was not
possible to interpret all the verbal communication of the recordings of the workshops, and the uninterpretable part was not included in the analysis. However, this part is so small (estimated to be roughly 1% of the total number of words, and impossible to calculate exactly, since one cannot judge exactly how many words are not heard in each instance) that it would only have had minor effects on the results of the analysis and would not have changed the overall conclusions.

If the speaking distribution had been equal between the participants, each participant (in the group of six participants) would have spoken an ideal fraction corresponding to one sixth of the number of words remaining after subtracting the number of words uttered by the supervisor from the total number of words spoken. The ideal fraction was compared to the actual fraction and the average of the absolute values for how much the participants deviated from the ideal fraction was calculated to be able to compare the “equality of speech” between the methods. A corresponding calculation was performed for the speaking frequency.

An overview of the areas of focus during the evaluation, the sources of evidence and the methods applied to analyse the data is presented in Table 7.

<table>
<thead>
<tr>
<th>Areas of focus</th>
<th>Sources of evidence</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>The participants’</td>
<td>Questionnaires</td>
<td>Quantitative analysis of the average value of each questionnaire statement</td>
</tr>
<tr>
<td>views</td>
<td>Transcribed audio recordings of group</td>
<td>Quantitative analysis of the correlations between pairs of statements in the questionnaires</td>
</tr>
<tr>
<td></td>
<td>interviews</td>
<td>Qualitative analysis</td>
</tr>
<tr>
<td>Outcome of the</td>
<td>Sheets of paper from the workshop</td>
<td>Quantitative analysis (during different phases of the methods) of the following:</td>
</tr>
<tr>
<td>ideation sessions</td>
<td>Transcribed audio recordings of sessions</td>
<td>- the number of non-redundant concepts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- the number of non-redundant ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- the number of ideas reused in later steps</td>
</tr>
<tr>
<td>Behaviour of the</td>
<td>Observations</td>
<td>Qualitative analysis</td>
</tr>
<tr>
<td>group</td>
<td>Transcribed audio</td>
<td>Coding of the following:</td>
</tr>
<tr>
<td></td>
<td>recordings of sessions</td>
<td>- the distribution of spoken words between different activities (Table 6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- the distribution of spoken words between participants</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- the speaking frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Calculation of the average deviation from an equal distribution of the spoken words and speaking frequency</td>
</tr>
</tbody>
</table>

4 RESULTS AND DISCUSSION

Section 4.1 presents how the initial design of the method was reached and describes what changes were made to the method after the first test in the creative team. The changes were introduced to meet needs that were identified during the evaluation of the initial design. The remainder of Section 4 is dedicated to a detailed presentation and discussion of the results from the evaluation of the tests of the different versions of the developed method and a comparison of these results with those from the test of the established methods which the developed method is based upon.

4.1 Design of the method

4.1.1 Requirements and demands set for the method

From the interviews and the discussions within the group, two overall requirements emerged for the method to be developed: a) the method must produce a great number of ideas with breadth and potential, and b) the method must be integrated into the product development routines of STA and implemented by STA after the project finishes. Prerequisites for this implementation are that the intended users must feel that the method is easy to initiate (a low usage barrier) and the inclusion of
different participants (from different organizations) on different ideation occasions must be easy. From the interviews and the discussions among the group members, it emerged that, to meet these demands, the method must be easy to use and understand, as well as attractive to the users, and preferably they should find it both useful and exciting. Despite previous findings that facilitators (e.g. Kramer et al. 2001) and training (Parnes 1959) can enhance the productivity of ideation, both of these alternatives were ruled out, because they would raise the usage barrier too much. Partly for the same reason, special equipment or software was ruled out. Another reason for excluding software was that organizations typically (e.g. in the present case) do not allow external software to be downloaded onto their employees’ computers. In addition, computational tools for the conceptual stage do not yet provide the same usability, the same ease of annotation of design sketches with rough dimensions, and the same possibility of making notes as do pen and paper (Yang 2009).

The results from the tests of established ideation methods showed that during the SIL method the participants were quite confused about what they were supposed to do and asked a large number of questions about the procedure during the entire SIL workshop (Petersson et al. 2016). It was therefore concluded that the SIL method would not be feasible to explore further since it would be impossible to achieve the goal of easy implementation with that method; therefore, the SIL method was excluded from the scope of this paper. Neither Method 635 nor the gallery method encountered similar problems and were therefore judged to be a better starting point for the method to be developed.

4.1.2 Individual and group work

Drawing upon the implications of AMMs, it was determined that a feasible design would be to start with a first step of individual ideation where the participants would document the ideas which they would first come up with as a reaction to the ideation topic. In this way, they would not be affected by the other participants’ ideas initially. After the first individual step, AMMs suggest that the participants should be exposed to other ideas than their own to activate not-yet-accessed parts of their memory network. Participants can be exposed to ideas generated by others in the group or external ideas such as pictures of objects from nature. Previous research suggests that the effectiveness of stimuli depends on the appropriateness of the stimuli with respect to the given topic, participants and context (Perttula and Sipilä 2007). It was decided that the group would use internally generated ideas during the ideation method because a) it was expected that these ideas would have a higher probability of appropriateness than selected external stimuli, b) someone would have had to make the effort of selecting the external stimuli before the ideation session, which would have been a barrier to using the method, and c) the different functional backgrounds of the participants would provide ideas with enough variety.

4.1.3 Sharing of ideas

There are several ways of exchanging ideas that can activate the memory network in different ways. Since the ideation method to be developed did not have any given time restrictions, it was decided that a mix of idea-exchanging strategies would be a good starting point. Linsey et al. (2011) suggested that “an improved process for idea generation consists of first using a gallery communication method to generate a large number of high quality product solutions and then moving to a rotational viewing method using words and sketches to develop the details of the product solutions and a large number of functional ideas”. However, in the light of AMMs, one should first aim to achieve a large pool of ideas with great variety that can activate different parts of the memory network. Therefore, the method developed in the present study starts with individual ideation with rotational viewing followed by an IP with gallery viewing. After each viewing phase, the participants exchange ideas through presentation and discussion in verbal interaction steps. The participants of the present study thought that presenting and discussing ideas were very important for reaching good solutions, and this feature of the method made it particularly attractive, since the participants really enjoyed interacting verbally (Petersson et al. 2016). Further, the presentation steps let each participant gain an overview of the entire set of ideas generated by the group. One difference between interactive and nominal groups is the fact that, after the ideation session, the members of an interacting group will have knowledge about the entire set of generated ideas, whereas a member of a nominal group will
only know about the subset of ideas that he or she has generated. In consideration of this, it is not very surprising that the participants in an interactive group typically feel more productive (Stroebe et al. 1992; Paulus et al. 1993) and enjoy the ideation session more (Furnham 2000). It is reasonable to assume that the sharing of ideas is an important factor in motivating individuals to participate and engage in the ideation session. The procedure of the first version of the developed method is shown in Table 8. It is acknowledged that GP2, the presentation and discussion of ideas from IP2, might not add very many new ideas, but rather is used to conclude the session for the group. However, knowing that others will view the ideas which one has come up with might enhance the idea generation in the earlier steps through social comparison (Perttula et al. 2006). Further, it might be possible that this last step may activate new thoughts among the participants, who may come up with ideas directly after the session, after a period of incubation, or later when interacting with some other stimuli.

Table 8 The execution of the first version, OKMv1, of the developed method. IP=individual phase, GP=group phase

<table>
<thead>
<tr>
<th>Phase</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP1</td>
<td>1) Steps 1-5 of Method 635 (see Table 3).</td>
</tr>
</tbody>
</table>
| GP1   | 2) The participants use five minutes to read through the ideas that have been added to the sheet of paper which they started out with. 
3) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand. 
4) After each presentation, the other participants give feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of ten minutes is allowed per sheet of paper (for presentation and feedback). |
| IP2   | 5) The sheets of paper from step 1 are put up on a wall or some other place where all the participants can easily view them. 
6) Each participant works individually for ten minutes to develop or combine ideas from the collection of ideas from step 1. New ideas are also welcome. New sheets of paper are used to document the ideas by means of sketches and/or text. |
| GP2   | 7) Each participant presents their own ideas from step 3. 
8) After each presentation, the other participants give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of five minutes is allowed per participant (for presentation and feedback). |

4.1.4 Form of representation

According to Linsey et al. (2011), embodied cognition theories suggest that external representations such as sketches are helpful for the performance of design tasks since they reduce the cognitive load, as the amount of information which an individual has to represent internally is reduced. On the other hand, whereas information such as geometry and configuration tends to be easier to express in drawings, other information such as abstract concepts might be easier to convey in words (Linsey et al. 2011). Additional information in words might also be helpful when other individuals interpret a sketch, which saves time during both rotational and gallery viewing. Particularly worth noting is that individuals have different preferences and some individuals simply do not like to sketch. Therefore, the participants were instructed to use both sketches and text, to the extent desired by them, when documenting their ideas. An advantage of the developed method is that all the ideas are documented in sketches and text by the participants continuously during the workshop, so that no special person needs to be appointed to record ideas, which facilitates implementation. Table 9 sums up the design principles used to set up the initial version of the new method.
Table 9 Design principles for ideation methods to be used in cross-functional inter-organizational groups, their justification and implementation

<table>
<thead>
<tr>
<th>Design principle</th>
<th>Justification</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used to design OKMv1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developing the method based on methods proven to be easy to understand</td>
<td>To increase the probability of implementation. To make it easy to involve new participants.</td>
<td>Combines elements from Method 635 and the gallery method.</td>
</tr>
<tr>
<td>The initial generation of a pool of ideas with great variety</td>
<td>Many different types of ideas to create associations with.</td>
<td>The method starts with individual ideation without interaction or other stimuli.</td>
</tr>
<tr>
<td>Exposure to stimulus ideas generated internally within the group</td>
<td>To activate new areas of each participant’s memory network. High probability of appropriateness. No additional work.</td>
<td>Participants are exposed to ideas from the other participants, silently during the round of IP1 and verbally during the group phases.</td>
</tr>
<tr>
<td>Mixing idea-exchanging strategies</td>
<td>Exchanging ideas in different ways may activate the memory network in different ways.</td>
<td>Combines rotational viewing and gallery viewing with verbally interactive steps.</td>
</tr>
<tr>
<td>The use of external representation in the form of sketches and text</td>
<td>To reduce the cognitive load. Sketch and text appropriate for different types of information. To facilitate the understanding of concepts. Personal preferences. Inherent documentation.</td>
<td>The participants themselves choose if they want to document their ideas with sketches or words.</td>
</tr>
<tr>
<td>Including time for discussion and debate</td>
<td>To increase the quantity and variety of ideas. Attractive to participants.</td>
<td>Individual phases are followed by a group phase.</td>
</tr>
<tr>
<td>Added after testing OKMv1 in the group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimizing the cycle time</td>
<td>Enough time for participants to exhaust their own ideas in the first round and to review and react to the stimulus ideas from other participants in the remaining rounds.</td>
<td>Time increased from five to ten minutes during each round of IP1.</td>
</tr>
<tr>
<td>Developing a strategy for distributing the verbal interaction</td>
<td>All perspectives and knowledge in the group should be considered.</td>
<td>Participants take turns to comment on other participants’ suggestions.</td>
</tr>
<tr>
<td>Added after reflections on the method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paying attention to stimulus ideas</td>
<td>Participants must attend to stimulus ideas to be inspired by them.</td>
<td>Rotational viewing during IP1.</td>
</tr>
</tbody>
</table>

4.1.5 Length of each round of rotational viewing – adjustment made in OKMv2

Two changes were made to the initial design of the method after the first test. The time allocated for each round of IP1 was prolonged from five to ten minutes, as the participants all agreed that five minutes was too short a time to be able to document all their own ideas during the first round, and to read through and understand the ideas on the sheet of paper which they had received and then come up with new ideas during the remaining rounds. This is in line with Nijstad et al. (2002) reasoning that stimulus ideas can interfere cognitively with a person’s train of thought if they belong to a different semantic area of memory. The train of thought is then cut off, leading to a loss of potential ideas. Therefore, the time allocated for each round must be long enough to allow the participants to reach the
end of their train of thoughts, as activated by exposure to the ideation topic and stimulus ideas. With each successive round, each sheet of paper included more and more information, and it took a longer time for the participants to acquaint themselves with this increased amount of information during the remaining rounds.

4.1.6 Distribution of spoken words – adjustment made in OKMv2

The second change concerned the unequal distribution of spoken words among the participants, which was found after the first version of the developed method had been tested and is dealt with in greater detail in Section 4.3. The distribution of spoken words is important, since previous research claims that it is essential that all the perspectives and knowledge are contributed in the group (Milliken et al. 2003). New information can activate new areas in the participants’ memory network, and to access and use the information which each participant possesses, all the participants must speak. Some persons are inclined to speak more and others less (e.g. Burke 1974), and therefore it was important to find a simple procedure that could even out the distribution of spoken words. Instead of letting the discussion be completely free, the change was introduced that, after each participant had presented his suggestions, the other participants took turns to comment on the presented ideas. Whenever questions and urgent comments arose, the participants were still allowed to speak out of turn. The steps of the second version, OKMv2, are shown in Table 10.

Table 10 The execution of the second version, OKMv2, of the developed method. IP=individual phase, GP=group phase

<table>
<thead>
<tr>
<th>Phase</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP1</td>
<td>1) Step 1-5 of Method 635 (Table 3), but with ten minutes in every round instead of five minutes.</td>
</tr>
</tbody>
</table>
| GP1   | 2) The participants use five minutes to read through the ideas that have been added to the sheet of paper which they started out with.  
3) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.  
4) After each presentation, the other participants take turns to give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of ten minutes is allowed per sheet of paper (for presentation and feedback). |
| IP2   | 5) The sheets of paper from step 1 are put up on a wall or some other place where all the participants can easily view them.  
6) Each participant works individually for ten minutes to develop or combine ideas from the collection of ideas from step 1. New ideas are also welcome. New sheets of paper are used to document the ideas by means of sketches and/or text. |
| GP2   | 7) Each participant presents their own ideas from step 3.  
8) After each presentation, the other participants take turns to give their feedback on the ideas (questions, improvements, potential, etc.). The remaining available time is used for discussions. A maximum of five minutes is allowed per participant (for presentation and feedback). |

4.2 The participants’ views

Section 4.2.1 and 4.2.2 present the results regarding the participants’ views on the versions of the developed ideation method and rely on data from the group interviews. Quotes are included to exemplify typical qualitative data that contributed to the results. The letter in brackets after each quote indicates who is quoted, according to Table 1.
4.2.1 The participants’ views after testing OKMv1

All the participants thought that OKMv1 was the best of all the methods which they had tried thus far, i.e. Method 635, the gallery method, the SIL method and OKMv1. They liked combining elements from Method 635 and the gallery method, since they thought that this gave them the opportunity to exhaust the ideation topic more thoroughly than they had been able to do with the methods tested previously. This is exemplified by the following quote:

“Directly after this [Method] 635 we got to do this [gallery method]. […] On the same ideation topic, and then we could exhaust the topic more.” (F)

All the participants felt that they had progressed further regarding the complexity and details of the concepts and ideas, and that the concepts and ideas had evolved in more directions and in greater detail than when they had tried the established ideation methods.

All the participants agreed that five minutes in each round of IP1 was too short a time. They had not had enough time to document their own initial ideas in the first round. Four participants mentioned specifically that, when the next round had started, they had forgotten their own remaining ideas and had concentrated on making associations with the ideas on the paper which they had received. However, reading through and understanding these ideas had consumed most of the minutes that had been available, and therefore there had not been much time left to build on these ideas. This is exemplified by the following quotes:

“Once you get the sheet from your neighbour you get a bit stressed, you don’t have the time to read through and understand all […] the innovations or ideas before you have to [add your own ideas], there is so little time left to try to come up with something new yourself.” (F)

“When I got [my neighbour’s sheet with ideas…]. I sat wondering what I should do with them, but I had my own ideas, of course, and […] his ideas] were completely different from what I was thinking about. […] But then I lost [my own ideas before I got to write them down]. […] Now I was supposed to understand what he had [suggested…].” (D)

Two participants mentioned specifically that they were surprised and impressed by the breadth of the ideas that had emerged during the session.

When asked if they would have preferred working in groups of two, for example, during IP2, all the participants agreed that they preferred working alone during the individual steps. All the participants thought that they could create ideas with more breadth by working individually, that working on their own saved time because working in pairs or teams required more time for explaining, and that there was enough group work during the presentation steps. Also mentioned was the fact that it was good to have time to develop one’s own ideas on one’s own. This is exemplified by the following quotes:

“We have better possibilities of [covering a wide range of ideas] if we work individually in this third step, in my view. The generation of ideas and then the presentation of them, that is when you have the common discussion. Then the group is working.” (D)

“I agree […] that it was a suitable mix of working individually and working together.” (E)

Two of the participants mentioned specifically that they had not been in a good creative mood on the day when OKMv1 had been tested, in one case because of a lack of sleep.

4.2.2 The participants’ views after testing OKMv2

All of the participants agreed that OKMv2 represented another step forward and was the best method that they had tried thus far. This is exemplified by the following quotes:
“I wrote that it is a step forward, that I think that this very fact that [we have more time] means that we can process, think about and grasp what people mean instead of just spurting out ideas. So I feel, as it were, that this is a method that, as I see it, could be put into practice.” (D)

“I agree, this was the best we accomplished, I believe. It felt completely right, all the way. [Sometimes] there were even some spare minutes, that was first-rate, I thought.” (C)

“Oh, best so far, definitely. […] I agree, I was hardly even close to thinking about drainage [before] and yet lots and lots of [ideas] emerge and then the way in which the quality is improved from [the Method 635 step] to the [gallery step], yes, it’s clear as a bell. Terrific.” (B)

The participants all thought that prolonging the time for the rounds of rotational viewing was excellent, but had somewhat different opinions about how to distribute the time. One participant thought that it would be better if the first round was longer than the others. The other participants, however, thought that ten minutes was appropriate for every round. All the participants thought that there was sufficient time to understand, consider and make associations with other people’s ideas, as well as to document one’s own ideas. This is exemplified by the following quote:

“No, ten [minutes] was excellent, it’s quite sufficient, because somehow I felt there was peace and quiet and we had the time to read through [the ideas] before starting to comment; [during OKMv1] we started to comment directly, as it were, one didn’t have the time.” (D)

Four of the participants mentioned that, during the last rounds of rotational viewing, they had actually had a few minutes left which they had thought could perhaps have been used for IP2, the second individual step, instead. None of the participants wanted to remove IP2, despite the fact that IP1 had been prolonged. They wanted to use IP2 to summarize or to develop concepts, or to be able to work on new ideas which they had created during GP1. Half of the participants thought that the time allocated for IP2 had been sufficient, whereas the other half felt that they could have used a few minutes more.

Three participants mentioned specifically that they were surprised by the large amount of ideas which they had conceived regardless of whether they had been very familiar or less familiar with the ideation topic. This is exemplified by the following quotes:

“I hadn’t expected to discover any ideas, but it felt as if my brain got going anyway, and I think that was good.” (F)

“Just to see visually what we [created], what we got down on paper, my goodness, we were never close to that any time before. And on a topic which I think, [although there is good knowledge of it within the group,] as for me, I have only scant knowledge of it, yes, well, incredibly good.” (C)

Letting all the participants sitting around the table speak in turn was appreciated, because this meant that everyone was given a hearing.

One participant suggested that the developed method could be improved by having each group member think of three ideas on the topic in question before the ideation session and bring it to the session as a starting point for IP1. The other participants, however, were reluctant to accept this suggestion, because they thought that it would be difficult to arrange a presentation and clarification of the ideation topic in advance of the session, that “crazy” ideas would disappear and that it would not be possible to achieve the quick responses and communication which the developed method made possible.

4.2.3 Questionnaire

The participants’ average levels of agreement with the statements in the questionnaire are shown in Figure 2. The responses to the statements presented in Figure 2 were also analysed separately for the academics and railway professionals. It turned out that the variation between all the individual participants was greater than that between the two subgroups, and therefore no conclusions about
possible differences between the two subgroups could be established. The only statement for which the variation between the individuals was less than that between the subgroups was the statement “Many ideas were new to me”. The academics exhibited a higher level of agreement with this statement during the test of OKMv2, which is perhaps not surprising, as the other participants had a longer experience from the railway sector. Moreover, there was a trend that the professionals from STA agreed to a lower extent than the others with the statement “Many interesting ideas were presented during the workshop”.

Fig. 2 The participants’ rankings for different statements about each method

Figure 2 shows that the engagement in all the ideation issues was quite high, although the engagement in the issue which was ideated on during OKMv1 was somewhat lower. This may be one explanation for the lower engagement felt during the OKMv1 ideation session compared to the engagement felt during the gallery method or OKMv2. However, when examining the individual ratings, it was found that it was the very low rating given by one of the two participants who had experienced a “bad day” which had pulled down the average level of engagement felt by the participants during the OKMv1 workshop. The same explanation applies to how satisfied the participants were with their own contribution.

When testing OKMv2, it was observed during IP1 that the participants had been less active during the last minutes of the last rounds. The participants reported that they had experienced difficulty coming up with new ideas at the end of these rounds and this is also reflected in the questionnaire. OKMv2 scored much lower on “If I had had more time, I would have come up with more ideas” compared to the other methods. Another way of judging whether a topic is more exhausted is to analyse the types of ideas appearing at the end of the ideation session. It was observed that the ideas that had emerged at the end of IP1 of OKMv2 had obviously been more crazy and unrealistic (e.g. killing plant life by using nuclear radiation or introducing animals that like water to address insufficient drainage at turnouts) than previous ideas. It has been reported in the literature that, when “the rate of suggested ideas has diminished almost to a standstill”, more intriguing and unusual ideas start to appear, but only a few people contribute then and most people would prefer to do something else (Byron 2012). It could therefore be argued that with OKMv2 this point in time was reached, which did not happen during the other tests, and therefore the topic dealt with during OKMv2 was more exhausted. The point in time when the participants run out of ideas is probably dependent on the ideation issue, but it can be concluded that it takes quite a long time to exhaust the types of questions
which have been dealt with in the present study; for example, in the case of OKMv2, the participants used 63 minutes during IP1. Once this point is reached, there is a choice of two alternatives: ending the ideation (or in the present case, advancing to the next step of the method) or trying to squeeze out unusual ideas at a lower rate while the members possibly start feeling that they are wasting time. Nijstad et al. (1999) found that an individual working alone or in a group stops brainstorming when they feel that it is no longer worth the effort, and this feeling is based on their subjective estimate of the probability of them being able to generate more ideas on the ideation topic. The continuation of ideation requires some effort to be made to overcome this stage, and perhaps additional measures must be taken. Although many of the ideas produced in this phase are obviously unrealistic, a few of them could lead to a breakthrough through association.

With regard to how the participants viewed the ideas generated during the ideation, OKMv1 scored highest on the number of interesting ideas presented and OKMv2 on the newness of the ideas to the participants. The variations concerning these parameters between the methods may, however, be a result of how familiar the participants were with the ideation topic. When rating the extent to which the presented ideas could be used practically, the participants awarded higher scores to the two versions of the novel method than they awarded to Method 635 and the gallery method. When analysing the data from the interviews, we could observe that this could be due to the discussions during the group phases and the possibility of refining ideas in several steps. OKMv2 scored highest on how useful the method was thought to be, which is reflected in the comments by the participants that this was the best of the methods that they had tried thus far. As expected, the developed method outperformed the other methods regarding user satisfaction.

Considering the correlations between different statements in the questionnaire, a few points are interesting to note. However, it should be kept in mind that the small number of participants and therefore data points makes the results uncertain and the findings should be regarded as tentative. The only correlation with any significance at all, see Figure 3, was between how engaged a participant felt during the workshop and how satisfied that participant was with his own contribution (R²=0.64). This tentatively indicates that there might be a causal relationship between the engagement felt by the participants and the effort which they make during the workshop. However, another possibility is that, if a participant feels more engaged, they will perceive their own contribution as better although it might not objectively have been so. Similarly, their engagement might bias their view of how useful the method is. Between the engagement felt by the participants and how useful they found the method, an R²=0.44 was found. Previous research concerning brainstorming found that group members enjoyed their work more and were more satisfied with their own performance than individual brainstormers, although they were actually less productive (e.g. Stroebe et al. 1992), and that one reason for this might be a tendency to appropriate ideas of others in the group (Paulus et al. 1993). One cannot make a straightforward comparison between these findings and the findings of the present study because of the different circumstances of the studies, but this underlines the importance of taking both objective and subjective results into account when determining what method is most promising, since the views of the participants might be biased by their engagement. It should be noted that many data points were clustered at a high rating for these statements, as is exemplified in Figure 3.

The absence of correlations for some pairs of statements is interesting. How useful the participants found the method and how satisfied they were with their own contribution during the workshop did not correlate (R²=0.18). A possible explanation for this is that the participants tried to judge how useful the method would be in general, while they rated their own contribution in the specific instance. How engaged the participants were in the ideation topic did not correlate to how engaged they felt during the workshop (R²=0.10) or how useful they found the method (R²=0.01). Other researchers (e.g. Isaksen 1998; Bolin and Neuman 2006) have suggested that ownership of the task or the strength of the incentive for high performance is important for the outcome of a brainstorm. One would intuitively think that a high engagement in the topic should lead to a high engagement during the workshop and a better outcome compared to a lower motivation to address the topic on the part of the group. The absence of a correlation between the engagement in the topic and the engagement felt during the workshop indicates that other factors are more important. However, other factors that could be expected to influence the engagement felt during a workshop, for example the extent to which the method resulted in ideas that could be used in practice (R²=0.06), how many ideas were new to the
participants ($R^2=0.01$) or how interesting the generated ideas were ($R^2=0.06$), were also uncorrelated. Neither were these factors correlated to how useful the participants found the method ($R^2=0.18$, $0.00$ and $0.12$, respectively). This may be because the right factors were not investigated in the questionnaire or because the relations between these parameters are more complex and cannot be explained by linear correlations between different pairs of statements. In any case, more instances of participants not being engaged in the topic would be required to obtain a reliable correlation.

![Linear fit (dashed line) between the participants' engagement during the workshop and how satisfied they were with their own contribution ($R^2=0.64$). “Do not agree at all” corresponds to 0 and “Completely agree” to 1](image)

**Fig. 3** Linear fit (dashed line) between the participants’ engagement during the workshop and how satisfied they were with their own contribution ($R^2=0.64$). “Do not agree at all” corresponds to 0 and “Completely agree” to 1

### 4.3 Distribution of spoken words

Table 11 shows the average deviation from an equal distribution of the spoken words and speaking frequency during the group phases. If the number of spoken words had been completely equally distributed among the participants, the deviation would have been zero. Table 11 shows that the average deviation from an equal distribution of the spoken words and speaking frequency among the participants decreased when the participants took turns to comment on the ideas (in OKMv2), suggesting that this might be a feasible procedure, although the same people as before still spoke the most (see Table 12).

<table>
<thead>
<tr>
<th>Method Phase</th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free discussion</td>
<td>Spoken words</td>
<td>7.3</td>
<td>6.4</td>
<td>9.4</td>
<td>4.3</td>
</tr>
<tr>
<td>Speaking frequency</td>
<td>6.0</td>
<td>7.2</td>
<td>10.0</td>
<td>4.1</td>
<td>4.6</td>
</tr>
</tbody>
</table>

All the participants thought that letting all the participants around the table take turns to speak was a good improvement. From observations it appeared that taking turns to comment on the ideas had helped to keep the discussion focused and had made it clear to the group what was going to happen next all the time, which had streamlined the time for interaction. The participants themselves jointly initiated a procedure whereby the sheet of paper containing suggestions presented by one participant was circulated among the other participants for them to comment on the ideas in turn. This made it clearer when the next person should start giving feedback and easier for the participants to avoid forgetting any comments which they might have thought of.
In Table 12 it can be observed that the supervisor spoke much less during GP1 and GP2 of OKMv1 and OKMv2 than during GP1 of the gallery method, indicating that the group worked more independently when applying the developed method. Although this is positive, since it suggests that the group can work through the method without a supervisor, it must be acknowledged that each member of the studied group has had the time to understand and learn the methods (despite the differences between them), as well as acquaint themselves better with the other group members. Therefore, at this stage one cannot conclude with certainty that a group with “beginners” would be able to accomplish similar results when applying OKMv2.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gallery GP1</th>
<th>OKMv1 GP1</th>
<th>OKMv1 GP2</th>
<th>OKMv2 GP1</th>
<th>OKMv2 GP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>12</td>
<td>19</td>
<td>9</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>C</td>
<td>16</td>
<td>6</td>
<td>9</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>16</td>
<td>23</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>E</td>
<td>23</td>
<td>26</td>
<td>29</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>F</td>
<td>26</td>
<td>23</td>
<td>22</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Supervisor</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

4.4 Ideation outcomes

The group generated concepts that included physical solutions to the problem of the ideation topic, organizational solutions and maintenance solutions. The concepts were described by words, a sketch, or by a combination of a sketch and words. Examples of concepts from each ideation session are shown in Figure 4 and Figure 5, respectively.

**Fig. 4** Examples of concepts from the OKMv1 ideation session on how to adjust the position of a modular turnout
In all the examples of concepts and ideas presented throughout this paper, the text or parts of the text have been translated from Swedish into English. Examples from the tests of Method 635 and the gallery method are to be found in Petersson et al. (2016).

From Table 13 it is obvious that there is a great difference in the idea generation rate between the individual and group phases. During IP1 and IP2, where the participants work individually, significantly more ideas were generated than during the group phases, as can be seen in the far right column in Table 13. It might be tempting to draw the immediate conclusion that, to maximize the total idea generation rate, only individual phases should be included. However, as was observed during IP1 of OKMv2, the participants eventually ran out of ideas. Although some ideas were expressed verbally, it can be seen from Table 14 that most of the spoken words were dedicated to the elaboration of ideas during all the group phases. The longer time the participants had to ideate individually, the longer was the time it took to present, explain and discuss the ideas in the subsequent step. After the group phases, the intention was that the participants would be able to create associations with what had been presented and come up with new ideas. Although it is difficult to quantify the extent of inspiration, it was found that the ideas that had turned up in IP2 in many cases had been based on some issue that had been brought up during GP1. The average concept generation rate was lower in IP2 than in IP1, but the average number of ideas per concept was higher in IP2, reflecting the fact that several of the participants aimed to combine ideas from previous phases with new ideas to achieve a complete solution to the problem of the ideation topic (as opposed to producing several unrelated ideas). An example is shown in Figure 6. The fraction of ideas from earlier phases that each participant reused in concepts generated in IP2 varied from 0 to 100%, but was on average 57% for OKMv1 and 54% for OKMv2, i.e. more than half. It is especially interesting that, although the participants had difficulties creating any more ideas during the last rounds of IP1 of OKMv2, despite exchanging ideas through rotational viewing, after the verbal interaction in GP1 they could squeeze out 32 new ideas in IP2.

Fig. 5 Examples of concepts from the OKMv2 ideation session on drainage in turnouts

In all the examples of concepts and ideas presented throughout this paper, the text or parts of the text have been translated from Swedish into English. Examples from the tests of Method 635 and the gallery method are to be found in Petersson et al. (2016).

From Table 13 it is obvious that there is a great difference in the idea generation rate between the individual and group phases. During IP1 and IP2, where the participants work individually, significantly more ideas were generated than during the group phases, as can be seen in the far right column in Table 13. It might be tempting to draw the immediate conclusion that, to maximize the total idea generation rate, only individual phases should be included. However, as was observed during IP1 of OKMv2, the participants eventually ran out of ideas. Although some ideas were expressed verbally, it can be seen from Table 14 that most of the spoken words were dedicated to the elaboration of ideas during all the group phases. The longer time the participants had to ideate individually, the longer was the time it took to present, explain and discuss the ideas in the subsequent step. After the group phases, the intention was that the participants would be able to create associations with what had been presented and come up with new ideas. Although it is difficult to quantify the extent of inspiration, it was found that the ideas that had turned up in IP2 in many cases had been based on some issue that had been brought up during GP1. The average concept generation rate was lower in IP2 than in IP1, but the average number of ideas per concept was higher in IP2, reflecting the fact that several of the participants aimed to combine ideas from previous phases with new ideas to achieve a complete solution to the problem of the ideation topic (as opposed to producing several unrelated ideas). An example is shown in Figure 6. The fraction of ideas from earlier phases that each participant reused in concepts generated in IP2 varied from 0 to 100%, but was on average 57% for OKMv1 and 54% for OKMv2, i.e. more than half. It is especially interesting that, although the participants had difficulties creating any more ideas during the last rounds of IP1 of OKMv2, despite exchanging ideas through rotational viewing, after the verbal interaction in GP1 they could squeeze out 32 new ideas in IP2.
From these results it is concluded that the verbally interactive steps including the presentation and discussion of ideas are relevant in the method. Kohn et al. (2011) found that interactive brainstorming groups generated combinations of previously generated ideas of higher utility than nominal groups. Nijstad et al. (1999) found that when groups brainstorm verbally without a time constraint, the productivity loss in relation to the productivity of nominal groups decreases, because the groups continue to ideate for a longer time. These findings indicate that verbal group interaction may be beneficial in group ideation if it is applied in the right instance, and may perhaps be especially beneficial in the longer ideation sessions that may be required in the field of engineering design. This may be especially fruitful when the absolute number of ideas is more important than the idea rate. One would expect that industrial companies would rather use some extra hours during ideation to assure that as many ideas as possible would be collected before starting the selection process. These extra hours would be much fewer than the total number of hours that would be used to develop a new product. Although the idea generation rate is an interesting parameter for evaluation purposes, the total number of ideas in a day or even a week is likely to be more relevant to companies and other organizations. It might therefore be more interesting to focus on how to make a group create more ideas on a topic which they think they have exhausted than mainly to consider the idea generation rate.

Table 13 Number of concepts and ideas, length of time and idea generation rate for each step of each method

<table>
<thead>
<tr>
<th>Method</th>
<th>Phase</th>
<th>Number of concepts</th>
<th>Number of new ideas</th>
<th>Average number of ideas/ concept</th>
<th>Time (min.)</th>
<th>Concept generation rate (concepts/min.)</th>
<th>New idea generation rate (ideas/min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>IP1</td>
<td>48</td>
<td>125</td>
<td>3.44</td>
<td>30</td>
<td>1.60</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>48</td>
<td>125</td>
<td>3.44</td>
<td>30</td>
<td>1.60</td>
<td>4.17</td>
</tr>
<tr>
<td>Gallery</td>
<td>IP1</td>
<td>26</td>
<td>66</td>
<td>2.92</td>
<td>15</td>
<td>1.73</td>
<td>4.40</td>
</tr>
<tr>
<td></td>
<td>GP1</td>
<td>-</td>
<td>15</td>
<td>-</td>
<td>25</td>
<td>-</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>IP2</td>
<td>3</td>
<td>5</td>
<td>NA</td>
<td>5</td>
<td>0.60</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>29</td>
<td>86</td>
<td>NA</td>
<td>45</td>
<td>0.64</td>
<td>1.91</td>
</tr>
<tr>
<td>OKMv1</td>
<td>IP1</td>
<td>41</td>
<td>100</td>
<td>3.31</td>
<td>30</td>
<td>1.20</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>GP1</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>37</td>
<td>-</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>IP2</td>
<td>10</td>
<td>22</td>
<td>5.80</td>
<td>10</td>
<td>1.0</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>2</td>
<td>10</td>
<td>2.5</td>
<td>33</td>
<td>0.06</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>53</td>
<td>144</td>
<td>3.85</td>
<td>110</td>
<td>0.48</td>
<td>1.31</td>
</tr>
<tr>
<td>OKMv2</td>
<td>IP1</td>
<td>77</td>
<td>183</td>
<td>3.05</td>
<td>63</td>
<td>1.22</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>GP1</td>
<td>-</td>
<td>17</td>
<td>-</td>
<td>54</td>
<td>-</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>IP2</td>
<td>10</td>
<td>32</td>
<td>7.80</td>
<td>10</td>
<td>1</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>GP2</td>
<td>-</td>
<td>15</td>
<td>-</td>
<td>31</td>
<td>-</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>87</td>
<td>247</td>
<td>3.60</td>
<td>158</td>
<td>0.55</td>
<td>1.56</td>
</tr>
</tbody>
</table>

*“New ideas” means those ideas which were presented for the first time in the ideation session in the step in question.

*Here “ideas” means all ideas, including ideas from earlier steps which have been reused and/or ideas suggested in other concepts.

*Not applicable (NA), because the participants continued working with their sheet of paper from IP1, while in OKMv1 and OKMv2 they received new sheets.

Table 14 Distribution of spoken words between activities (%)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Gallery</th>
<th>OKMv1</th>
<th>OKMv1</th>
<th>OKMv2</th>
<th>OKMv2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Going off at a tangent</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Elaboration</td>
<td>80</td>
<td>78</td>
<td>80</td>
<td>74</td>
<td>82</td>
</tr>
<tr>
<td>Ideas</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Criticism</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
It was expected that the developed method would outperform the other methods with regard to the quantity of ideas generated, and this expectation was fulfilled if one considers the total number of ideas. However, one can question whether the novel method outperformed the other methods concerning the overall concept and idea generation rate. The generation rate for new ideas during IP1 tended to decrease as the length of this step increased, which is consistent with previous research on verbal brainstorming showing that the idea flow declines with time (e.g. Paulus and Dzindolet 1993; Byron 2012). However, although the idea generation rate decreased somewhat, not least due to the minutes with less activity during the last rounds of IP1, the second version of the novel method had a surprisingly good idea generation rate during this step, considering the fact that the time was doubled compared to that allocated for the first version. All of the participants thought that it was better to spend some extra time, despite a lower idea generation rate, to achieve a greater absolute number of ideas.

Method 635 and the similarly executed IP1 of OKMv1 were of equal length, but the former generated more ideas. Several parameters can explain this result, but the fact that two of the six participants experienced a “bad day” when testing OKMv1 and were unhappy with their own effort probably contributed. Nevertheless, the group was very satisfied with the results of the workshop, especially with the breadth of the ideas generated during the workshop. Therefore, another explanation may be that the scope of the topics dealt with in Method 635 and OKMv1 differed somewhat in width, despite efforts to find relevant topics of a similar scope, and that in relation to the topic the performance of IP1 of OKMv1 was good.

When testing OKMv1, the participants felt that they had not had enough time to follow through on their own ideas created during the first round, but had been interrupted by the ideas on the sheet of paper given to them by their neighbour after 5 minutes. Linsey et al. (2011) hypothesized that “rotational viewing encourages the participants to spend more time understanding other [teammates’] ideas”, which is in line with the experience from the present study. Attending to the ideas of others is a critical factor for the observation of stimulation effects in groups (Paulus et al. 2015). The first ideation phase of the developed method actually consist of a) attending to the content of the sheet of paper and b) generating new ideas. To allow the participants both to comprehend the existing ideas on the sheet of paper received and to generate more ideas, the length of time for each round was successfully increased to ten minutes. However, it is important not to prolong the duration of each round unnecessarily, since previous research has shown that individuals under time pressure work at a faster rate and that the less time there is available, the greater is the task focus shown by groups; however, the quality of the performance may suffer if too high a pressure is applied, since information is processed less thoroughly then and a narrower range of alternatives are considered (Karau and Kelly 1992). Too long a time might also allow the participants to start thinking about disadvantages of their ideas and decide not to share them with the group. It was concluded that ten minutes was the right length of time for the topics considered in the present study. Linsey et al. (2011) used ten minutes for individual idea generation prior to four subsequent 7.5 minute periods where each group of five participants viewed either all the ideas (gallery viewing) or a new subset of ideas during each period.
The total time of 40 minutes in the study performed by Linsey et al. (2011) is shorter than the time used for IP1 in the present study, but longer than the total ideation time reported for the typical study in social psychology literature, which does not deal with design problems. An intuitive assumption is that the appropriate total and cycle times will depend on the ideation topic and the size and experience of the group concerned, and developing guidelines for setting these times as a function of these parameters is an interesting area for future research.

4.5 Formalization of learning and limitations

The present study is one of several efforts to understand better how ideation methods can be utilized within the area of engineering design, and the study has focused especially on a real-world context with ideation topics decided by the participants themselves. Besides the quantitative outcomes from the ideation sessions, the present study has considered the opinions of the participants in detail, which is unusual in research on ideation methods and can give important insights into the factors determining whether or not such methods will be implemented in industry. This study is limited to a single group and, therefore, statistical generalizations are not applicable. The findings and conclusions of the paper are valid for this particular group. Using another group might have resulted in different findings due to different group characteristics exerting different influences. Possible influencing factors in this connection are the personalities of the group members, their experience, backgrounds, gender, age and other personal characteristics. Further, the group acted in the railway sector, and studying a cross-functional inter-organizational group from another sector could result in other findings. Through the guidance of ADR, the findings have been generalized to design principles that can be used as a starting point for developing ideation methods in a similar situation that other researchers might face within other organizations. The last stage of ADR requires the researcher to make a conceptual move from the specific and unique to the generic and abstract on three levels (Sein et al. 2011). The present study has developed an instance-specific solution that has been reconceptualized into a class of solutions. The developed ideation method has been viewed as belonging to the class of ideation methods for cross-functional inter-organizational groups. Preliminary design principles for this class of solutions were formulated when designing the initial version of the developed method. These were confirmed during the tests and two additional design principles were identified during the test of the initial version of the method. Further, it was realized that the method had inherently encouraged the participants to attend to the other participants’ ideas and that this was an important design principle. The design principles are to be found in Table 9.

Although AMMs explain how information in the long-term memory is accessed, other factors exist that can further stimulate or interfere with activation, such as motivation. Stimulation requires that the subject concerned should attend to the stimulus idea (e.g. Dugosh et al. 2000), which is a process that requires the person to be motivated to do so. The motivation for a participant to create ideas or share the ideas created might be low if there is some reason why that participant might encounter negative consequences through doing so. One factor which could lower the motivation of a participant is a fear of other participants stealing good ideas from them due to the inclusion of competing actors in the group (Petersson et al. 2016). Almefelt and Claesson (2015) applied systematic design methods in a collaborative project including an automotive manufacturer and 35 automotive suppliers and found that, although the application of the methodology was in general successful, in some teams where the parties represented competing suppliers, the work progress was hindered by conflicts of interest. Therefore, no competitors (actors competing in the same functional segment) were involved in the creative team of the present study, but there is still a risk that some participants might have kept some ideas for themselves owing to a fear of having them spread outside the group.

Because only a single group participated in this study, we were forced to change the ideation topic between the ideation sessions. Although the different topics were chosen with care so that they would be equivalent, and although they were found to lead to similar engagement among the participants (see Figure 2), it cannot be excluded that they influenced the ideation outcomes differently. The number of non-redundant ideas and concepts was carefully counted according to a specified procedure to make the count as fair as possible between the methods. However, even if great care is taken to count ideas in a neutral way, subjectivity will to some extent influence the counting, and in the present case the different nature of the concepts made the identification of single ideas challenging. Because of these
uncertainties originating from the change of ideation topic and the counting of ideas, our aim has not been to make conclusions about which method is the “best” with respect to the idea generation rate, but rather to focus on interesting trends in the data that can be useful in helping us to understand how to apply ideation methods in an optimal way within the area of engineering design.

The tests of the methods were performed on different occasions. This was the only possible way to organize the tests, because a) the participants were busy professionals located in different parts of Sweden and b) the evaluation of the first test had to be completed before the second version could be tested. Therefore, the conditions of each test may have varied with respect to the mood and fitness of the participants, and this was especially evident during the test of OKMv1, where two participants explicitly mentioned that they had experienced a bad day with respect to creativity.

This study has focused on developing and testing a method suitable for application in a functionally heterogeneous group. However, we have not tested the method in a homogeneous group and compared the results. One of the starting points for developing the method was that participants with different functional backgrounds would be able to provide different types of ideas as a reaction to the ideation topic in the individual ideation phase at the start of the method, and that these ideas would be used as stimulus ideas for the other participants throughout the method. If the participants of a group have a similar background, competence and perspective with respect to the ideation topic, it is expected that the ideas that the group can produce at the start of the method will have less variety and hence not serve the purpose of functioning as stimulus ideas for the group. Consequently, the outcome is expected to be worse compared to that for a group that is heterogeneous with respect to the ideation topic. In the case of a very homogeneous group, external stimulus ideas might be a better approach. It has, however, not been within the scope of the present study to investigate this, but it is a very interesting area for future research.

Several researchers have found that studying team performance in work settings is difficult because it is hard to obtain access to such settings (Shah 1998; Paulus et al. 2015). However, studies of ideation methods in organizations and industry are important for understanding how the use of ideation methods in such contexts can increase the effectiveness of teams working with innovations (Shah et al. 2000; Paulus et al. 2015). From our experience, the company or organization must be convinced that they will profit from this type of study in order to provide the researchers with access. This implies that the researcher has to adapt to the situation and the on-going work at the company or organization, and that the main aim of the activity for which the researcher is given access is that it should be fruitful for the organization. Further, setting up a completely controlled experiment necessarily means procedures that would never take place in real life, and therefore the study necessarily becomes less realistic, although it still takes place in a work setting. Therefore, we found that it is critical to find a balance between letting the participants work in a realistic setting and exerting enough control of the experiment to be able to accomplish a meaningful evaluation. Performing studies of ideation methods in work settings will necessarily imply more uncertainties and a smaller sample than performing such studies in controlled experiments with students, and therefore statistical significance is a challenge. In our view, such studies are still worthwhile, since they can shed light on findings from statistical studies and indicate if these findings are relevant in real work settings, as well as identify relevant topics for future research.

5 FUTURE WORK

The present paper has presented the results of a study which has developed an ideation method within one cross-functional inter-organizational group. The next step is to make field tests where the method will be tested in real-life projects on real problems by participants who are new to the method.

5 CONCLUSIONS

The learning derived from the development of an ideation method in a single group was formalized into the following design principles for ideation methods to be used in cross-functional inter-organizational teams:

- developing the method based on methods proven to be easy to understand,
- the initial generation of a pool of ideas with great variety,
exposure to stimulus ideas generated internally within the group,
mixing idea-exchanging strategies,
the use of external representation in the form of sketches and text,
including time for discussion and debate,
optimizing the cycle time,
developing a strategy for distributing the verbal interaction,
paying attention to stimulus ideas.

The method developed according to these design principles was found by the participants in the group to be more useful and to generate more ideas that could be used in practice than Method 635, the gallery method and the SIL method.

During the tests of the developed method, the group generated concepts that included physical solutions to the problem of the ideation topic, organizational solutions and maintenance solutions. The concepts were described by words, a sketch, or by a combination of a sketch and words. Considerably more ideas were generated during the individual phases than during the group phases. Most of the spoken words were dedicated to the elaboration of ideas during all the group phases. The developed method outperformed the other methods with regard to the quantity of ideas generated (considering the total number of ideas), but not with regard to the overall concept and idea generation rates. All the participants thought that it was better to spend more time, despite a lower idea generation rate, to achieve a greater absolute number of ideas. Although the participants had difficulties creating any more ideas during the last rounds of the first individual ideation phase when trying the developed method, they could squeeze out 32 new ideas in a subsequent individual ideation phase after a group phase of verbal interaction. The fraction of ideas from earlier phases that each participant reused in concepts generated in a subsequent individual ideation phase after a group phase of verbal interaction varied from 0 to 100%, but was on average 57% for the first version of the developed method and 54% for the second version, i.e. more than half. The average concept generation rate was lower in the second individual ideation phase than in the first such phase, but the average number of ideas per concept was higher in the second such phase.

ACKNOWLEDGEMENTS

The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged. The OptiKrea partners Trafikverket (STA), Vossloh Nordic Switch Systems and Infranord are gratefully acknowledged for their contributions to the project. Special thanks are extended to the participants of the creative team.

REFERENCES


Petersson AM, Lundberg J (2016) Applying action design research (ADR) to develop concept generation and selection methods. Accepted to the 26th CIRP Design Conference
Paper IV
EVALUATING AN IDEATION METHOD IN A REAL-LIFE CONTEXT: A FIELD TEST
FROM THE RAILWAY SECTOR

Anna Malou Petersson Jan Lundberg
Division of Operation, Maintenance and Acoustics
Department of Civil, Environmental and Natural Resources Engineering
Luleå University of Technology
Luleå, Sweden

ABSTRACT
A great number of ideation methods are available to assist the engineer in the conceptual phase of product development. Previous research on idea generation suggests that, in order to understand how ideation can be successful in reality, the context must be taken into account. Therefore, evaluating promising ideation methods in the intended use setting should be an important complement to other research studies in the field of engineering design. In the present study, an ideation method which had previously been developed by the present authors in close collaboration with industry, a government agency, and academia was evaluated in a typical use setting. Testing the method in a use setting led to unexpected events and to new insights into the method. During the most ideation-intensive stage of the method, one of the groups in the use setting generated suggestions at a rate similar to that of the small representative group in which the development of the method had taken place, whereas the other three groups in the use setting had a lower suggestion generation rate. The findings indicated that the participants’ views on the method correlated with those of the group which had developed the method.

Keywords: ideation; ideation methods; real-life settings; field tests; railway

1 INTRODUCTION
A great number of ideation methods, proposed by both practitioners and researchers, are available to assist the engineer in the conceptual phase of product development. Ideation methods, and especially brainstorming [1], have been extensively studied in the social psychology literature, but although much can be learnt from these studies, it is impossible to transfer the results directly to the area of engineering design for at least three different reasons. Firstly, these studies let the participants use words to express ideas, whereas designers use sketches to communicate. Secondly, the typical psychology study involves topics that do not require domain-specific knowledge for the solution of problems, whereas engineers deal with problems whose solutions do require such knowledge. Thirdly, whereas these studies typically involved the participation of students, engineers engaged in product development typically work in cross-functional teams involving participants with complementary skills and knowledge. These differences imply that ideation methods have to be studied and developed in the context of engineering design to extract and provide guidance as to how engineers can make the best use of these methods to generate concepts. Linsey et al. [2] examined how two key factors influenced the outcome of an ideation session involving the design of a device to shell peanuts. These two factors were the ways in which a group’s ideas are displayed and the form of communication between group members. They found that over a 40 minute ideation session, 50% more ideas were generated with the most favourable combination of key factors than with the least favourable. Linsey and Becker [3] made a complementary study letting nominal groups ideate on the same problem given the same amount of time. They found that real groups develop a larger number of ideas than equivalent nominal groups in several combinations of the key factors [3]. These studies show that using ideation methods in an optimal way has a great potential to increase the number of ideas that a group can contribute during concept generation. This is particularly important since past research on brainstorming in the psychology field has found that the quantity and quality of ideas are highly correlated [4,5], and some supporting evidence for this finding has been obtained in the area of engineering design [2,6].
The OptiKrea Group consisted of six participants representing the infrastructure manager (IM), a turnout representative of the context which the method was developed in, a turnout manufacturer, a maintenance contractor and academics with domain-specific knowledge. Considering previous relevant research on ideation methods, associative memory models and the findings from tests of established ideation methods performed in the same group [15], a number of preliminary design principles were formulated and implemented in an ideation method that combines individual phases of rotational and gallery viewing with phases of verbal group interaction [14]. The method was thereafter refined through tests which were performed in the OptiKrea Group and which addressed issues concerning turnouts that the OptiKrea Group proposed. The instructions of the method are presented in Table 1.

### Table 1. Ideation method instructions

<table>
<thead>
<tr>
<th>General directions</th>
<th>Stage 1: Rotational viewing</th>
<th>Stage 2: Presentation of ideas and feedback from other participants</th>
<th>Stage 3: Gallery viewing</th>
<th>Stage 4: Presentation of ideas and feedback from other participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) The instructions herein presuppose that the group has understood and agreed upon the ideation topic.</td>
<td>a) During 10 minutes each participant comes up with at least three ideas on how to address the ideation topic. Each participant documents their ideas with sketches and/or text on a sheet of A3 paper.</td>
<td>a) The participants use 5 minutes to read through the ideas that have been added to the sheet of paper which they started out with.</td>
<td>a) The sheets of paper from Stage 1 are put up on a wall or some other place where all the participants can easily view them.</td>
<td>a) Each participant presents their own ideas from stage 3.</td>
</tr>
<tr>
<td>b) Maximize the number of ideas (old and new).</td>
<td>b) Each participant sends their sheet of paper to the neighbour on their left.</td>
<td>b) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.</td>
<td>b) Each participant works individually, developing or combining ideas from the collection of ideas from Stage 1. New ideas are also welcome.</td>
<td>b) After each presentation the presented sheet of A3 paper is circulated among the participants and each participant gives feedback on the ideas. The remaining available time is used for discussions.</td>
</tr>
<tr>
<td>c) Avoid negative criticism during ideation.</td>
<td>c) The neighbour reads through the ideas and adds at least three improvements, combinations of the ideas and/or new ideas on the sheet of paper during 10 minutes. One is allowed to ask the neighbour on one’s right what they meant by an idea that is incomprehensible.</td>
<td>c) After each presentation, the presented sheet is circulated among the participants and each participant gives feedback on the ideas. The remaining available time is used for discussions.</td>
<td>c) The aim is still to collect as many ideas as possible!</td>
<td>c) A maximum of 5 min. is allowed per participant (for presentation and feedback).</td>
</tr>
<tr>
<td>d) Appoint one person to be responsible for keeping track of the time with the help of a timer.</td>
<td>d) A maximum of 10 min. is allowed per sheet of paper (for presentation and feedback).</td>
<td>d) Appoint one person to be responsible for keeping track of the time with the help of a timer.</td>
<td>d) The sheets pass between all the participants (i.e. Stage 1 ends when one receives the sheet of paper which one started out with).</td>
<td>d) A maximum of 5 min. is allowed per participant (for presentation and feedback).</td>
</tr>
</tbody>
</table>

Viswanathan and Linsey [7] found that design methods can have a different impact on novices and experts. Studies performed on ideation methods have, however, mainly involved inexperienced users, predominantly students, working under experimental conditions, although some studies of professionals working under experimental conditions do exist (e.g. [8,9]). A few studies of professionals using ideation methods in a real-life context exist (e.g. [10,11]). Jackson and Poole [11] found that idea generation in groups in naturally occurring contexts is more complex than typically imagined. The idea generation rate per time unit or member was considerably lower than that in laboratory studies.

Previous research on idea generation [10-12] suggests that, in order to understand how ideation can be successful in reality, the context must be taken into account. Many researchers believe that the use of systematic conceptual design methods in industry is limited (e.g. Shah [13]). Gish and Hansen [12] showed that contextual factors can be important in idea work and one reason for methods not being adopted in industry could be that they are not adapted to the context. Therefore, evaluating promising ideation methods in the intended context should be an important complement to other research studies in the field of engineering design research.

In the present study, the aim was to evaluate, in a typical use setting, an ideation method which had previously been developed by the present authors in close collaboration with industry, a government agency and academia [14]. The objectives were to explore how the method works in a typical use setting, and to judge if the findings from the development of the method in a small group can be generalized to other groups in similar contexts.

### 2 BACKGROUND

The ideation method tested in the present study had previously been developed within the scope of a research project called OptiKrea and run by Luleå Railway Research Center at Luleå University of Technology. The OptiKrea project was initiated with the intention of promoting the technical development of turnouts, especially from the point of view of designers. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining, utilizing and researching the railway. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions to the problem would be found. A parallel goal of the project was to develop an ideation method that combines individual phases of rotational and gallery viewing with phases of verbal group interaction [14]. The method was thereafter refined through tests which were performed in the OptiKrea Group and which addressed issues concerning turnouts that the OptiKrea Group proposed. The instructions of the method are presented in Table 1.

The ideation method tested in the present study had previously been developed within the scope of a research project called OptiKrea and run by Luleå Railway Research Center at Luleå University of Technology. The OptiKrea project was initiated with the intention of promoting the technical development of turnouts, especially from the point of view of designers. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining, utilizing and researching the railway. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions to the problem would be found. A parallel goal of the project was to develop an ideation method that combines individual phases of rotational and gallery viewing with phases of verbal group interaction [14]. The method was thereafter refined through tests which were performed in the OptiKrea Group and which addressed issues concerning turnouts that the OptiKrea Group proposed. The instructions of the method are presented in Table 1.

### Table 1. Ideation method instructions

<table>
<thead>
<tr>
<th>General directions</th>
<th>Stage 1: Rotational viewing</th>
<th>Stage 2: Presentation of ideas and feedback from other participants</th>
<th>Stage 3: Gallery viewing</th>
<th>Stage 4: Presentation of ideas and feedback from other participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) The instructions herein presuppose that the group has understood and agreed upon the ideation topic.</td>
<td>a) During 10 minutes each participant comes up with at least three ideas on how to address the ideation topic. Each participant documents their ideas with sketches and/or text on a sheet of A3 paper.</td>
<td>a) The participants use 5 minutes to read through the ideas that have been added to the sheet of paper which they started out with.</td>
<td>a) The sheets of paper from Stage 1 are put up on a wall or some other place where all the participants can easily view them.</td>
<td>a) Each participant presents their own ideas from stage 3.</td>
</tr>
<tr>
<td>b) Maximize the number of ideas (old and new).</td>
<td>b) Each participant sends their sheet of paper to the neighbour on their left.</td>
<td>b) Each person presents the ideas on the sheet of paper which they started out with and, if necessary, the other participants help to explain anything which the presenter has not been able to understand.</td>
<td>b) Each participant works individually, developing or combining ideas from the collection of ideas from Stage 1. New ideas are also welcome.</td>
<td>b) After each presentation the presented sheet of A3 paper is circulated among the participants and each participant gives feedback on the ideas. The remaining available time is used for discussions.</td>
</tr>
<tr>
<td>c) Avoid negative criticism during ideation.</td>
<td>c) The neighbour reads through the ideas and adds at least three improvements, combinations of the ideas and/or new ideas on the sheet of paper during 10 minutes. One is allowed to ask the neighbour on one’s right what they meant by an idea that is incomprehensible.</td>
<td>c) After each presentation, the presented sheet is circulated among the participants and each participant gives feedback on the ideas. The remaining available time is used for discussions.</td>
<td>c) The aim is still to collect as many ideas as possible!</td>
<td>c) A maximum of 5 min. is allowed per participant (for presentation and feedback).</td>
</tr>
<tr>
<td>d) Appoint one person to be responsible for keeping track of the time with the help of a timer.</td>
<td>d) A maximum of 10 min. is allowed per sheet of paper (for presentation and feedback).</td>
<td>d) Appoint one person to be responsible for keeping track of the time with the help of a timer.</td>
<td>d) The sheets pass between all the participants (i.e. Stage 1 ends when one receives the sheet of paper which one started out with).</td>
<td>d) A maximum of 5 min. is allowed per participant (for presentation and feedback).</td>
</tr>
</tbody>
</table>
The participants in the OptiKrea Group were satisfied with the results and the final version of the method, and the number of ideas generated when applying the method was satisfactory [14]. However, the act of assisting in the development of the method might have biased the team’s satisfaction with it and it was deemed desirable to evaluate the method in other groups in a similar context. Two participants in the OptiKrea project took part in a work package in an EU project called In2Rail. In2Rail aims to create the technology necessary to complete the Single European Railway Area. The first author of this paper was invited to conduct a two day workshop where 17 In2Rail project participants would apply the developed ideation method to four tasks that were part of the work package. One of the aims set for the development of the ideation method was that the method should be easy to use for individuals from different organizations. Therefore, it was expected that the method would be easy to apply during the In2Rail workshop. Further, it was expected that the participants of the In2Rail workshop would share the views of the OptiKrea Group on the method, and that they would generate a comparable number of suggestions.

3 METHOD

3.1 Participants and groups

A total of 17 people participated in the In2Rail workshop. On the first day, 16 people took part in the workshop and on the second day the number of participants was 15. Fourteen people took part on both days and three people only during one day. The participants came from six different European countries and represented five different IMs, four different universities and four different companies. The participants were asked to answer a number of questions on a questionnaire regarding their background and prior experience of ideation methods, and the results of the questionnaire are shown in Table 2. The ten participants who had used ideation methods previously were asked what ideation methods they had experience of. Apart from the two people who had been members of the OptiKrea Group, none of the participants had tried any other method that resembles the developed method. During the first day, the participants were assigned by the work package leader to two different groups according to what task group they belonged to in the In2Rail project. Some participants took part in both task groups, and then they were assigned to the group considering the topic which they were judged to have most knowledge about. During the second day, the participants were split into two new groups with the aim of reaching a good balance between the participating organizations in each group. This assignment to different groups was carried out by one of the task leaders, taking the participants’ opinions into consideration. Table 3 shows the details of each In2Rail group, as well as the details of the OptiKrea Group for comparison. Two participants in the OptiKrea Group also participated in the In2Rail project. They were assigned to different groups throughout the workshop, so that each of the four In2Rail groups had one participant from the OptiKrea Group.

<table>
<thead>
<tr>
<th>Table 2. Participants’ replies to questions about their background and prior experience of ideation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What type of organization do you represent?</td>
</tr>
<tr>
<td>IM</td>
</tr>
<tr>
<td>Supplier</td>
</tr>
<tr>
<td>Academia</td>
</tr>
<tr>
<td>Another organization</td>
</tr>
<tr>
<td>2. How many years of experience do you have from the railway sector?</td>
</tr>
<tr>
<td>0-5</td>
</tr>
<tr>
<td>5-10</td>
</tr>
<tr>
<td>10-15</td>
</tr>
<tr>
<td>15-20</td>
</tr>
<tr>
<td>20-25</td>
</tr>
<tr>
<td>&gt;25</td>
</tr>
<tr>
<td>3. Have you used ideation methods before?</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>4. If you answered “Yes” to Question 3, how often have you used ideation methods?</td>
</tr>
<tr>
<td>I have used them a few times</td>
</tr>
<tr>
<td>A few times per year</td>
</tr>
<tr>
<td>Every month</td>
</tr>
<tr>
<td>Every week</td>
</tr>
<tr>
<td>Other frequency</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. Details of the groups (OKG=OptiKrea Group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Number of participants representing</td>
</tr>
<tr>
<td>IM</td>
</tr>
<tr>
<td>Supplier</td>
</tr>
<tr>
<td>Academia</td>
</tr>
<tr>
<td>Other organizations</td>
</tr>
<tr>
<td>Number of organizations represented</td>
</tr>
<tr>
<td>Number of participants</td>
</tr>
</tbody>
</table>

*One participant in Group 2 arrived 7.5 minutes after the group had started applying the ideation method.

3.2 Procedure

The workshop was planned in accordance with the prerequisites given by the leaders of the work package. Each of the two groups selected on the first day worked on a different ideation topic. Each of the two groups selected on the second day worked on both the topics chosen for that day, and the reason for this was that the two topics had points in common and synergies could be expected. The ideation topics are
presented in Table 4. According to the plan drawn up, the different task groups in the In2Rail project were supposed to prepare the ideation session related to their task by executing a topic analysis and goal-setting prior to the workshop, and each day was to begin with a presentation of the outcome of the preparation. Thereafter, the ideation method was to be applied to each topic on the first and second day. On each day, after the ideation method had been executed, each group was to organize their ideas into different categories and finally present these to the other group. Before the first ideation session, the ideation method instructions were shown by means of a projector and explained to the participants for about ten minutes. The instructions had also been sent to the participants prior to the workshop. The participants were told that only six rounds would be performed in Stage 1. Group 2 and 4 were supervised by one of their participants, who had taken part in the OptiKrea project and was familiar with the method. Group 1 and 3 were supervised by the first author, who did not take part in the ideation, but gave instructions, when necessary, and answered questions. Each group worked in a separate meeting room, and was provided with sheets of A3 paper, tape and pairs of scissors. Each ideation session was audio-recorded.

### Table 4. Ideation topics

<table>
<thead>
<tr>
<th>Group</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Novel turnout locking mechanism</td>
</tr>
<tr>
<td>2</td>
<td>Embedded and integrated sensors</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>Radical turnout concept</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>Self-inspecting/correcting/adjusting turnout</td>
</tr>
</tbody>
</table>

### 3.3 Views of the participants

After the ideation method had been applied each day, each participant answered a questionnaire presenting different statements about the method. This was the same questionnaire as had been applied in the OptiKrea Group. The participants answered on a continuous scale from “Do not agree at all” to “Agree completely”, by making a mark on a 100 millimetres long line. The position of the mark was measured in millimetres from “Do not agree at all” with a ruler and divided by the total length of the line. We calculated the average group value for each statement and the average values for the subgroups IM, suppliers, academia and other organizations. It was known that competitors in different market segments participated in the EU project. Therefore, in addition to the statements from the original questionnaire applied in the OptiKrea Group, a question was posed which asked the participants whether they had kept ideas to themselves due to the presence of competitors, since the participants in the OptiKrea Group had mentioned that they would have kept ideas to themselves if competitors had been present [15]. This effect had been observed in a parallel project conducted by the second author of the present paper and has also been reported by other researchers [16]. The participants were able to give written comments on the method in free text at the end of the questionnaire.

### 3.4 Outcome of the ideation sessions

The sheets of paper that the participants had used for documentation while applying the ideation method were copied and analysed. The sheets had contents of different types. An inductive approach was used in developing guidelines for categorizing the content. The entire content was parsed several times. After the first round, the categories of content were tentatively defined. In the second round, a check was performed as to how the content fitted into these categories and the categories were more clearly defined. New categories were defined, as some content did not fit into the original categories. The third round involved formulating guidelines as to what category a certain content should belong to, and a final check was performed to ensure that all the contents belonged to a category. The guidelines are shown in Table 5. The number of pieces of content in each category generated by each group was counted.

### Table 5. Guidelines for counting pieces of content

1. As a starting point, the pieces of content are identified on the sheets as text and sketches that belong together and are marked off from other text and sketches by the participants.

2. The pieces of content belong to one of three main categories:
   - A. Topic analysis and requirements: This category includes content that only treats issues related to the background of the topic, goals/objectives and requirements for the solution.
   - B. Suggestions on how to address the ideation topic: Each suggestion constitutes a solution or an approach to a solution of the problem or a subproblem of the ideation topic on its own.
   - C. Irrelevant content: This type of content does not have any relevance to the ideation topic.

3. Suggestions (main category B) are categorized into three subcategories:
   - a. Physical suggestions: Physical solutions suggested to address the ideation topic.
   - b. Organizational suggestions: Suggestions attempting to address the ideation topic by making some change in the organization which the ideation topic belongs to.
   - c. Data management suggestions: Suggestions attempting to address the ideation topic by making some change concerning how data are managed.

4. Suggestions (main category B) may be a mix of physical, organizational and data management suggestions. If this is the case, the major contribution in the suggestion determines what category it should belong to.

5. Physical suggestions are either abstract or concrete suggestions. If the suggestion has no information on how to realize the solution, it is an abstract suggestion. If there is any information on how to realize the solution, it is a concrete suggestion.

6. Some suggestions mention alternatives. If the alternative is a variant of a minor function or detail that does not change the principal idea of the solution, the alternative belongs to the original suggestion. Otherwise it is counted as a separate suggestion.

7. The output of these suggestions are classified which frame the problem and, although they do not specifically address the ideation topic as described, still meet the higher level need.

8. Pieces of content that include the same information are counted only once.
Group 2 was supposed to ideate on embedded sensor technology, but reframed the topic to include the properties which they considered important to measure and any type of method for measuring them. Consequently, a considerable part of the documented content had the character of a list. Two additional subcategories had to be used for the content generated by Group 2, i.e. two subcategories in addition to the subcategories needed to categorize all the pieces of content generated by the other groups. All the pieces of content generated by Group 2 that listed properties and measurement techniques were counted in an additional subcategory. Several concepts had the character of a list, but had additional information such as implications, questions and objections. These were counted separately. The total number of properties and measurement techniques was also counted. Thereafter, the remaining content was categorized according to Table 5.

4 RESULTS AND DISCUSSION

4.1 Deviations from the method instructions

Most of the participants had not contributed to the topic analysis and goal-setting stages in advance. Therefore, more time than planned had to be dedicated to the common introduction during the workshop and, consequently, less time was available to perform the ideation sessions. This led to the following deviations from the method instructions (Table 1).

- Group 1 performed only Stage 1 and 2.
- Group 2 combined Stage 3 and 4 with categorization. The participants used Stage 3 to summarize, based on Stage 1 and 2, what they had found to be most promising and/or important. During the presentation of the summaries during Stage 4, the group simultaneously categorized the presented content.
- Group 3 and 4 completed five rounds during Stage 1.
- Group 4 set a maximum limit of five minutes per participant for presentation and feedback during Stage 2.
- Group 3 and 4 performed Stage 3 and 4 as one group. Stage 3 and 4 were used as they were used by Group 2 during Day 1. The participants were told, however, that they could also add any new ideas which they thought of.

The ideal time consumption (as specified by the instructions) and the actual time consumption of each stage in each group are shown in Table 6. Between the different stages, there was typically a short break while the participants stretched their legs or fetched a cup of coffee. A few times, participants also left the group during a stage, e.g. to take a phone call. On average, Group 1, 2 and 3 used a little less time than the ideal time for presentation and feedback during Stage 2. However, the participants presenting suggestions at the beginning of Stage 2 in these groups used more than ten minutes. As the remaining participants realized that the group was running out of time, many of them chose to present their ideas very quickly and not to give any comments or feedback on other participants’ suggestions. Therefore, the participants presenting suggestions at the end of Stage 2 had less time at their disposal than those presenting suggestions at the start, and unfortunately this means that interesting suggestions and viewpoints might not have been presented to the group. Very little time was available for Stage 4 in Group 3 and 4, and consequently the presentation assumed more the character of a discussion. Nevertheless, the group was able to classify its suggestions into several categories during the approximate time of 20 minutes which was available.

The deviations from the method instructions gave us additional insights. The design of the method allowed it to be adapted to a specific situation by easy adjustments. An advantage of the method from this point of view is that it is built of stages that can be viewed as different modules. By removing modules, changing the number of rounds or the time per round, the method can rapidly be adjusted to the current situation. We also found that it was an advantage that the method has a “pen-and-paper” approach, since the facilities did not allow the use of technical equipment.

The deviations from the method instructions were used in another way than intended. Instead of generating new suggestions, the majority of the participants in the In2Rail workshop chose to use the time for summarizing the content of Stage 1 and 2, and either put together a “complete” solution from different suggestions or listed their favourite suggestions. From an ideation perspective
these stages appeared almost useless, since very few participants contributed new suggestions. Interestingly, the participants did not make any comments regarding the possibility of excluding Stage 3 and 4 from the method, either verbally or in the questionnaire. It appears that they found these stages useful despite the lack of new suggestions and used them as a bridge from ideation to categorization of suggestions.

One important goal of the method design was that it should be possible for participants who are unfamiliar with the method, or new to ideation methods in general, to use the method directly. The participants were introduced to the method instructions shown in Table 1 for about ten minutes and were then asked to start the ideation session. They received further short directions from the person supervising each group before starting the ideation session and between stages when necessary. Although the participants were slightly hesitant at the start of the ideation sessions during the first day, they were able to apply the method directly to the ideation topics. There were few questions about Stage 1 and 2 of the method, whereas Stage 3 appeared to be more confusing to the participants, as in that stage they asked more questions about what they were supposed to do. A more useful approach would probably be to give paper copies of the instructions to every participant, as this would remove the need for remembering the instructions shown before the ideation session.

4.2 Questionnaire
Figure 1 shows the questionnaire results from the In2Rail workshop compared to those from the OptiKrea Group. Figure 2 shows the questionnaire results from the In2Rail workshop according to the types of organization which the respondents represented. Table 7 shows the statements included in the questionnaire. Overall it can be concluded from the questionnaire that the In2Rail project participants were satisfied with using the method. Figure 3 shows that there is a correlation between the extent to which the OptiKrea participants and the extent to which the In2Rail participants agreed on average with the statements in the questionnaire. This is an indication that the views of the participants in the OptiKrea project can be generalized to other cross-functional inter-organizational groups within the railway sector.

In Figure 2, it is apparent that the supplier representatives kept more ideas to themselves than the other subgroups. This can, however, be misleading, since only two participants reported that they represented suppliers. Of these two, one agreed to a great extent with the statement, pulling up the average, whereas the other did not award a conspicuous score for this statement. On the other hand, one of the participants representing an IM awarded a conspicuously high score for this statement, and two other IM representatives awarded a fairly high score compared to the other participants. It was not expected that the IM representatives would keep ideas to themselves, and rather the opposite was expected. One reason could be that these individuals had had novel ideas presented to them by manufacturers, for example, and had agreed to be confident about these ideas. The supplier representatives on average agreed to a much greater extent than the other subgroups with the statement that they would have come up with more ideas if they had had more time, and both the participants representing a supplier agreed to a similar degree with that statement.

4.3 Ideation outcomes
Since Stage 3 and 4 were used for summarizing and categorizing the outcome of Stage 1 and 2, rather than generating new suggestions, this section concerns the content from Stage 1.

Table 8 shows how the pieces of content generated by each group are distributed between the main categories specified in Table 5. Irrelevant content only appeared once. The conclusion can therefore be drawn that the participants focused on the ideation topic at hand. Some content concerned topic analysis, including the discussion of goals and requirements for the solution to be developed. This content is relevant to the ideation topic, but should ideally have been treated in the topic analysis phase before the ideation started. Therefore, this content is a sign that the topic was not adequately clarified and the goals of the workshop were not adequately specified before the ideation started, which is also reflected in the fact that most participants had not contributed to the topic analysis and it was not possible to assign enough time for a thorough problem clarification. Some content concerned topic analysis, including the discussion of goals and requirements for the ideation topic at hand. Some content concerned the solution to be developed. This content is relevant to the ideation topic, but should ideally have been treated in the topic analysis phase before the ideation started. Therefore, this content is a sign that the topic was not adequately clarified and the goals of the workshop were not adequately specified before the ideation started, which is also reflected in the fact that most participants had not contributed to the topic analysis and it was not possible to assign enough time for a thorough problem clarification.

Group 1 worked with novel mechanisms for locking turnouts, and dealt with the three functions of actuation, locking and detection. Table 9 shows that most suggestions concerned locking and detection, and only one suggestion addressed actuation. Six suggestions involved the performance of two or all of the three functions with the same feature. During the discussion of the topic before ideation, the group focused on a frequent problem involving a stop signal being triggered at a turnout without any apparent fault having occurred. It seems reasonable to surmise that this led the group to focus on ways to lock the turnout and detect whether it is in the correct configuration, in order to avoid traffic disturbances, and this might explain why actuation received less attention. Another possible reason could be that the participants in the group had more knowledge in the area of locking and detection.

Table 10 shows the distribution of the suggestions generated by Group 2 between the subcategories. In total, 178 different measurable properties and 34 measurement techniques were given by Group 2 (dispersed among all the suggestions).
Group 3 and Group 4 worked on the same topics during the same length of time in Stage 3 and 4. From Table 8 it can be seen that Group 3 generated 40 suggestions, while Group 4 generated 63 suggestions, despite the fact that Group 3 had eight participants and Group 4 had seven participants. In Table 11 the distribution of these suggestions between the two tasks and the different subcategories and groups is shown. It can be observed that Group 4 generated a considerably higher amount
Figure 3. Linear fit ($R^2$=0.84) between the extent to which the OptiKrea participants (average of the group) and the extent to which the In2Rail participants (average of the groups) agreed with each statement in the questionnaire

Table 8. Number of pieces of content from Stage 1 in the different main categories according to group

<table>
<thead>
<tr>
<th>Category</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic analysis and requirements</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Suggestions on how to address the ideation topic</td>
<td>35</td>
<td>43</td>
<td>40</td>
<td>63</td>
</tr>
<tr>
<td>Irrelevant content</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Suggestions with the character of a list are not included here.

Table 9. Number of suggestions generated by Group 1 in Stage 1, according to subcategory and function

<table>
<thead>
<tr>
<th>Function</th>
<th>Subcategory</th>
<th>Physical (concrete*)</th>
<th>Organizational</th>
<th>Data management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actuation</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Locking</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Shared</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*No abstract physical suggestions were generated by Group 1.

Table 10. Number of suggestions generated by Group 2 during Stage 1, according to subcategory

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggestions in list-form</td>
<td>21</td>
</tr>
<tr>
<td>Suggestions in list-form with additional information such as implications, questions and objections</td>
<td>19</td>
</tr>
<tr>
<td>Physical suggestions</td>
<td>33</td>
</tr>
<tr>
<td>abstract</td>
<td>1</td>
</tr>
<tr>
<td>concrete</td>
<td>32</td>
</tr>
<tr>
<td>Organizational suggestions</td>
<td>4</td>
</tr>
<tr>
<td>Data management suggestions</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 11. Number of suggestions generated by Group 3 & 4 in each subcategory according to group and ideation topic

<table>
<thead>
<tr>
<th>Subcategory</th>
<th>Radical design</th>
<th>Self-adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 3</td>
<td>Group 4</td>
</tr>
<tr>
<td>Physical suggestions</td>
<td>21</td>
<td>49</td>
</tr>
<tr>
<td>abstract</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>concrete</td>
<td>17</td>
<td>47</td>
</tr>
<tr>
<td>Organizational suggestions</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Data management suggestions</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>53</td>
</tr>
</tbody>
</table>

The personalities and knowledge of the participants differed between the groups. Overall, the groups generated mostly concrete physical suggestions. However, Group 3 and 4 generated more abstract suggestions than Group 1 and 2, which is probably due to the nature of the ideation topics.

- Compared to Group 4, Group 3 might have put more effort into creating suggestions on self-adjustment, and this might have been a more difficult topic to generate ideas on.
- Group 3 might have generated suggestions with many more inherent ideas and details than Group 4, which is not captured in the counting of suggestions.
- The participants in Group 3 might have had a less stimulating/inspiring experience during the group work and might have devoted less effort to generating ideas compared to Group 4, as they, on average, kept more ideas to themselves during the workshop and felt less engaged during the workshop (Figure 1).
Suggestions” in the present study correspond to “concepts” in the tests of the developed method in the OptiKrea Group [14] and the terms are regarded as synonymous in the present paper, although slightly different guidelines for counting “concepts” were used in the OptiKrea tests. Table 12 compares the number of suggestions generated during Stage 1 in the In2Rail groups with the number of concepts generated in the OptiKrea Group. Since the time used and the number of participants varied between the groups, the suggestion generation rate per participant has been calculated. One participant in Group 2 arrived 7.5 minutes into Stage 1 and it is uncertain how this affected the quantity of suggestions. However, calculating the suggestion generation rate per person for Group 2 for eight participants instead of nine only changes the rate slightly from 0.08 to 0.09. On the other hand, the fact that Group 2 used an unknown part of their time for generating suggestions in list-form makes it less meaningful to compare their suggestion generation rate with the rate of the other groups. Group 4 and the OptiKrea Group had approximately the same suggestion generation rate per person, whereas the other groups in the In2Rail project had a lower rate than the OptiKrea Group. At least four different aspects are important to highlight with respect to this comparison.

• Previous research on brainstorming has found that larger groups are less productive [17], and a similar effect might exist during non-verbal ideation.
• The participants in the OptiKrea Group had worked together during the development of the method and had previously tried similar ideation methods together, and therefore they were more familiar with each other and the method, which probably had a positive effect on the number of generated concepts.
• The ideation topics were different and might have given different opportunities for generating suggestions.
• In the OptiKrea Group, all the participants used their native language when applying the ideation method, whereas in the In2Rail groups the majority of the participants did not. Generating and discussing ideas in another language than one’s native language might be more difficult.

### 4.4 Comparing the outcome with the expectations

There were several expectations as to the outcome of the field test. Firstly, it was expected that it would be easy to apply the ideation method during the In2Rail workshop. It turned out that the groups were able to use it without training, but asked several questions, especially regarding Stage 3. The groups might have been able to work more independently if the method instructions had been available to each participant all the time. Further, having the supervisor available makes it easy for the participants to ask questions, and it is possible that if the groups had been on their own, they would have found out what to do without asking.

Secondly, it was expected that the In2Rail participants would share the views of the OptiKrea participants on the method, since the latter had been chosen carefully to be representative of a typical use setting. A correlation of $R^2=0.84$ was found between the average questionnaire scores of the In2Rail and OptiKrea groups, which indicates that an agreement exists between the views of the In2Rail and OptiKrea groups. However, a more thorough statistical analysis is required to confirm this finding, and one limitation in this connection is the small number of participants. The OptiKrea participants on average judged the method to be more useful, were more engaged during the workshop and more satisfied with their own contribution than the In2Rail participants did and were on average (Figure 1). This might be explained by the fact that the participants in the OptiKrea Group were more familiar with each other and similar ideation methods than the participants in the In2Rail workshop. Another possible explanation is that the OptiKrea Group was positively biased towards the method since they had been part of its development.

Thirdly, it was expected that the number of suggestions generated by the In2Rail groups would be comparable to that generated by the OptiKrea Group. Jackson and Poole [11] found in their study of ideation procedures in groups in naturally occurring contexts that their idea generation rate was lower than the rates reported in laboratory studies. During Stage 1 one of the In2Rail groups had a suggestion generation rate per person that was comparable to that of the OptiKrea Group, whereas the other In2Rail groups had lower rates.

These findings suggest that developing ideation methods for a certain context in a small group can be feasible, but a prerequisite might be that the participants should be carefully chosen to form a group that is representative of the context. It would be very interesting to perform future studies where the outcome of methods developed among non-representative participants (e.g. students) would be compared to the outcome of methods developed in a representative group after both these types of methods have been applied in a typical use-setting. Nonetheless, it can be concluded that testing the present method in a typical use setting which was different from the development context provided us with valuable insights into the method and suggestions for the future direction of our research.

The present study also highlights challenges that arise when applying a method outside an experimental environment.
5 CONCLUSIONS

The field test gave several insights into the ideation method, concerning how the method could be adapted to a certain situation, how stages were used in other ways than intended, the unequal distribution of the used time among the participants during the group review stages, and the feasibility of using a “pen-and-paper” approach.

Most of the content from the ideation sessions consisted of suggestions on how to address the ideation topic and the groups generated mostly concrete physical suggestions. The suggestion generation rate per person of one of the In2Rail groups during Stage 1 was comparable to that of the OptiKrea Group, whereas the other three In2Rail groups had lower rates.

The participants in the In2Rail groups were satisfied with the method. A correlation of R²=0.84 was found between the extent to which the OptiKrea participants and the extent to which the In2Rail participants agreed with different statements concerning how the method could be adapted to a certain situation, how stages were used in other ways than intended, the unequal distribution of the used time among the participants during the group review stages, and the feasibility of using a “pen-and-paper” approach.

REFERENCES


Acknowledgments

The OptiKrea participants partnering in the project, the Swedish Transport Administration, Vossloh Nordic Switch Systems and Infranord, are gratefully acknowledged for their contributions to the project. Finally, we would like to thank the participants in the In2Rail workshop.
Paper V
Conceptual design method for the development of railway products in the context of the Swedish deregulated railway market

Anna Malou Petersson* and Jan Lundberg
Division of Operation, Maintenance and Acoustics, Department of Civil, Environmental and Natural Resources Engineering, Luleå University of Technology, SE-971 87 Luleå, Sweden
*Corresponding author: E-mail: anna.malou.petersson@ltu.se; tel.: +46(0)722419179; fax: +46(0)920-492818.

1 INTRODUCTION

In its aim to establish a Single European Railway Area, the European Union (EU) has encouraged competition and the opening of the railway market through a number of gradual legislative proposals (Alexandersson, 2013). Sweden is one of the European countries that have implemented the most far-reaching reforms in the railway sector and can be considered to be several years ahead of current EU policy (Alexandersson, 2013). Therefore, Sweden is an interesting example for other countries to learn from and can show the way on how different issues related to the liberalization of the market can be addressed. The deregulation of the Swedish railway system started in 1988 with the vertical separation of train operations from the railway infrastructure. Railway infrastructure ownership, investments and maintenance responsibility were transferred from the public utility Swedish State Railways to a national authority, the Swedish Rail Administration (SRA). Public procurement of train operation on railway lines began shortly afterwards. In 1998, new policies led to the split-up of SRA into a client and contractor and competition in the area of railway maintenance was initiated in 2001. In 2010, SRA and several other Swedish authorities were merged to form the Swedish Transport Administration (STA). STA is responsible for the strategic planning of roads, railways, ship transport and aviation, as well as investment in and the operation and maintenance of roads and railways in Sweden.

A consequence of the deregulation process is that the responsibility for design and production of railway products originally belonging to the Swedish State Railways were transferred to companies and procured. For example, the design and production of turnouts were taken over by a company in 1990. In 1993, SRA initiated the public procurement of turnouts through frame agreements. Purchases made by public authorities such as STA are regulated by Swedish legislation. The Swedish Act on Procurement within the Water, Energy, Transport and Postal Services Sectors (APWETPSS) (Lag om Upphandling inom Områdena Vatten, Energi, Transporter och Posttjänster, 2007) applies to purchases concerning railway-related activities (Sundstrand, 2013). The APWETPSS is based on Directive 2004/17/EC of the European Parliament and of the Council (EU, 2004), and corresponding rules apply in other countries belonging to the EU.

Although STA doesn’t have a responsibility for the design and production of railway products within their organization, they deal with the conceptual design phase (the early stages of product development where requirements are formulated, ideas generated and concepts selected) in at least two different situations. Firstly, if STA recognize a need to introduce a new or improved product in the infrastructure they manage, they can procure this product based on stated requirements. In practice, this often means that the development of this product is outsourced to the company that receives the contract, since these products are typically not “off-the-shelf” products. Secondly, they can take part in development work in research projects. These can take place on a national level or, increasingly common, as partners in an EU research project.

Railway infrastructure and particularly track components are expensive assets with long life spans and the punctuality and effectiveness of train operations heavily depend on their seamless operation. The OptiKrea project, run by Luleå Railway Research Center at Luleå University of Technology, was initiated with the intention of promoting the technical development of railway products, especially from the point of view of maintenance and life cycle cost, through the collaboration between different railway market actors. In 2012, turnout-related failures were among the top ten causes of hours of disruption in Sweden (Trafikverket 2013). Turnouts cause at least 13% of the maintenance costs associated with the Swedish railway (Nissen 2009). Therefore, the turnout was chosen as a case object for the project. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions would be found.
A parallel goal of the project was to develop conceptual design methods, which would facilitate innovation, would be tailor-made for the railway sector and could be used in the future. At the heart of the project is the so-called “creative team”, which consists of representatives from each collaborating actor. The team is cross-functional in the sense that it represents the different functions that are important when developing a turnout, i.e. research, design, manufacturing, management, maintenance, and disposal.

In the present paper, the overall goal is to develop practical methods that STA can use to enhance collaboration during the early phases of product development to share knowledge and information between relevant actors while restricted to the rules by PPA as well as to decide on what concept(s) to take forward in a product development or procurement process. In Figure 1, a generic product development process is presented. The present study focuses on the parts concerning task definition and evaluation and selection of concept.

The paper is separated in two different parts. Firstly, the conditions for product development on the Swedish deregulated railway market are investigated. Secondly, a practical method that can facilitate collaboration during the conceptual design phase is developed and explained and justified in its intended context. The method is applied in a case study concerning turnouts. The application of the method is evaluated and discussed. Especially, we are interested in finding out how the different competences of the participants manifested itself during the application. By involving participants representing different relevant actors during the topic clarification phase, we expect that many viewpoints on the product-to-be-developed will surface and thus help in covering a wide range of aspects in the requirement list.

2 Method

2.1 Participants

The study took place in the creative team, a group of six participants comprising four railway professionals and two academic. The professionals each had 20-30 years of experience from the railway sector. Two worked with turnout-related issues at STA, one at an international company manufacturing turnouts and one at a contractor performing maintenance. The academics had experience from railway-related projects performed in close cooperation with STA and industry. Details of the participants are provided in Table 1. The participants were carefully chosen so that each participant had deep knowledge of the organization they represents which makes it possible for them to relate the application of the method to a real-life context and judge how useful it would be in other typical use contexts. In Sweden, the infrastructure manager typically runs its research and development projects in cooperation with or through research institutes and universities, and therefore academic researchers are relevant participants in the group.

<table>
<thead>
<tr>
<th>#</th>
<th>Organization</th>
<th>Position</th>
<th>Field of expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Contractor</td>
<td>Product engineer</td>
<td>Maintenance method and product development</td>
</tr>
<tr>
<td>B</td>
<td>Supplier</td>
<td>Development manager</td>
<td>Product development of turnouts, turnout engineering</td>
</tr>
<tr>
<td>C</td>
<td>STA</td>
<td>Specialist</td>
<td>Maintenance management of turnouts</td>
</tr>
<tr>
<td>D</td>
<td>STA</td>
<td>Specialist</td>
<td>Track and turnouts</td>
</tr>
<tr>
<td>E</td>
<td>Academia</td>
<td>Professor</td>
<td>Mechanical engineering design, maintenance, product development</td>
</tr>
<tr>
<td>F</td>
<td>Academia</td>
<td>Postdoc. research fellow</td>
<td>Applied acoustics and signal processing</td>
</tr>
</tbody>
</table>

Figure 1. Generic product development process
2.2 Collection of background information

Background information on the current situation regarding product development and procurement of railway products was acquired through semi-structured individual interviews among the professionals and semi-structured group interviews in the creative team as well as comments given by the participants during the executed workshops throughout the OptiKrea project. The interviews and workshops were recorded and transcribed. The transcription was qualitatively analysed. Information was also collected from relevant documents.

2.3 Development of method

The development of the method suggested in the present paper was based on findings in the literature, the collected background information and the findings from applying an alternative method. This alternative method had been proposed by the last author of this paper (a brief example is given in Lundberg (2000)) and is based on the method by Pahl et al., (2007) which incorporates utility theory in a systematic engineering design process. The method by Pahl et al. (2007) is well-tried in industry and academia and it normally appeal to engineers since it is structured and systematic. Some parts of this method are kept in the suggested method and have been refined through tests in the creative team. The development took place in close discussion with the professionals participating in the project. The authors of the paper had the role of leading the project and managing the research part. The creative team proposed issues concerning turnouts and addressed these issues by applying the methods proposed by the authors. The authors analysed the data from the applications and proposed changes to the method under development. The creative team gave their views on the methods after the applications and commented on any changes which were proposed.

2.4 Application of the method

The group applied the method to a real-life issue that they were working on within the scope of the OptiKrea project: How can the alignment of the track be prevented from deteriorating in turnouts? The application was performed during a two-day workshop. During the first day, the method was explained to the group and discussed, and the problem clarification stage was executed. During the second day, the group performed the goal-setting, requirements, concept screening and concept scoring stages. The total time used for applying the method was approximately 6 hours.

After going through the problem clarification, goal-setting and requirements stages the group was provided five concepts (V1-V5) that had been conceptually developed to a similar extent by the last author of the present paper as possible solutions to the topic. These concepts were based on ideas from ideation workshops that had taken place earlier during the OptiKrea project regarding the same topic. These concepts were used when during the concept screening and concept scoring stages.

2.5 Evaluation

After the tests of the methods, the participants answered a questionnaire and a group interview regarding the participants’ experience of the applied method took place to capture the participants’ immediate reactions to the ideation method that they had tried. The questionnaire presented different statements about each method, and the participants assessed their level of agreement with these statements, answering on a continuous scale from “Do not agree at all” to “Agree completely”, by making a mark on a line which was about 100 millimetres in length and where 0 represented “Do not agree at all”. The position of the mark was measured in millimetres from 0 with a ruler and divided by the total length of the line. The average value for each statement was calculated and the individual results were also compared to find out if any interesting patterns were present. The group interviews were audio-recorded and the transcribed recordings, along with the questionnaires, served as the basis for analysing the participants’ views on the methods.

To judge if the participants contributed different viewpoints, the contributions from each participant during the topic clarification and requirement specification stages, were compared. Firstly, all contributions were identified. A contribution was identified as one suggested requirement for the
requirement specification stage. A contribution from the topic clarification stage was identified as a “piece of information” that in turn was a sentence limited to one topic. Examples of raw data in the form of quotes that was counted as pieces of information are found in Table 2-Table 4. Contributions with the same meaning were merged. It was analysed how many participants had mentioned a certain contribution. The total number of contributions by each participant was counted, as well as how many contributions were made in each order of the talking sequence when the participants took turns presenting their contributions. The information that was actually presented to the group during the workshop was included in this analysis, not any answers that the participants had written down but did not share with the group. Although the participants are encouraged to provide any information they think of during this stage, in order to capture a wide range of aspects, only the contributions that were agreed to be relevant and taken forward by the group were part of the analysis. Otherwise the analysis would give an unrealistic picture of the method, as participants might have generated a lot of different contributions that give a positive impression in the analysis but are actually useless when it comes to developing the desired product.

Table 2. Replies to question 2: Who wants the problem to be solved and why?

<table>
<thead>
<tr>
<th>Who</th>
<th>Why</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>STA/Infrastructure manager</td>
<td>Wants to provide a reliable railway.</td>
<td>B</td>
</tr>
<tr>
<td></td>
<td>An irregular track alignment generates impacts that break down the turnout.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wants a predictable deterioration so that maintenance work can be planned.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Wants a method to perform track alignment with low LCC.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>Maintenance costs and traffic standstill are too large.</td>
<td>E</td>
</tr>
<tr>
<td></td>
<td>They have to pay to fix the failures.</td>
<td>F</td>
</tr>
<tr>
<td>Train operators/The customers of STA</td>
<td>Do not want speed reductions or disruptions in traffic.</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>Rolling stock breaks down</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Make demands on comfort and safety</td>
<td>C</td>
</tr>
<tr>
<td>Contractors</td>
<td>Viability, be able to carry out calculable maintenance with the right competence.</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>It should be clear to the machine drivers what they are supposed to do.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Replies to question 3: What are the (root) causes of the problem?

<table>
<thead>
<tr>
<th>Cause</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Settlements in the substructure/underballast. Drainage problem and similar.</td>
<td>A</td>
</tr>
<tr>
<td>Irregularities in the steel in the turnout propagate down in the macadam that in turn break down and give instability. (Also need to find the right standard level of maintenance of steel and track alignment)</td>
<td>B</td>
</tr>
<tr>
<td>The deterioration of the system depend on load from the use that lead to deformations. Uncertain where the load originates from.</td>
<td>C</td>
</tr>
<tr>
<td>The geometry changes so that the ballast pressure is exceeded</td>
<td>D</td>
</tr>
<tr>
<td>Either the track alignment methodology or the engineering design with the design of the area of the sleepers, or a combination.</td>
<td>E</td>
</tr>
<tr>
<td>Different load distributions between turnout and ordinary track.</td>
<td>F</td>
</tr>
<tr>
<td>Different track alignment methods for turnout and ordinary track.</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Replies to question 12: What are the future trends regarding technology, environmental aspects and ergonomics?

<table>
<thead>
<tr>
<th>Trend</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modularisation</td>
<td>B</td>
</tr>
<tr>
<td>Short time in the track for maintenance and installation</td>
<td>B</td>
</tr>
<tr>
<td>Lower energy usage and environmental emissions during manufacture and operation</td>
<td>D</td>
</tr>
<tr>
<td>Recycability</td>
<td>D</td>
</tr>
<tr>
<td>The self-inspecting turnout</td>
<td>E</td>
</tr>
</tbody>
</table>
3 THE CURRENT SITUATION

In this section, the findings regarding the conditions related to the development of new or modified railway products on the Swedish deregulated railway market are presented. The findings are based upon the interviews and workshops conducted among the participants.

3.1 Results

From the collection of background information, the following conditions concerning the procurement of new or modified railway products have been identified:

- Railway infrastructure and particularly track components are expensive assets with long life spans.
- The infrastructure manager has no in-house production of railway products but products are purchased through procurement.
- On a national level, the infrastructure manager is the sole purchaser of railway assets.
- All public utilities (including STA) must follow the APWETPSS.
- In research and development projects the APWETPSS does not apply. Learning from such project can be transferred to regular procurements.
- The delivery of a certain product is procured for a time span of several years through frame agreements.
- The procured product, as specified by the requirements given by STA in the tender request, does typically not exist in reality during the procurement. Hence, the information in the tender offers are marred by uncertainties which introduces a risk that the eventually delivered product will not meet the requirements as specified in the tender request to the stated cost.
- Once a winner of the procurement has been found, STA can work together with this supplier in developing a product that meets the requirements stated in the tender request.
- The requirements or stated evaluation process cannot be changed after the tender request has been distributed without major consequences such as repeated procurement.
- The tender award criteria and evaluation outcome should be expressed in a monetary value.
- Typically, the only value with no uncertainty in the tender request is the purchase price, and this is one of the reasons that the purchase price so far has been given the majority of the weight during procurement.
- It is difficult to formalize the trustworthiness of tenderers in a fair way that can be used during the procurement to sort out unrealistic offers and determine which company has the best necessary qualities to succeed in delivering a product that fulfills the requirements and simultaneously take the price into account.
- The current tender process presupposes that nothing can go wrong during the product development process, i.e. there is no redundancy in the process if something goes wrong.
- The requirements should preferably be written as functional requirements.
- The suppliers take a risk in developing a product that might be sold to only one buyer.
- The suppliers take a risk putting a too low price to win the procurement that will not cover the actual development costs.

Regarding the design of the product, the following conditions have been identified:

- It is not within the mission of STA to develop new products. STA should specify the scope of the product, whereas companies should develop and provide the actual product.
- The conditions specified in the tender request will direct the development.
The design of the product has an impact on the direct costs (e.g. investment and maintenance costs), which are financed by taxes, and the indirect societal costs (e.g. costs originating from a failure to provide public transport for citizens).

- The product to be developed must fit into the existing infrastructure and existing maintenance contracts. There are several interfaces to the existing infrastructure that must be taken into account, as well as existing maintenance equipment.
- The development of railway products must be performed in accordance to legislation.
- EN50126 is to be followed by STA when introducing new or modified products. EN50126 is a framework that describes the process to be followed to assure that all aspects on an asset from “the cradle to the grave” are considered. Different methods can be used to provide the information required by EN50126.
- According to law the Common Safety Method (CSM) must be applied if a considerable change is introduced.
- There is always an alternative when STA evaluates a situation. The alternative must not be a similar product, it can also be e.g. the option of driving a car instead of taking the train.

3.2 Discussion

During product development there is a risk that non-suitable requirements will be specified, and consequently not the most suitable concept be chosen. Whereas a normal product development process can to some extent compensate for this at a later stage, it is of outmost importance that this stage is completely covered in the case of a procurement case, since public procurement according to PPA does not allow the requirement list to be changed without major consequences once a tender request has been issued. As STA has outsourced production as well as maintenance of railway assets knowledge required to write the requirement specification has to some extent been transferred to suppliers and contractors. Further, STA need to specify requirements that assure that the product will meet legislation and has a responsibility to provide public utility while being economical with tax payers’ money. In conclusion, working out a requirement list and tender request is a delicate process during which many aspects need to be considered. The situation is further complicated by the fact that uncertainties and risks that are an inherent part of product development processes are not addressed in the procurement process that has been used so far since it does not provide any redundancy should undesired events occur. The situation on the deregulated railway market differs from that in a typical product development process, and this calls for an adapted process.

3 THE DEVELOPED METHOD

Figure 2 shows the overall conceptual design process developed within the OptiKrea project for taking an identified need to one or several concepts to take forward in the product development process. In the present paper the method for clarifying the problem and selecting concept(s) are presented. These phases are connected since the requirements formulated before concept generation are used to select among the concepts after concept generation.

The focus on costs led to the approach of making a decision based on a life cycle cost (LCC) calculation combined with a calculation of the societal cost (SC) predicted to be caused by introducing a concept. The LCC concept is an accepted approach in the railway community. By presenting quantified values, one can gain a better understanding of the type of changes in the design and maintenance strategy that can be planned to lower the lifecycle cost (Nissen 2009). The cost of maintenance plays an important role in the LCC analysis of assets like track infrastructure, where the operation and maintenance phase comprises major share of the system’s life cycle (Patra 2009). Therefore, it is relevant to consider the cost-of-ownership per year.
The evaluation shows that modification of requirements, weight and/or overall design is necessary. The evaluation results in further development of selected existing concepts.

Figure 2. Schematic representation of the process that the method presented in the present study is part of. The left side follows the process for research projects and the right side follows the case of procurement. Stages with bold frame are part of the present paper.

In one sense, the societal cost can be viewed as part of the LCC, as it describes costs that originate from a failure of the product. However, we have chosen to separate between direct and indirect costs. Direct costs include all costs that STA actually has to cover. The indirect societal costs are not financed by STA directly. It is however within the mission of STA to provide an infrastructure that limits societal costs. The LCC and SC values are combined into a concept score. Depending on the purpose of applying the method, the values can also be presented separately, which gives a more nuanced understanding of the implications of the evaluated concepts.

STA has rules and regulations regarding how LCC and SC are to be calculated and what input values that should be used. Since infrastructure managers in different countries apply different sets of rules and regulations concerning societal and life cycle costs, we do not discuss in detail how these calculations should be set up. The models need to be adapted to the specific context.

The LCC and SC are combined with requirements stating minimum values for a range of properties. Typically, STA is not prepared to pay more than what stated required minimum value would cost if it doesn’t decrease the LCC and SC costs. However, in some cases, when relevant soft parameters do not influence neither LCC nor SC in a meaningful way and STA wants to reward a better solution than the minimum requirements, then the values of the properties representing these values are transferred onto a monetary scale representing a monetary value that is included in the concept score. This procedure was however not tested in the present study but emerged as this need was identified after the study.

The method instructions are shown in Table 5 and each stage is described in Sections 3.1-3.5.
Table 5. Instructions to the method.

<table>
<thead>
<tr>
<th>Topic mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Short introduction to the topic by the initiator</td>
</tr>
<tr>
<td>2. All participants work individually with the questions in Table 3 during 10-20 minutes.</td>
</tr>
<tr>
<td>3. Each question is gone through at a time. The participants take turns presenting their response to the question. The contributions are documented where everyone can see them (e.g. computer connected to projector or whiteboard), discussed and summarized before the group proceeds to the next question.</td>
</tr>
<tr>
<td>4. Continue until all questions have been gone through.</td>
</tr>
<tr>
<td>5. Establish preliminary LCC and SC models and make an LCC and SC calculation with respect to today’s existing product/situation (if relevant) to see which the largest cost drivers are.</td>
</tr>
<tr>
<td>6. Check that the topic mapping and especially the (root) causes to the situation are well enough illuminated and understood to justify ideation. Otherwise analyze what must be done before proceeding to ideation.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Goal-setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The objective of formulating a goal-setting is to make sure that all participants have the same interpretation of what the project should achieve and to act as a reminder during the project. The goal-setting should be 1-3 sentences long (Pahl and Beitz, 2007).</td>
</tr>
<tr>
<td>2. All participants individually formulate a goal-setting (their idea of the outcome of the project)</td>
</tr>
<tr>
<td>3. The participants take turns presenting their formulations and agree upon a joint goal-setting.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requirement specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. All participants work individually formulating requirements under each (relevant) headline in the checklist by Pahl et al. (2007, page 149) for 10-20 minutes.</td>
</tr>
<tr>
<td>2. Each headline is gone through at a time. The participants take turns presenting their stated requirements. The requirements are documented where everyone can see them and discussed. Similar requirements are merged, and a check is performed that all contributions are indeed requirements (can be answered with yes/no) and if they are necessary. Summarize the content under each headline before proceeding to the next step.</td>
</tr>
<tr>
<td>3. Identify possible “soft parameters” that is not valued by the LCC or SC models and where a value above the minimum requirement value is sought after. Transfer the value to a monetary scale:</td>
</tr>
<tr>
<td>- The requirement value should correspond to zero in the monetary scale.</td>
</tr>
<tr>
<td>- The maximum limit of the monetary scale should be the maximum sum the purchaser is prepared to pay for an improvement, and correspond to the value of the property connected to the requirement they would expect at this price.</td>
</tr>
<tr>
<td>- The values between the minimum required value and the maximum value that the purchaser is prepared to pay is extrapolated.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decide weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weights for the evaluation criteria, LCC and SC, are determined through agreement. The participants reflect on how big influence they desire the SC to have on the concept score in the specific case.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concept screening</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Go through the requirements list and choose the requirements to be used for screening of concepts.</td>
</tr>
<tr>
<td>- Prioritize requirements that will not be based solely on guesses.</td>
</tr>
<tr>
<td>- Prioritize the most important requirements.</td>
</tr>
<tr>
<td>- Choose as few requirements as possible, preferably maximum 10 requirements.</td>
</tr>
<tr>
<td>2. The group uses a screening chart (adapted from Pahl and Beitz, 2007) to select the concepts that should be further developed. As soon as an idea receives a minus on any criteria, it is eliminated and not screened further. Reason for elimination is given in the screening chart.</td>
</tr>
<tr>
<td>3. Information is often insufficient at this stage and then the group has to make a decision if it has time to gather the needed information or if the concept should be selected/eliminated despite lack of information at the current stage.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concept scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A concept must fulfill all requirements to proceed.</td>
</tr>
<tr>
<td>2. Calculation of LCC according to the agreed LCC model.</td>
</tr>
<tr>
<td>3. Calculation of SC according to the agreed model.</td>
</tr>
<tr>
<td>4. Calculation of values connected to soft parameters according to the specified monetary scales.</td>
</tr>
<tr>
<td>5. Calculation of concept score according to Equation 1 (see Section 3.5).</td>
</tr>
<tr>
<td>6. The lowest rating indicates what concept(s) are most promising to develop further.</td>
</tr>
</tbody>
</table>
Table 6. Checklist with questions to be answered by each participant individually during the topic clarification.

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What is the core of the problem and wherein lies the greatest need?</td>
</tr>
<tr>
<td>2</td>
<td>Who wants the problem to be solved and why?</td>
</tr>
<tr>
<td>3</td>
<td>What are the (root) causes of the problem?</td>
</tr>
<tr>
<td>4</td>
<td>What hidden wishes, demands and exceptions are involved?</td>
</tr>
<tr>
<td>5</td>
<td>What are the tasks which the product has to solve?</td>
</tr>
<tr>
<td>6</td>
<td>Which possibilities are open and which possibilities are closed for the achievement of the new concept?</td>
</tr>
<tr>
<td>7</td>
<td>What properties should the concept have/not have?</td>
</tr>
<tr>
<td>8</td>
<td>What alternative products exist today?</td>
</tr>
<tr>
<td>9</td>
<td>What are the demands according to standards and regulations?</td>
</tr>
<tr>
<td>10</td>
<td>What are the wishes and demands regarding possibilities for product upgrading?</td>
</tr>
<tr>
<td>11</td>
<td>Are there other aspects that should be taken into account?</td>
</tr>
<tr>
<td>12</td>
<td>What are the future trends regarding technology, environmental aspects and ergonomics?</td>
</tr>
</tbody>
</table>

3.1 Topic clarification

In a group composed by participants representing different actors, the topic clarification stage is especially important, since the participants have different backgrounds and typically different apprehensions of the issue. This diversity makes it important to achieve a common basis to start from, but also represents a major advantage of a cross-functional group. A cross-functional group with participants from different organizations is more likely to cover a wide range of issues related to the topic compared to a homogeneous group. However, some strategy to facilitate information sharing between the participants is required, since an expert in one domain often do not know what information is of use to an expert in another domain (Strauss et al. 2011) and since it has been found that less information is shared within a team when its members possess diverse sets of knowledge and (Mesmer-Magnus and DeChurch, 2009). A large body of research shows that a group covers more ideas if the individuals in the group generate ideas on their own instead of interacting as a group (see e.g. Byron, 2012; Mullen, Johnson and Salas, 1991). Especially, the individuals will cover a wide range of information if they first ideate individually before interacting with each other, as they will then not have been influenced by the other people in the group. Hence, the method starts with letting the participants individually answer the questions shown in Table 6. The questions are developed from Lundberg (2000) and originally the most questions are inspired by Pahl et al (2007, page 150). All questions might not be relevant to all product development situations, but by checking this list one can be confident to have covered many important aspects. Letting the word be free, a few individuals typically dominate the discussion and perspectives represented by other members of the group are lost. Therefore, during the group review, the participants take turns stating and explaining their contributions.

In railway contexts, the product to be developed can be compared to an existing situation or solution. Therefore, it is valuable to make a comprehensive analysis of the LCC and SC of the existing product or situation to identify the major cost drivers. This will provide important input to the team that is to generate concepts as well as highlight where effort should be put in the writing of the tender request. In situations involving procurement, the LCC and SC models to be used in the tender request cannot be changed later in the evaluation process and hence requires special attention. It is important to remember that the design of the LCC and SC will guide the development of concepts/products among suppliers. The development of LCC and SC models should not be seen as an isolated event during a specific project but be reused and continuously modified to reflect the learning from different projects and procurements. In this way, the procurement can be refined to become a powerful tool in guiding the suppliers to develop products that actually meet the objectives of the infrastructure manager.

3.2 Goal-setting

The goal-setting is important in making sure that the participants have actually gained a common basis and agree on the overall goal of the process. The goal-setting should shortly, in maximum 3 sentences,
specify the overall goal of the project (Pahl and Beitz, 2007). Optimally, the participants are quite in agreement at this stage but if not, the group needs to discuss and agree upon a goal-setting before continuing. Compromises might be necessary, but the purchaser generally has the last words. Another important point about the goal-setting is that it should specify a vision of what the end-product fulfills, but not specify how it should be accomplished, as this will lock the development to certain solution principles.

3.3 Requirement specification

After working with the topic clarification and goal-setting, it is time for the group to consider the requirements regarding the product. This step is performed in a similar way as the topic clarification step. To their help, the participants have a list of headings and examples as compiled by Pahl et al. (2007, page 149); geometry, kinematics, forces, energy, material, signals, safety, ergonomics, production, quality control, assembly, transport, operation, maintenance, recycling, costs and schedules. In the list of headings used in the present study recycling was replaced by “environment” to cover a wider range of environmental issues such as pollution. “Design” and “marking and notations” were added to take any wishes regarding the esthetics into account. Finally, “handling” has been added to the list to cover procedures which must be taken into account during installation and operation, for example, and “operation” has been removed, since it is regarded as being covered by other headings. In many cases the group can restrict itself to the requirements that concern the modification of the product.

After the requirement specification, the process divides depending on if it is a research project or procurement, see Figure 2. In a research project, the group moves on to the ideation phase where solutions addressing the topic are collected and/or generated. These solutions need to be organized and a concept screening is typically performed to sort out a number of ideas that should be developed further conceptually. Different methods can be used during this phase, and the result should be a number of different concepts that are developed to a similar degree. In Section 3.4 we suggest one such method. How well conceptually developed the concepts are determines what method is most suitable. The method described in Section 3.4 assumes that the concepts are developed to a similar extent and that some information is gathered about them after ideation has taken place.

A workshop with representatives from relevant actors will provide a wide range of aspects on the product. However, in the case of procurement, the infrastructure manager needs to further review the requirements and add other relevant requirements before issuing a tender request. The requirement list and the LCC and SC models need to be compiled in a tender request. Each supplier’s individual process results in a tender offer delivered to the infrastructure manager. These tenders each represent a concept to be evaluated in comparison with other concepts according to the specified procedure.

3.4 Concept screening

Concept screening is performed by using a screening chart based on that by Pahl et al. (2007, page 108) and adapted to the specific context. The elimination criteria have been modified and are described in Table 7. The reason for elimination should be given in the screening chart to facilitate later understanding of why a certain concept was chosen or discarded.

An important modification to the screening chart proposed by Pahl et al (2007, page 108) is that in the suggested method the product development team does not consider if a concept fulfills all the demands of the requirements list. Instead, the team selects a subset of approximately ten requirements that are considered to be of high importance and to be possible to judge at the current stage. This was found by the group to streamline the concept screening, as many properties, and hence the fulfillment of related requirements, are difficult to judge at the current stage of development.
Table 7. Description of elimination criteria

<table>
<thead>
<tr>
<th>No.</th>
<th>Elimination criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Addresses the core problem</td>
<td>Check if the concept actually has the potential to solve the main problem connected to the product development process.</td>
</tr>
<tr>
<td>2</td>
<td>Realisable in principle</td>
<td>Check if the concept is realizable in respect of performance, layout, etc (Pahl et al, 2007). Cost is also considered under this criterion.</td>
</tr>
<tr>
<td>3</td>
<td>Selected requirements are fulfilled</td>
<td>The group checks if the concept fulfills a subset of about ten requirements from the requirements list that has been selected by the group because of their high importance and possibility to judge at the current stage.</td>
</tr>
<tr>
<td>4</td>
<td>Preferred or discarded of different reasons</td>
<td>This is a check if there are certain reasons to prefer or discard a concept. Such reasons can be e.g. the possibility to use an earlier developed product, that the concept is known to fit with the existing infrastructure or that the organization has no knowledge of the design principles used in the concept.</td>
</tr>
</tbody>
</table>

3.5 Concept scoring

The concept scoring method is similar for both research projects and procurement. During research projects, however, the stage can be applied iteratively. New concepts or modifications of concepts might be generated in the evaluation phase. In the procurement case, the information delivered from the suppliers in tenders need to be of high enough quality to act as input in the LCC and SC models, and all requirements must be addressed.

The concept score is calculated by:

\[ TC_i = LCC_i + w_{sc,i}SC_i - \sum_{j=1}^{n} SP_{ij} \]

Equation 1

where \( TC_i \), \( LCC_i \), \( SC_i \), and \( SP_{ij} \) is the total cost, LCC, SC and the estimated monetary worth of the soft parameter(s) of the \( i^{th} \) concept, respectively, \( w_{sc,i} \) is the weight of SC compared to LCC for the \( i^{th} \) concept and \( n \) is the number of soft parameters. The worth of the soft parameters are subtracted from the LCC and SC, since it by definition is better the greater it is, whereas TC is better the smaller it is. Note that TC is not a real cost, but an aggregation of different parameters on a monetary scale and represents a score designed to pick out the most suitable concept based on the need of the infrastructure manager.

Due to the possible restrictions concerning how LCC and SC can be calculated, the team responsible for the purchase have the opportunity to assign the SC a weight compared to LCC if they think that the SC is inappropriately low valued compared to LCC considering the needs in the certain case.

The concept score should also be calculated for the existing product or solution, so that the proposed concepts can be compared to the “benchmark” solution. The concept with the lowest total cost (TC) is the winner of this evaluation. In the case of procurement, this winner is strict, unless the purchaser has specified any certain rules in the tender request. This can be the case if e.g. two bids are so close that the uncertainty in the calculation is greater than their difference.

4 APPLICATION OF METHOD

4.1 Topic clarification

The group used about 17 minutes to write down their replies to the questions in Table 6. Examples of their replies are shown in Table 2-Table 4. These examples indicate that the participants contributed different viewpoints on the questions. The quantitative analysis of the all contributions shows a similar pattern. The group in total generated 92 different pieces of information during the problem clarification. As can be seen in Table 8, 80 were mentioned by only one person and none were mentioned by more than four people.
Figure 3 shows the total number of pieces of information across all categories contributed by each participant. Figure 4 shows the total number of pieces of information depending on the order of talking sequence. Figure 5 shows the number of times each person spoke in a certain order during the topic clarification stage.

It was found that most contributions mentioned during the topic clarification and requirement specification stages were unique for each participant. It is possible that the participants might have written down more requirements than those they mentioned, and especially they might have refrained from mentioning a requirement that had already been contributed by someone else. During most of the headlines, the participants took turns mentioning their requirements, and they rotated who started. Figure 4 shows the number of contributions made depending on order in the talking sequence. It appears as during the topic clarification the three first people to speak contributed more information than the three last. During requirement specification, it appears only the last and 6th person contributed less than the others. Although the number of contributions declined, new contributions were made also by the last participants in each round and this indicates that they were able to provide relevant information that had not been presented yet. It is believed that this is possible due to their different knowledge and perspective of the topic. However, not all participants spoke the same number of times in each order. Figure 5 shows how many times each participant spoke in a certain order during topic clarification. E.g., participant F spoke three times in position four and four times in position six whereas participant D spoke only two and one times in those positions. Figure 3 shows that each person contributed a different number of pieces of information and requirements. Therefore, these results might be affected by in which order each person spoke and the relevance of their knowledge to the topic at hand, as well as their personality.

<table>
<thead>
<tr>
<th>Table 8. The number of participants mentioning a specific demand/piece of information during the topic clarification and requirement stage, respectively</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Figure 3. The total number of pieces of information and demands mentioned by each person.
4.2 Goal-setting

After reviewing and discussing the goal-settings written down by each participant the group agreed on the following goal of the product development process:

“A turnout that maintains its track geometry, is possible to adjust when the track position changes, in a cost efficient way (LCC) that gives few [traffic] disturbances.”

4.3 Requirements

The group used approximately 10 minutes to write down requirements individually. The group compiled 58 relevant requirements. As is shown in Table 8, 48 requirements were mentioned by only one person. No requirement was mentioned by more than four participants. Figure 3 shows how many requirements each person contributed. The total number of requirements mentioned depending on the number in the speaking sequence is shown in Figure 4 (only the headlines during which the participants took turns are taken into account).
4.4 Concept screening

During the concept screening, discussions around one of the concepts led to one additional concept that were added, V6.

The group selected a subset of demands from the requirements list to be used during the concept screening, see Table 10. Figure 6 shows the screening chart that was used. Three concepts (V2, V3 and V4) were eliminated because they did not fulfill at least one of the selected requirements in Table 10.

Table 10. Subset of requirements chosen by the group to be used during concept screening (quotes)

<table>
<thead>
<tr>
<th>No.</th>
<th>Demand from requirements list</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Can be adapted to all types of turnouts with a fixed crossing</td>
</tr>
<tr>
<td>2</td>
<td>Must fulfill the minimum limits for alignment (EN, TDOK)</td>
</tr>
<tr>
<td>3</td>
<td>The speed of deterioration (the vertical and lateral movement as a function of time) must be less than today</td>
</tr>
<tr>
<td>4</td>
<td>The base design must be able to absorb energy without getting plastic deformation.</td>
</tr>
<tr>
<td>5</td>
<td>Prefabricated</td>
</tr>
<tr>
<td>6</td>
<td>Delivered in sections (modularized)</td>
</tr>
<tr>
<td>7</td>
<td>Easy to transport to installation location</td>
</tr>
<tr>
<td>8</td>
<td>Easy to adjust the track alignment</td>
</tr>
<tr>
<td>9</td>
<td>Maximum LCC the same as today</td>
</tr>
<tr>
<td>10</td>
<td>Prototype manufactured within one year</td>
</tr>
<tr>
<td>11</td>
<td>Mean time between failure at least 20 years (steel)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concept</th>
<th>Addresses the core problem (+/-/?), Realisable in principle (+/-/?), Selected requirements are fulfilled (+/-/?), Preferred or discarded from different reasons (put a cross and write reason in the comment field)</th>
<th>Reason for elimination/information that is missing/comments</th>
<th>DECISION (+/-/?), (+), (+), (-), (-), (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>++ + ***                                                                                                                             *Depending on the root cause Provided USP are used Other beam types can be used</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>++ + -                                                                                                                               Does not fulfill requirement No. 3</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>++ + -                                                                                                                               Does not fulfill requirement No. 10</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>++ + -                                                                                                                               Would require an adjustment device to fulfill req. no. 3 Does not fulfill requirement No. 5</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>V5</td>
<td>++ +                                                                                                                                Might require an adjustment advice to fulfill req. no 8 Might not be needed to adjust the alignment ***More information would be beneficial</td>
<td>+</td>
<td>***</td>
</tr>
<tr>
<td>V6</td>
<td>+ + +</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

Figure 6. Screening chart used during the concept screening process.
4.5. Concept scoring

The group decided to take only costs that are connected to track alignment into account in the LCC and SC calculations. The LCC was calculated according to:

\[
LCC = \frac{C_c}{T} + \frac{C_{\text{inst}}}{L} + \frac{C_{\text{align}}}{I_{\text{align}}} + \left(C_{\text{m,n}} - S_{\text{track}}\right)
\]

Equation 2

where \(C_c\) is the cost price, \(C_{\text{inst}}\) is the installation cost, \(l\) is the length of life, \(C_{\text{align}}\) is the cost for performing a track alignment, \(l_{\text{align}}\) is the track alignment interval, \(C_{\text{m,n}}\) is other maintenance costs based on solution V0, and \(S_{\text{track}}\) is savings in other maintenance costs with a concept compared to V0 due to an improved track alignment. All parameters are given on a yearly basis.

The SC was represented by the societal cost caused by failures of turnouts and the societal cost related to closing of the track when performing planned maintenance. The average number of failures per turnout and year with the benchmark solution V0 is \(n_{\text{fail},V0}\). Of these, a certain share \(A_{\text{stop}}\) results in a traffic stop (no trains are allowed to pass the turnout). Of the share \(A_{\text{stop}}\), a share \(A_{\text{align}}\) depends on failures related to the track alignment. Hence, the number of alignment related failures per turnout per year with V0 can be calculated. To estimate the cost of these failures the number of failures is multiplied with the average time length \(t_{\text{fail}}\) for taking care of an alignment related failure and the societal cost for unplanned stops per time unit \(c_{\text{unplan}}\). This cost concerns V0. To make an estimation of the cost of unplanned stops of concept V1, V5 and V6, this cost was multiplied with the expected share of today’s number of failures \(A_{\text{unplan}}\) that would be the case with each of these concepts. The estimation of this share was based on the assumption that a design that implied a more stable alignment would result in lower number of failures. The societal cost related the closing of track due to maintenance related to track alignment was found by multiplying the societal cost per time unit for planned stops \(c_{\text{plan}}\) with the average time to perform track alignment \(t_{\text{align}}\), divided by the track alignment interval. All parameters are given on a yearly basis. The mathematical expression becomes:

\[
SC = n_{\text{fail},V0}A_{\text{unplan}}A_{\text{stop}}A_{\text{align}}t_{\text{fail}}c_{\text{unplan}} + \frac{c_{\text{plan}}A_{\text{align}}}{t_{\text{align}}}
\]

Equation 3

The weight of SC was decided to be 2 in this specific case and this decision was taken to give the societal cost some extra weight compared to the LCC since some participants thought that the hourly rates for a standstill were too low and not reflecting the actual effects on society. Table 11 shows the resulting LCC, SC and concept score. The input data was estimated by the group. According to these calculations, V6 is the most promising concepts in terms of cost per year. Despite an increase in purchase price for V5 and V6 compared to V0 and V1, these concepts still got a lower LCC mainly due to the expectation that no track alignment would be required during their lifetime. In this case, the effect of the societal cost was small.

### Table 11. LCC, SC and score of each concept.

<table>
<thead>
<tr>
<th>Concept</th>
<th>V0</th>
<th>V1</th>
<th>V5</th>
<th>V6</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCC (SEK/year)</td>
<td>170000</td>
<td>140667</td>
<td>131111</td>
<td>114545</td>
</tr>
<tr>
<td>SC (SEK/year)</td>
<td>22400</td>
<td>14000</td>
<td>1000</td>
<td>500</td>
</tr>
<tr>
<td>Score (SEK/year)</td>
<td>214800</td>
<td>168667</td>
<td>133111</td>
<td>115545</td>
</tr>
</tbody>
</table>
4.6 Evaluation of application

4.6.1 Questionnaire results

Figure 7 shows to what extent the participants on average agreed with different statements about applying the method. Table 12 shows the statements in the questionnaire. For comparison, the result from applying the method based on Pahl et al (2007) is included. Overall, the questionnaire results indicate that the participants were more satisfied using the developed method than the method based on Pahl et al (2007). Especially, the participants found it to be considerably more time efficient and easier to carry out. Interestingly, the participants found the developed method to have greater potential for improvement than the systematic method. A reason could be that the developed method was easier to understand and therefore easier to vision how it could be improved. Another reason is that when applying the Pahl et al (2007) method, the participants did not have any other method/application to compare the process with.

There are some interesting patterns in the detailed analysis of the individual participants’ levels of agreement. Regarding the question about assurance of choosing the right solution that has the greatest potential to solve the issue, the participants representing STA had a considerable lower level of agreement compared to the other participants (on average 0.30 compared to 0.82). The participants representing STA also had a considerable lower level of agreement concerning the group making an objective choice of what solutions to be further developed (on average 0.36 compared to 0.82). A possible explanation is that more time would have been required to work with the requirement list and the STA representatives had most experience from the effect of a deficient requirement list on the end result.

![Figure 7. Questionnaire results](image)

Table 12. Statements in the questionnaire

<table>
<thead>
<tr>
<th>No.</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I am engaged in the topic.</td>
</tr>
<tr>
<td>2</td>
<td>I think the method was easy to carry out.</td>
</tr>
<tr>
<td>3</td>
<td>I feel that the criteria we chose for evaluating the concepts are relevant.</td>
</tr>
<tr>
<td>4</td>
<td>I feel confident that we have selected the concepts that have the greatest potential to solve the problems associated with the topic.</td>
</tr>
<tr>
<td>5</td>
<td>Eventually, we as a group agreed on which concepts should be selected for further development.</td>
</tr>
<tr>
<td>6</td>
<td>I think that we made an objective choice regarding which concepts should be developed further.</td>
</tr>
<tr>
<td>7</td>
<td>I think the method is time-efficient.</td>
</tr>
<tr>
<td>8</td>
<td>I felt engaged during the workshop.</td>
</tr>
<tr>
<td>9</td>
<td>I am satisfied with my own contribution during the workshop.</td>
</tr>
<tr>
<td>10</td>
<td>I think the method is useful.</td>
</tr>
<tr>
<td>11</td>
<td>I think the method can be improved.</td>
</tr>
</tbody>
</table>
4.6.2 Interview results

Overall, the participants during interviews expressed that they were satisfied with the method. Their comments regarding the method can be summarized in three different areas that all participants highlighted as positive. Firstly, the method provides a structured way of thinking through the entire process and the structure fits fairly well to what STA wants to accomplish. It provides an approach to how concepts should be graded. Secondly, the topic clarification gives a good picture of the problem covering a lot of aspects and helps the group in quickly agreeing on the issue. The topic clarification lays the foundation for making a good evaluation of the concepts. If the pre-stages are performed satisfactory then the concept screening and scoring goes smoothly. Thirdly, the method provides documentation of the process which makes it easy to return to a process several times and to reuse the process in a future development project or procurement.

The method process was not found by the participants to be neither too banal nor too complex. The LCC and SC models can however be more sophisticated. Reflections on the part of STA representatives on the results was that it might be difficult to convince the purchase department that the most expensive concept should be purchased, even if it got the best score.

5 DISCUSSION

The proposed method is a step towards taking the knowledge of different railway actors into account during the development of railway products. One major advantage of using the proposed approach is that it provides structure to collaboration during product development. Other methods could also be used, if they facilitate collaboration through a clear structure that makes meetings efficient and facilitates the contribution of all participants’ knowledge. It was found that the different participants made different contributions which indicate that it is indeed possible to collect a great range of information by involving representatives from different relevant actors.

During such collaboration, it must be remembered that each person has different aims of joining the group. The present study took place during a research project, and no competing actors were present. In a procurement situation, all possible suppliers must be invited to such a workshop according to PPA. In this situation, it can be expected that competing suppliers would not want to mention any information that might give their competitors an advantage. On the other hand, they would at this stage have the possibility to influence how STA will formulate the requirements by providing relevant background information. While the process leader should keep the different aims of the participant in mind, such a workshop would be an opportunity for the community around the product to be procured to influence the development and provide STA with important information before the tender request is formalized. The suppliers would receive information on STA’s view on the product and future needs which can hint them on in what direction they should develop their methods to meet the upcoming tender request.

Although the proposed method provides a structure during the conceptual design phase, there are several issues that are not addressed in its current form. Most importantly, uncertainty and risks have not been taken into account. It is very easy to guess an input figure and then believe that the outcome is the truth. Guessing both costs and technical parameters are difficult. It is of decisive importance to incorporate uncertainty in the process to be able to make decisions based on an understanding of the risk of choosing a certain concept.

The intended users of the suggested method are non-experts when it comes to the conceptual design phase and decision support methods. The participants found the method to be neither too banal nor too complex. In our view non-expert users should apply transparent methods that allow them to control the decision support process and understand why differences between alternatives arise while complex methods should be applied by expert users that can judge if they are appropriate for the topic at hand and understand why a result emerges.

6 FUTURE WORK

The structure of the method is adapted to the infrastructure manager’s point of view. However, the process should be further aligned with EN50126 in the future and applied during field tests to identify
possible improvements. An extension of the method includes incorporating uncertainty in the process. It is of interest to understand what incentives it creates during procurement and how the LCC and SC model can be used to identify cost drivers and drive the product development in the desired direction. Connected work concerns how tender requests and contracts should be designed to generate trustworthy tenders and avoiding suppliers to promise more than what they are able to deliver.

ACKNOWLEDGEMENTS

The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged. The OptiKrea partners Trafikverket (STA), Vossloh Nordic Switch Systems and Infranord are gratefully acknowledged for their contributions to the project.

REFERENCES

Paper VI
Applying action design research (ADR) to develop concept generation and selection methods

Anna Malou Petersson*, Jan Lundberg
Luleå University of Technology, SE-971 87 Luleå, Sweden
* Corresponding author. Tel.: +46-(0)920-491734; fax: +46-(0)920-492818. E-mail address: anna.malou.petersson@ltu.se

Abstract
During conceptual design, concept generation and selection methods can be used to facilitate the generation of new ideas and the selection of the most promising suggestions. The present paper is an effort to understand better how action design research can be utilized to develop such methods. Using action design research, methods were developed through iterative cycles of building, testing and evaluation in the context of a deregulated railway market, through a close collaboration between actors from industry, a government agency and academia. The approach was found to be feasible and to provide generalization of the context-specific findings through the formulation of design principles.

Keywords: concept generation; concept selection; methods; action design research; collaboration; design principles; In2Rail

1. Introduction
During conceptual design, concept generation and selection methods can be used to facilitate the generation of new ideas and the selection of the most promising suggestions. A great number of concept generation and concept selection methods, proposed by practitioners as well as researchers, are available to assist the engineer in the process of inventing and creating new products. Many researchers believe that conceptual design methods are used to a small extent in industry (e.g. Shah et al. [1]). Gish and Hansen [2] showed that contextual factors can be important in idea work and we believe that one reason for methods not being adopted in industry may be that they are not adapted to the context. In the present paper, we report from a project where action design research (ADR) was used to develop concept generation and selection methods in close cooperation with actors from industry, a government agency and academia. No previous studies have, to the best of the authors’ knowledge, explicitly used ADR to generate prescriptive knowledge on concept generation and selection methods through interaction between researchers and engineers in a real-world context. However, other researchers have applied action research to study developed methods, e.g. a computer support system for systematic design [3] and methods for service innovation in a multi-disciplinary context [4]. An important difference between the approaches applied in these and the present case is that we have let the target context shape the proposed methods through cycles of building, intervention and evaluation until the participants felt no need for further refinement.

ADR was proposed by Sein et al. [5] in an effort to blend design research with action research. Design science is the study of artefacts in their context [6], whereas action research is intervention in a social situation in order to both improve this situation and learn from it [7]. The purpose of ADR is to generate prescriptive design knowledge through learning from the intervention of building and evaluating an artefact in an organizational setting to address a problem [5]. Following Sein et al. [5], artefacts are viewed as ensembles for the purpose of this paper, meaning that the interaction between design efforts and contextual factors throughout the design process is manifested in the form, structure, goals, and conceptualization of the developed artefact. According to Rogerson and Scott [8], it is uncommon that a design-based intervention will turn out as planned at the first attempt, especially when addressing a social situation with practitioners involved. Therefore, ADR is problem-driven and
aims to build design principles based on iterative cycles in the same context [5]. ADR was proposed within the field of information systems. An artefact in this field is a specific bundle of hardware and software that is assembled to fulfil information needs. It would therefore be reasonable to assume that the development of artefacts for information handling in other fields could also profit from ADR.

As the project described in the present paper aimed to develop concept generation and selection methods in a certain context through interaction between researchers and engineers, it was deemed that ADR would be useful in guiding the study, as it allows the researcher to deliver a practical outcome for the involved organization whilst simultaneously meeting academic standards. The present paper is an effort to understand better how ADR can be utilized in the area of engineering design by exemplifying how the principles of ADR can be applied to the development of concept generation and selection methods, and reflect on the feasibility of using ADR for the development of such methods.

2. Background of the project

The turnout is a vital part of the railway infrastructure and a failure of a turnout, especially in a critical location, can cause significant delays and societal costs. Therefore, the OptiKrea project, run by Luleå Railway Research Center at Luleå University of Technology, was initiated in 2012 with the intention to promote the technical development of turnouts, especially from a maintenance and life-cycle-cost point of view, as a representative case of the development of railway products. The deregulation of the Swedish railway has resulted in different market actors managing, supplying, maintaining, utilizing and researching the railway. The idea behind the OptiKrea project was that, by integrating the different perspectives on and knowledge about the turnout that the different actors possess, better solutions would be found. A parallel goal of the project was to develop working methods, including concept generation and selection methods, which would facilitate innovation, would be tailor-made for the railway sector, and could be used in the future.

3. Applying ADR in the project

ADR consists of four stages [5]: 1) problem formulation; 2) building, intervention (i.e. using the artefact in the target environment), and evaluation (BIE); 3) reflection and learning; and 4) formalization of learning. Each stage is based on certain principles and involves the execution of certain tasks. An overview is shown in Fig. 1. In this section, each stage is described briefly and exemplified through a description of the development of concept generation and selection methods in the OptiKrea project. This paper deals only briefly with the results and findings of this development work to illustrate the ADR process, but a full account of the concept generation method can be found in Petersson and Lundberg [9,10] and the concept selection method is treated fully in Petersson and Lundberg [11].

3.1. Problem formulation

The first ADR stage is problem formulation, guided by the principles of practice-inspired research and a theory-ingrained artefact [5]; i.e. existing knowledge should be used to create an artefact that addresses a practical need interwoven with a context. The first two tasks in this stage are to identify and conceptualize the research opportunity and to formulate initial research questions.

The trigger for the present study was the insight that, since different functions involved during the life cycle of railway products are spread out over different actors, a lack of suitable collaboration strategies increased the risk of suboptimization during product development. Specifically, we were interested in the conceptual design phase, since it has a considerable influence on the subsequent steps during product development with regard to cost, quality and performance (e.g. [12]). By involving representatives from several actors during the conceptual design phase, more viewpoints on the product to be developed would be shared and thus a better product with higher maintainability would eventually be designed. A way to address this issue is to use concept generation and selection methods to structure the interaction between participants representing different actors. The initial research question formulated was therefore how concept generation and selection methods should be designed to capture the benefits of collaboration during the conceptual design phase in the given context. It was very easy to recruit different railway actors to the project, which was interpreted as a sign that a real need had indeed been identified.

The third task of this stage is to cast the problem tentatively as an instance of a class of problems, as this facilitates for the researcher to generate knowledge that can be applied to the class of problems that the specific problem exemplifies [5]. No concept generation or selection methods had previously been developed specifically for groups with members possessing different functional knowledge and representing different organizations. According to Straus et al. [13], there is no knowledge of what happens in inter-organizational groups, as different actors bring different cultures and agendas with them to such groups. Therefore, it was concluded that design principles for concept generation
and selection methods to be used in cross-functional inter-
organizational groups would be an interesting research topic,
as deregulation and outsourcing have become increasingly
common. As a result, the research of this part of the project
was framed as addressing the following class of field
problems: concept generation and concept selection in cross-
functional inter-organizational groups.

Task 4 involves identifying contributing theoretical bases
and prior (technology) advances. A literature review was
performed on idea generation, decision making and methods
for the conceptual design phase. A test of three established
concept generation methods was executed among the project
participants [14] to find empirical evidence as to how a cross-
functional inter-organizational group interacts during concept
generation.

Task 5 involves securing long-term organizational
commitment. The participants were carefully selected in
consideration of their background, expertise and function
within the organization, so that they could give a
representative view of what would be useful for the actor
which they represented. Formal letters of intent regarding the
contributions from each participating organization were
signed, as well as intellectual property agreements. The
organizational commitment was followed up during the entire
project and secured through adherence to the project plan,
keeping the participants updated, and showing them that
concrete results were achieved which each organization could
benefit from.

The last task of this stage involves clarifying roles and
assigning responsibilities. On the Swedish deregulated
railway market, we identified different stakeholders with
respect to the development of infrastructure-related products:
the infrastructure manager (the Swedish Transport
Administrator (STA)), product manufacturers, contractors
performing maintenance of the products, and academia. In
Sweden, the infrastructure manager typically runs its
development and research projects in cooperation with or
through research institutes and universities, and therefore
academic researchers were relevant stakeholders. All these
stakeholders in the case of turnouts were represented in the
project group of seven participants. Two academics were
given the role of leading the project and managing the
research part, and formed the “scientific team”. The other
participants and one of the academics in the scientific team
formed the “creative team”, in which issues concerning
turnouts would be proposed and addressed by means of the
methods proposed by the scientific team. The scientific team
would analyse the data from the interventions and propose
changes to the method under development. The creative team
would give their views on the methods after the interventions
and comment on any changes which the scientific team would
propose.

3.2. Building, intervention and evaluation (BIE)

Carried out as an iterative process in a target environment,
the BIE phase interweaves the building of the artefact,
intervention in the organization, and evaluation, and the
outcome of the BIE stage is the realized design of the artefact
[5]. This stage is guided by the principles of reciprocal
shaping, mutually influential roles, and authentic and
concurrent evaluation. The principle of reciprocal shaping
implies that increased understanding of the organizational
context influences the design of the artefact, and the artefact
influences the practices in the organizational context. The
principle of mutually influential roles emphasizes the different
types of knowledge which the project participants bring with
them and the mutual learning among the different participants
[5]. The principle of authentic and concurrent evaluation
emphasizes that the evaluation should be on-going and
interwoven with the activities throughout the BIE stage [5].
The evaluation should occur spontaneously in the
organizational context, and opportunities should be sought
following natural controls when possible, rather than in a
controlled environment [5,8].

The first task of the BIE stage is to define an initial
knowledge-creation target. In this project, our target was
structured methods that would encourage information sharing
in a cross-functional inter-organizational group, leading to an
enhanced idea flow and facilitating the selection of the most
promising alternatives.

The next task involves selecting or customizing the form of
BIE to be implemented. According to Sein et al. [5], the form
of BIE to be used should be determined by whether the
innovation is mainly technological or organizational, and they
suggest a generic BIE scheme for each case [5]. In the case of
concept generation and selection methods, we have re-
interpreted this to mean that the form of BIE to be used
should depend on whether the focus is mainly on the method
design or organizational intervention. We customized the
form of BIE to suit our specific case and the given
circumstances, and adapted the BIE according to the
possibilities appearing in the course of the project. As an
example, the BIE scheme adopted in the development of the
cross-functional inter-organizational group.

In the case of the development of the concept generation
method, the iteration cycles started in the creative team (the
“alpha phase”). The requirements for the method emerged
during interviews and discussions within the project group.
Connecting relevant findings from the literature review and
the evidence derived from tests of established concept
generation methods in the creative team [14] with the stated
requirements and demands, tentative design principles were
formulated in parallel with setting up an initial concept
generation method. The suggested method was discussed in
the group and a few changes were introduced before the
performance of interventions. An evaluation of each of the
two BIE cycles was made by means of questionnaires,
interviews, material from the ideation sessions, observations
and transcribed recordings from the ideation sessions. The
creative team was quite satisfied with the method already after
the first iteration, but two changes were introduced to give the
participants enough time in certain steps of the method and to
address the unequal distribution of spoken words among the
participants. After the second iteration, the participants were
satisfied with the method and did not want to make any
further modifications. The method also gave satisfactory

Carried out as an iterative process in a target environment,
the BIE phase interweaves the building of the artefact,
intervention in the organization, and evaluation, and the
outcome of the BIE stage is the realized design of the artefact

In the first BIE cycle, it was quickly found that the findings from the literature review, lead to an improved method. It was suspected that the method had become too cumbersome and that an “encounter with reality” would reveal why it had not been implemented and, together with the practical experience from the design process. The method by Lundberg [15] was developed which incorporates utility theory in a systematic engineering design process. The method by Lundberg [15] was developed with a brief example is given in Lundberg [15]. The test in the beta phase resulted in unexpected results concerning other relevant parameters, such as the number of ideas generated. The successfully introduced changes were conceptualized as design principles and added to the tentative design principles. At this stage, the question arose as to whether the act of taking part in the development of the method would bias the creative team’s satisfaction with it, and it was desirable to try it in a wider context and preferably with several groups. We were provided with the opportunity to do this in an EU project called In2Rail. In2Rail aims to integrate new and advanced technologies into innovative rail product solutions. We were invited to conduct a two day workshop with our methods in one of the work packages concerning turnouts. This step allowed the performance of a comprehensive intervention, corresponding to the “beta phase”, which involved judging the value of the method in a typical use setting. During this beta phase, 17 people from different infrastructure managers, universities and companies participated, representing seven different European countries. Each day, the participants were split into two different groups that applied the method to a total of four different tasks. The evaluation in this phase was performed using questionnaires, observations, recordings and workshop material. The test in the beta phase resulted in unexpected events that gave us new insights into the method. One of them was that the design of the method allowed easy adaption to a specific situation by making small adjustments, leading to the addition of yet another design principle. One important goal of the method design was that it should be possible to use the method directly with participants who were unfamiliar with it or new to ideation methods in general. It turned out that, after receiving the instructions, the groups were able to execute the method with only a little help. According to the questionnaires, the participants found the method to be useful.

In the case of the concept selection method, the starting point was a method that had been proposed by the last author of this paper (a brief example is given in Lundberg [15]). The method was developed from a method by Pahl et al. [16] which incorporates utility theory in a systematic engineering design process. The method by Lundberg [15] was developed for the same context as that described in the present paper, but without interaction with practitioners, and had never been implemented. It was suspected that the method had become too cumbersome and that an “encounter with reality” would reveal why it had not been implemented and, together with the findings from the literature review, lead to an improved method. In the first BIE cycle, it was quickly found that the method was too complicated. Besides being difficult for the participants to grasp, small mistakes in the procedure had a great impact on the outcome in terms of the score which each concept received. It appeared that all the calculations gave a false sense of confidence in the outcome. After the first cycle, the scientific team carefully considered all the data from the intervention and all the results from the evaluation and proposed focusing on life cycle and societal costs, since railway products are typically expensive items that can last for several decades and affect the efficiency with which passengers and freight are transported; this new focus led to a streamlining of the selection procedure. After the second cycle, the scientific team proposed some minor changes to the revised methods, which the participants in the creative team agreed upon, and after that they agreed that they did not need to try the method one more time. The revised method was easier to understand and execute, and the participants found it more useful and, especially, more time-effective; they also appeared to have a more realistic understanding of its advantages and disadvantages.

3.3. Reflection and learning

The reflection and learning stage, a continuous stage that parallels the first two stages, moves conceptually from building a solution for a certain case to applying the learning derived from that process to a broader class of problems [5]. Continuous reflection also allows adjustment of the research process according to increased understanding of the ensemble constituted by the artefact [5]. This stage is guided by the principle of guided emergence, which emphasizes that the ensemble will reflect not only the preliminary design created, in this case, by the scientific team, but also its ongoing shaping by organizational use, different perspectives and the participants, and by the outcomes of authentic and concurrent evaluation [5]. In other words, the reflection and learning stage emphasizes the importance of incorporating the outcome of addressing Principle 1-5 (see Figure 1) in the final artefact.

The first task of the reflection and learning stage is to reflect on the design and redesign during the project. The concept generation and selection methods developed in the present study have been adapted continuously according to the evaluations and analyses that have taken place, to reflect the increasing understanding of both the organizational context and the emerging methods. Task 2 involves evaluating the adherence to the principles and task 3 is to analyse the intervention results according to the stated goals. These two tasks were performed throughout the course of the project and at the end of the project.

3.4. Formalization of learning

The fourth ADR stage involves the formalization of learning and draws on the principle of generalized outcomes, which emphasizes that generalization is a challenge since the artefact has been developed to address a specific situation [5]. Generalization is accomplished by viewing the developed artefact as a solution that addresses a problem, and through making a conceptual move from “the specific and unique to
the generic and abstract” on three levels [5]: 1) generalization of the problem instance, 2) generalization of the solution instance, and 3) derivation of design principles from the design research outcome, i.e. recommendations on how methods addressing the same type of problems should be designed. The first and third task of this stage involve abstracting the learning into concepts for a class of field problems and articulating the outcomes as design principles. The developed concept generation method has been viewed as belonging to the class of concept generation methods for cross-functional inter-organizational groups. Reconceptualizing the learning from this specific instance into design principles for the class of solutions to which it belongs was performed tentatively in the first proposal for the new method. The tentative design principles were confirmed during the alpha and beta cycles, and three additional design principles emerged as a result of the interventions.

The class of field problems for the concept selection method had to be reframed. We realized in the course of the project that the method would be relevant for rather expensive products with a rather long life span to be purchased for the provision of a public utility, or a product that is to be purchased in a large quantity for the provision of a public utility. The design of the product has an impact on the direct costs (e.g. investment and maintenance costs), which are financed by taxes, and the indirect societal costs (e.g. costs originating from a failure to provide public transport for citizens). Therefore, we decided that the concept selection method should belong to the class of methods that facilitate the selection of one or more large-investment products that are to contribute to the provision of a public utility. The design principles were formulated for this class of solutions, emphasizing the importance of a cross-functional team in the selection process.

The second task involves sharing the outcomes and assessment with practitioners. The practitioners involved in the project were continuously updated on its progress and emerging outcomes. A step that belongs to this task, but has not yet been executed, is the implementation of the methods in the routines of STA. Furthermore, the outcome was shared through presentations, workshops and publications.

Task 4 involves articulating the learning in the light of the theories selected in Stage 1 to formalize the results for dissemination. It is an important step to relate the findings to previous research and theories, and this step is not restricted to ADR. The developed artefacts and the learning from the development process have been viewed in the light of the relevant literature identified during Stage 1.

4. Reflections on using ADR

The descriptive and practical nature of the present paper may enable other researchers facing a similar situation in other organizational contexts to understand if ADR can be feasible in their case. ADR has proved to be a feasible framework for the presented project. Through the guidance of ADR, it has been possible to structure the study in a meaningful way. Especially, ADR has contributed an understanding of how the learning derived from the study can be formalized. However, although we found it worthwhile to use the method, there are some issues that we would like to highlight.

Rather than employing one or more control groups, ADR relies on comparing the social situation before an ADR intervention with that after the ADR intervention. In the case of the OptiKrea project, the overarching problem was the lack of interaction during the conceptual design phase on a deregulated railway market. The OptiKrea project can be viewed as a case of the overarching problem, as it was an intervention addressing this problem limited to a group of participants representing different actors involved in turnouts. The participants in the creative team agreed that working in this way in the future would be very beneficial, and according to them the methods should be beneficial for other railway products as well, since the problem in the case of turnouts is very similar to that encountered with other railway products. In the beta phase, the organizational context was reframed as a European organizational context, having previously been framed as a Swedish context. In retrospect, we would have benefited from defining a wider organizational context clearly from the start. The goal of the OptiKrea project was to deliver methods that would be ready for implementation at the infrastructure manager. Implementing the methods at STA is the most appropriate wider organizational context in this case, and one should arrange for a follow-up of the outcome of the implementation. However, it seems very likely that there is now a higher probability of the implementation being successful than if the method had been developed without interaction with practitioners. The concept selection method has so far only been tested in the creative team, and it is very likely that a further need for improvements will be discovered after the implementation.

Applying ADR was successful in that the participants thought that the intervention led to a much better outcome than that which they would normally reach in other kinds of meetings dealing with a similar problem. It was encouraging that the method worked well, but from these interventions we cannot draw the conclusion that the developed method is the “best” method, and can merely conclude that it is one possible solution. There are probably a large number of methods which would have resulted in a better outcome than not using a method at all. We would recommend defining the situation in which interventions are to be made and the wider organizational context at the start of the ADR work, and tentatively defining how to measure the success or failure of the outcome. For instance, in the case of the present project, is it a question of the concept generation method actually generating a higher quantity of ideas than previous strategies, or a question of the participants merely considering the method to be a better and more useful way of working? Moreover, how is this to be measured or assessed? It might be very difficult to measure or assess a former situation (i.e. that existing before intervention) at a later stage. We found that testing the developed method among participants not familiar with it in a typical use setting is very fruitful in judging its feasibility and acquiring additional insights, since the group participating in the development of the method might be at least partly biased.
In the case of the concept generation method, we conducted a more comprehensive preliminary study by testing several established methods and reviewing relevant literature before building the first version of the method. In the case of the concept selection method, we started out with the method developed by Lundberg [15] and, after the first cycle, the findings derived from reviewing relevant literature were incorporated. After the first cycle of the development of the concept generation method in the creative team, only minor changes were introduced in the method. On the other hand, large parts of the concept selection method were eliminated and changed after the first cycle of its development in the creative team. This implies that a thorough preliminary study will be rewarded by a reduction in the work needed between the BIE cycles. However, the changes made in the concept selection method were not radical. We are concerned that ADR might lead to incremental improvements rather than a radical change in the artefact once the first version has been built, although it is impossible to judge this only on the basis of the present case. If this is so, it is even more important to conduct a thorough preliminary study and consider or try different approaches, because once the first version of the artefact is built, the direction of development is determined.

The organizational commitment worked very well. We recommend taking the necessary steps to lay the foundation for a good collaboration between the participants and to align the participants’ goals. We believe that the interacting nature of ADR, in which the participants’ views are continuously taken into account, plays an important part in building this commitment. On the other hand, the scientific researchers should challenge the views of the participants whenever necessary and explain why they are doing so. As the researchers conducting the study and evaluating the scientific part, we believe that it is more important to try to acquire knowledge of the organizational context and its interplay with the artefact on the basis of unexpected events and sudden changes in plan, rather than to try to adhere to the original plan. This exemplifies how authentic and concurrent evaluation, as opposed to evaluation using controlled settings, plays an important role in the application of ADR. However, this makes it more difficult to compare methods to each other, and statistical methods are probably more capable of determining what methods are inherently better. On the other hand, ADR appears to be a better approach to adapting methods to certain contexts while simultaneously understanding how those methods perform in reality, and to identifying context-dependent factors that are not part of controlled experiments. Therefore, we found that ADR and statistical research methods complement each other.

4. Conclusions

From our experience of using ADR in a project developing concept generation and selection methods in close collaboration with practitioners, we draw the conclusion that ADR is a feasible complement to other research methods in the field of engineering design. In our case, we found the main advantages of ADR to be the structure provided, the understanding gained of how the learning from the study can be formalized, and the authentic and concurrent evaluation. To apply the method with success, we recommend taking the necessary steps to secure long-term commitment from the participants, defining the situation where interventions are to be made and the wider organizational context, and specifying how to measure the success or failure of the outcome at the start of the project. Finally, researchers aiming to test ADR should be aware that the solution developed might not be the “best” solution, but might instead be one possible solution optimized for the given context.

Acknowledgements

The financial support provided by Luleå Railway Research Center at Luleå University of Technology is gratefully acknowledged. Moreover, the OptiKrea partners, the Swedish Transport Administration, Vossloh Nordic Switch Systems and Infranord, are gratefully acknowledged for their contributions to the project.

References

[9] Petersson AM, Lundberg J. Developing an ideation method to be used in cross-functional inter-organizational teams by means of action design research. To be submitted.