Using Bayesian belief networks for reliability management : construction and evaluation: a step by step approach

Citation for published version (APA):
Houben, M. J. H. A. (2010). *Using Bayesian belief networks for reliability management : construction and evaluation: a step by step approach*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR675503

DOI:
10.6100/IR675503

Document status and date:
Published: 01/01/2010

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 16. Nov. 2023

Using Bayesian Belief Networks for Reliability Management

Construction and Evaluation: a Step by Step Approach

Using Bayesian Belief Networks for Reliability Management

Construction and Evaluation: a Step by Step Approach

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op woensdag 25 augustus 2010 om 16.00 uur

door

Maurits Johan Henry Antony Houben

geboren te Deurne

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. M.J. Newby
en
prof.dr.ir. A.C. Brombacher

Copromotor:
dr.ir. P.J.M. Sonnemans

Houben, M.J.H.A.

Using Bayesian Belief Networks for Reliability Management – Construction and Evaluation: a Step by Step Approach / By M.J.H.A. Houben. – Eindhoven: Technische Universiteit Eindhoven, 2010.
– Proefschrift –

A catalogue record is available from the Eindhoven University of Technology Library

ISBN 978-90-386-2276-7

NUR 992

Keywords: Product development / Bayesian Belief Networks / Reliability management / Capital goods / Systems analysis

Printed by: University Printing Office, Eindhoven

This research has been funded by the Innovation-Oriented Research Programme ‘Integrated Product Creation and Realization (IOP IPCR)’ of the Netherlands Ministry of Economic Affairs.

Acknowledgements

This thesis is the result of more than four years of research. During these years of research, many people have supported me. I would like to take this opportunity and thank a number of them in particular.

Firstly, I would like to thank my supervisors for their support in this project. I especially want to thank my first supervisor prof. Martin Newby. Despite his very dense time schedule, especially during his visits to Eindhoven, he was always willing to provide feedback. This resulted in numerous phone calls with London in the time when he was not in the Netherlands. Throughout the years, his confidence in my research provided me with the reassurance I sometimes needed. Martin: thanks for all your support! I would also like to thank dr. Peter Sonnemans, my daily supervisor in this research. His help with various problems that turned up throughout the course of my research was invaluable. His ability to put me on the right tracks again, when I came across a dead-end in my research, positively helped me to finish my PhD: thank you! I also want to thank my second supervisor, prof. Aarnout Brombacher for his support of my research. His feedback, looking at my research from a different perspective, has helped me to improve my thesis greatly.

Secondly, I want to thank all the other members of my PhD committee. All comments that prof. Rob Kusters, prof. Martin Neil and prof. John Quigley provided, have led to strong improvements in this thesis, concerning both the quality of the contents, and the readability of the thesis.

I furthermore want to thank those who have facilitated the research. I especially want to thank dr. Guillaume Stollman from Philips Healthcare, for his cooperation and for the opportunity to carry out my research at Philips Healthcare. He was always supportive of my research and as an advisory member of my PhD committee he has provided me with many valuable comments. I would also like to thank all people at Philips Healthcare who have in any way contributed to my research for the time and effort that they made available. Without their help, this research would not have been possible.

I also gratefully acknowledge the support of the Innovation-Oriented Research Programme 'Integrated Product Creation and Realization (IOP IPCR)' of the Netherlands Ministry of Economic Affairs, and I want to thank all people that were involved in the project "life cycle oriented design of capital goods". I am also grateful for the opportunity that was provided to valorise the results of the research.

I would also like to thank the students Jean-Paul Widdershoven, Jaap Houben, and John Visée who contributed to this research. They all have provided valuable insights and John particularly contributed to chapter 6.

Thank you to all the colleagues in the sub department of Quality and Reliability Engineering, and the later sub department of Business Process Design. They include Elke den Ouden, Lu Yuan, Ilse Luyk – de Visser, Jan Rouvroye, Hanneke Driessen, Jeroen Keijzers, Alex Alblas, Aravindan Balasubramanian, Christelle Harkema, Eva Hopma, Kostas Kevrekidis, Aylin Koca, Joël Luyk, and Liesbeth van de Water.

Pap en mam, ik ben jullie heel erg dankbaar voor alle steun die jullie mij hebben gegeven, en voor jullie stimulans om er altijd uit te halen wat er in zit. René en Joris: ik ben blij dat ik twee broers heb op wie ik altijd mag rekenen. Jeroen en René: ik ben er trots op dat jullie mijn paranimfen willen zijn.

Eelke: jou wil ik bedanken voor al je begrip, geduld, vertrouwen en onvoorwaardelijke steun. Zonder jou was het mij niet gelukt om dit proefschrift te schrijven.

Maurits Houben
2010

Summary

Using Bayesian Belief Networks for Reliability Management – Construction and Evaluation: a Step by Step Approach

In the capital goods industry, there is a growing need to manage reliability throughout the product development process. A number of trends can be identified that have a strong effect on the way in which reliability prediction and management is approached, i.e.:

- The lifecycle costs approach that is becoming increasingly important for original equipment manufacturers
- The increasing product complexity
- The growth in customer demands
- The pressure of shortening times to market
- The increasing globalization of markets and production

Reliability management is typically based on the insights, views, and perceptions of the real world of the people that are involved in the process of decision making. These views are unique and specific for each involved individual that looks at the management process and can be represented using soft systems methodology. Since soft systems methodology is based on insights, view and perceptions, it is especially suitable in the context of reliability prediction and management early in the product development process as studied in this thesis (where there is no objective data available (yet)).

Two research objectives are identified through examining market trends and applying soft systems methodology.

The first research objective focuses on the identification or development of a method for reliability prediction and management that meets the following criteria:

- It should support decision making for reliability management
- It should be able to also take non-technical factors into account
- It has to be usable throughout the product development process and especially in the early phases of the process.
- It should be able to capture and handle uncertainty

This first research objective is addressed through a literature study of traditional approaches (failure mode and effects analysis, fault tree analysis and database methods), and more recent approaches to reliability prediction and reliability management (REMM, PREDICT and TRACS).

The conclusion of the literature study is that traditional methods, although able to support decision making to some extent, take a technical point of view, and are usable only in a limited part of the product development process. The traditional methods are capable of taking uncertainty into account, but only uncertainty about the occurrence of single faults or

failure modes. The recent approaches are able to meet the criteria to a greater extent: REMM is able to provide decision support, but mainly on a technical level, by prioritizing the elimination of design concerns. The reliability estimate provided by REMM can be updated over time and is clearly usable throughout the product development process. Uncertainty is incorporated in the reliability estimate as well as in the occurrence of concerns. PREDICT provides decision support for processes as well as components, but it focuses on the technical contribution of the component or process to reliability. As in REMM, PREDICT provides an updateable estimate, and incorporates uncertainty as a probability. TRACS uses Bayesian belief networks and provides decision support both in technical and non-technical terms. In the TRACS tool, estimates can be updated and uncertainty is incorporated using probabilities. Since TRACS is developed for one specific case, and an extensive discussion on the implementation process is missing, it is not readily applicable for reliability management in general. The discussion of literature leads to the choice of Bayesian belief networks as an effective modelling technique for reliability prediction and management. It also indicates that Bayesian belief networks are particularly well suited in the early stages of the product development process, because of their ability to make the influences of the product development process on reliability already explicit from the early stages of the product development process onwards.

The second research objective is the development of a clear, systematic approach to build and use Bayesian belief networks in the context of reliability prediction and management.

Although Bayesian belief network construction is widely described in the literature as having three generic steps (problem structuring, instantiation and inference), how the steps are to be made in practice is described only summarily. No systematic, coherent and structured approach for the construction of a Bayesian belief network can be found in literature. The second objective therefore concerns the identification and definition of model boundaries, model variables, and model structure.

The methodology developed to meet this second objective is an adaptation of Grounded Theory, a method widely used in the social sciences. Grounded Theory is an inductive rather than deductive method (focusing on building rather than testing theory). Grounded Theory is adapted by adopting Bayesian network idioms (Neil, Fenton & Nielson, 2000) into the approach. Furthermore, the canons of the Grounded Theory methodology (Corbin & Strauss, 1990) were not strictly followed because of their limited suitability for the subject, and for practical reasons. Grounded Theory has been adapted as a methodology for structuring problems modelled with Bayesian belief networks. The adapted Grounded Theory approach is applied in a case study in a business unit of a company that develops and produces medical scanning equipment.

Once the Bayesian belief net model variables, structure and boundaries have been determined the network must be instantiated. For instantiation, a probability elicitation protocol has been developed. This protocol includes a training, preparation for the elicitation, a direct elicitation process, and feedback on the elicitation. The instantiation is illustrated as part of the case study.

The combination of the adapted Grounded Theory method for problem structuring, and the probability elicitation protocol for instantiation together form an algorithm for Bayesian belief network construction (consisting of data gathering, problem structuring, instantiation, and feedback) that consists of the following 9 steps (see Table 1).

Table 1: Bayesian belief network construction algorithm

1. Gather information regarding the way in which the topic under discussion is influenced by conducting interviews
2. Identify the factors (i.e. nodes) that influence the topic, by analyzing and coding the interviews
3. Define the variables by identifying the different possible states (state-space) of the variables through coding and direct conversation with experts
4. Characterize the relationships between the different nodes using the idioms through analysis and coding of the interviews
5. Control the number of conditional probabilities that has to be elicited using the definitional/synthesis idiom (Neil, Fenton & Nielson, 2000)
6. Evaluate the Bayesian belief network, possibly leading to a repetition of (a number of) the first 5 steps
7. Identify and define the conditional probability tables that define the relationships in the Bayesian belief network
8. Fill in the conditional probability tables, in order to define the relationships in the Bayesian belief network
9. Evaluate the Bayesian belief network, possibly leading to a repetition of (a number of) earlier steps

A Bayesian belief network for reliability prediction and management was constructed using the algorithm. The model's problem structure and the model behaviour are validated during and at the end of the construction process.

A survey was used to validate the problem structure and the model behaviour was validated through a focus group meeting. Unfortunately, the results of the survey were limited, because of the low response rate (35%). The results of the focus group meeting indicated that the model behaviour was realistic, implying that application of the adapted Grounded Theory approach results in a realistic model for reliability management.

The adapted Grounded Theory approach developed in this thesis provides a scientific and practical contribution to model building and use in the face of limited availability of information. The scientific contribution lies in the provision of the systematic and coherent approach to Bayesian belief network construction described above. The practical contribution lies in the application of this approach in the context of reliability prediction and management and in the structured and algorithmic approach to model building. The case study in this

thesis shows the construction and use of an effective model that enables reliability prediction, and provides decision support for reliability management throughout the product development process from the earliest stages of the process. Bayesian belief networks provide a strong basis for reliability management, giving qualitative and quantitative insights in relationships between influential variables and reliability.

Samenvatting

Using Bayesian Belief Networks for Reliability Management – Construction and Evaluation: a Step by Step Approach

In de kapitaalgoederen industrie bestaat een groeiende behoefte om bedrijfszekerheid te managen gedurende het product ontwikkelproces. Er kan een aantal trends worden geïdentificeerd die een sterke invloed hebben op de manier waarop voorspellingen en management van bedrijfszekerheid worden benaderd:

- De levenscyclusbenadering die steeds belangrijker wordt voor producenten van kapitaalgoederen
- De toenemende complexiteit van producten
- De toenemende klanteneisen
- De toenemende tijdsdruk op het productontwikkelingsproces
- De toenemende globalisering van zowel afzetmarkten als productieprocessen

Bedrijfszekerheidmanagement wordt typisch gebaseerd op inzichten, visies en percepties van de wereld van mensen die betrokken zijn in het besluitvormingsproces. Deze inzichten zijn uniek en specifiek voor elke betrokkene die naar het management proces kijkt, en kunnen worden gerepresenteerd door middel van 'soft systems methodology'. 'Soft systems methodology' is gebaseerd op inzichten, visies en percepties. Dit maakt het bijzonder geschikt in de context van bedrijfszekerheidvoorspelling en -management vroeg in het ontwikkelingsproces zoals wordt bestudeerd in dit proefschrift (waarbij er (nog) geen objectieve data beschikbaar zijn).

Er zijn twee onderzoeksdoelen geïdentificeerd d.m.v. het bestuderen van de trends en d.m.v. het toepassen van 'soft systems methodology'.

Het eerste onderzoeksdoel richt zich op het identificeren of ontwikkelen van een methode voor bedrijfszekerheidvoorspelling en -management die aan de volgende criteria voldoet:

- De methode moet het nemen van beslissingen met betrekking tot bedrijfszekerheidmanagement ondersteunen
- De methode moet in staat zijn om ook niet-technische factoren mee te nemen
- De methode moet bruikbaar zijn gedurende het product ontwikkelproces, in het bijzonder in de vroege stadia van het proces.
- De methode moet in staat zijn om onzekerheid mee te nemen en te adresseren

Het eerste onderzoeksdoel wordt behandeld door middel van een literatuurstudie van traditionele (faalwijzen- en gevolgenanalyse, foutenboom analyse en database methodes) en recentere benaderingen voor bedrijfszekerheidvoorspelling en -management ('REMM', 'PREDICT' en 'TRACS').

De conclusie uit de literatuurstudie is dat traditionele methodes vanuit een technisch perspectief kijken, en dat ze slechts bruikbaar zijn in een beperkt deel van het productontwikkelingsproces (hoewel ze in staat zijn om het nemen van beslissingen (in een bepaalde mate) te ondersteunen). De traditionele methodes zijn in staat om onzekerheid mee te nemen, maar alleen onzekerheid m.b.t. het optreden van individuele fouten of faalwijzen. De recente methodes voldoen in grotere mate aan de criteria, maar voornamelijk op een technisch niveau, d.m.v. het geven van prioriteit aan het oplossen van design problemen. De inschatting van bedrijfszekerheid die wordt gegeven door REMM kan worden geüpdate in de tijd en is bruikbaar gedurende het product ontwikkelproces. Onzekerheid is zowel opgenomen in de inschatting van bedrijfszekerheid, als in het optreden van problemen. PREDICT ondersteunt het nemen van beslissingen met betrekking tot processen en componenten, maar richt zich op de technische bijdrage van de processen en componenten aan bedrijfszekerheid. Zoals ook het geval bij REMM, geeft PREDICT een inschatting die geüpdate kan worden, en neemt onzekerheid mee door middel van het gebruik van kansen. TRACS gebruikt Bayesiaanse netwerken en ondersteunt het nemen van beslissingen met betrekking tot zowel technische als niet-technische aspecten. In TRACS kunnen inschattingen worden geüpdate, en wordt onzekerheid meegenomen door het gebruik van kansen.
Aangezien TRACS is ontwikkeld voor een specifieke case, en een uitgebreide discussie met betrekking tot het implementeren van Bayesiaanse netwerken ontbreekt, is het niet direct toepasbaar voor het managen van bedrijfszekerheid in het algemeen. De bespreking van de literatuur leidt tot de keuze voor Bayesiaanse netwerken als effectieve modelleertechniek voor het voorspellen en managen van bedrijfszekerheid. Het geeft ook aan dat Bayesiaanse netwerken in het bijzonder geschikt zijn in de vroege stadia van het productontwikkelingsproces, omdat ze invloeden van het product ontwikkelproces op bedrijfszekerheid dan al expliciet kunnen maken.

Het tweede onderzoeksdoel is het ontwikkelen van een duidelijke en systematische benadering van het bouwen en gebruiken van een Bayesiaans netwerk in de context van het voorspellen en managen van bedrijfszekerheid.

De constructie van Bayesiaanse netwerken zoals besproken in de literatuur bestaat uit 3 generieke stappen (probleemstructurering, concretisering en gevolgtrekking). Echter, in de literatuur wordt de uitvoering van deze stappen slechts summier besproken. Er kan geen systematische, coherente en gestructureerde benadering van het bouwen van Bayesiaanse netwerken worden gevonden. Het tweede onderzoeksdoel richt zich daarom op het identificeren en definiëren van modelgrenzen, modelvariabelen en modelstructuur.

De methodologie die is ontwikkeld om aan het tweede onderzoeksdoel te voldoen is een bewerking van Gefundeerde Theorie, een methode die gebruikt wordt in de sociale wetenschappen. Gefundeerde Theorie is een inductieve in plaats van een deductieve methode (de focus ligt op het tot stand brengen van een theorie in plaats van op het bewijzen ervan). Gefundeerde Theorie is bewerkt door de idiomen voor Bayesiaanse netwerken (Neil, Fenton & Nielson) op te nemen in de benadering. Verder zijn de canons van Gefundeerde Theorie

(Corbin & Strauss, 1990) niet strikt gevolgd vanwege hun beperkte geschiktheid voor het onderwerp, en vanwege praktische overwegingen. Gefundeerde Theorie is bewerkt als een methodologie voor het structureren van problemen die gemodelleerd worden met behulp van Bayesiaanse netwerken. De bewerkte Gefundeerde Theorie benadering is toegepast in een casestudie in een businessunit van een bedrijf dat medische scanapparatuur ontwikkelt en produceert.

Wanneer de variabelen, structuur en grenzen van het Bayesiaans netwerk zijn vastgesteld, moet het netwerk worden geconcretiseerd. Voor concretisering is een ‘kans elicitatie protocol’ ontwikkeld (protocol om uitspraken met betrekking tot kansen te ontlokken). Dit protocol bevat een training, voorbereiding voor de elicitatie, een direct elicitatieproces, en feedback op de elicitatie. De concretisering vormt een deel van de casestudie.

De combinatie van de bewerkte Gefundeerde Theorie voor probleemstructurering en het kans elicitatie protocol voor concretisering vormen samen een algoritme voor de constructie van Bayesiaanse netwerken (bestaande uit data verzamelen, probleemstructurering, concretisering en feedback), dat bestaat uit de volgende 9 stappen (zie Tabel 1).

Tabel 1: Algoritme voor de constructie van Bayesiaanse netwerken

1.	Het verzamelen van informatie m.b.t. de manier waarop het onderwerp van discussie wordt beïnvloed door het uitvoeren van interviews
2.	Het identificeren van factoren (‘nodes’) die het onderwerp van discussie beïnvloeden door het analyseren en coderen van de interviews
3.	Het definiëren van variabelen door de verschillende toestanden (toestandsruimte) van de variabelen te identificeren met behulp van codering en gesprekken met experts
4.	De relaties tussen de verschillende variabelen karakteriseren door analyse en codering van de interviews en door gebruik te maken van de idiomen.
5.	Het aantal conditionele kansen dat moet worden geëliciteerd beheersen met behulp van het ‘definitie/synthese’ idioom (Neil, Fenton & Nielson, 2000)
6.	Het Bayesiaans netwerk evalueren wat mogelijkerwijs leidt tot herhaling van (een aantal van) de eerste 5 stappen
7.	Het identificeren en definiëren van de conditionele kanstabellen die de relaties in het Bayesiaans netwerk definiëren
8.	Het invullen van de conditionele kanstabellen om de relaties in het Bayesiaans netwerk te definiëren
9.	Het Bayesiaans netwerk evalueren wat mogelijkerwijs leidt tot herhaling van (een aantal van) de eerdere stappen

Een Bayesiaans netwerk voor bedrijfszekerheidvoorspelling en -management is gebouwd m.b.v. het algoritme. De probleemstructuur en het gedrag van het model zijn gevalideerd in de loop van en aan het eind van het constructieproces.

De probleemstructuur is gevalideerd met behulp van een enquête, het modelgedrag is gevalideerd door middel van een focusgroep meeting. De resultaten van de enquête waren beperkt, vanwege het beperkte aantal reacties (slechts 35% van alle geënquêteerden heeft gereageerd). De resultaten van de focusgroep meeting duiden aan dat het gedrag van het model realistisch is, wat impliceert dat toepassing van de bewerkte Gefundeerde Theorie benadering resulteert in een realistisch model voor bedrijfszekerheidmanagement.

De bewerkte Gefundeerde Theorie benadering die is ontwikkeld in dit proefschrift levert een wetenschappelijke en praktische bijdrage aan modelbouw met beperkte informatie en aan het gebruik van modellen in deze context. De wetenschappelijke bijdrage ligt in het verschaffen van een systematische en coherente benadering van het bouwen van Bayesiaanse netwerken, zoals beschreven in Tabel 1. De praktische bijdrage ligt in de toepassing van deze benadering in de context van bedrijfszekerheidvoorspelling en -management en in de gestructureerde en algoritmische benadering van modelbouw. De case studie in dit proefschrift laat de constructie en het gebruik zien van een effectief model dat bedrijfszekerheid voorspelt, en die het nemen van beslissingen voor bedrijfszekerheidmanagement ondersteunt gedurende het productontwikkelingsproces vanaf de vroege stadia. Bayesiaanse netwerken bieden een sterke basis voor bedrijfszekerheidmanagement en geven kwalitatieve en kwantitatieve inzichten in relaties tussen invloedsvariabelen en bedrijfszekerheid.

Table of Contents

List of Abbreviations

ANOVA	Analysis Of Variance
BBN	Bayesian Belief Network
BU	Business Unit
C&D	Concept & Definition
CPT	Conditional Probability Table
D&D	Design and Development
DOC	Direct Operating Costs
DSS	Decision Support Systems
DoE	Design of Experiments
EESA	Extreme Evidence Sensitivity analysis
FMEA	Failure Modes and Effects Analysis
FTA	Fault Tree Analysis
GT	Grounded Theory
LCC	Lifecycle costs
M&I	Manufacturing & Installation
MCDA	Multiple Criteria Decision Analysis
MPD	Marginal Probability Distribution
MTBF	Mean Time Between Failures
NPT	Node Probability Table
OEM	Original Equipment Manufacturer
PCP	Product Creation Process
pdf	probability distribution function
PDP	Product Development Process
PLC	Product lifecycle
PREDICT	Performance and Reliability Evaluation with Diverse Information Combination and Tracking
REMM	Reliability Enhancement Methodology and Modelling
RPN	Risk Priority Number
SE	Evidence Sensitivity
SESA	Subtle Evidence Sensitivity analysis
SODA	Strategic Options Development and Analysis
SP	Parameter Sensitivity
SSM	Soft Systems Methodology
TRACS	Transport Reliability Assessment and Calculation System
TTM	Time to Market

1. Introduction

In the capital goods industry, the focus of the suppliers of the capital goods (original equipment manufacturers (OEMs)) is shifting from only selling a product towards taking care of the upkeep of the system during its entire lifecycle. Marinai, Probert & Singh (2004) state that "'Power by the Hour™' (trade mark held by Rolls-Royce) type of contracts, which includes the capital cost plus a blend of financing and maintenance after the engine's sale, are increasingly being demanded". Also, Walls, Quigley & Marshall (2006) identified that "customers have moved towards a progressive reliability assurance framework embedded within service level contractual agreements". In this context, availability is the main performance indicator of the product and has to be addressed.

Availability of a product throughout its lifecycle is dependent on:

- *Reliability of the product*
  This is a result of choices of the OEM in the process of product definition until product installation in the field (for an elaborate discussion on the product lifecycle (PLC), see section 1.2).
- *Maintenance activities*
  These activities are performed when the product is in the field (after installation).

The focus in this thesis will be on the first stages of the PLC, from product definition until product installation, in order to gain insights in the way in which the OEM may influence and manage reliability in the early stages of the PLC. This scope is in line with the IOP-IPCR project "lifecycle oriented design of capital goods", which focuses on providing techniques to balance lifecycle costs (LCC) and system availability. This can help OEMs have to find a good balance between the costs of a system from design until installation, and the operational costs in the field (e.g. costs of maintenance and upkeep), which together represent a large part of the LCC (Blanchard & Fabrycky, 2006).

The focus on LCC and the relation between reliability and LCC will be discussed in more detail in the next section. Also, a number of trends in the capital goods industry that strongly influence the approach that has to be taken towards reliability management (Magniez, 2007) will be discussed.

1.1. Trends in the Capital Goods Industry

Next to the transition of the capital goods industry from "manufacturing industry" to a "service industry" (putting the focus on an LCC approach, see subsection 1.1.1), Brombacher, Sander, Sonnemans & Rouvroye (2005) identify four trends that have an influence on the way in which reliability is approached. These are:

1. Increase in product complexity (subsection 1.1.2)
2. Increasing customer demands (subsection 1.1.2)
3. Increasing time to market (TTM) pressure (subsection 1.1.3)
4. Increasing globalization, market-wise as well as production-wise (subsection 1.1.4)

In this section, first, the need for reliability management, as a result of the relation between LCC and reliability will be further discussed. After that, the effects of the different trends on the approach towards reliability management will be discussed in different subsections.

1.1.1. Lifecycle costs approach

The current focus on availability and LCC in the capital goods industry introduces both an opportunity (see e.g. Wise & Baumgartner, 1999) and a risk for OEMs. Integrating services in their product offerings creates the need for defining service contracts and performance measures, in this way providing a possibility to improve profitability and generate income for the OEM. However, there is also an important risk involved in integrating service in the product offering. In such a case, high service costs through high maintenance cost and penalties for the lack of availability (Walls et al. 2006) would reduce the income of the OEM. As such, reliability as a cost driver (Ormon, Cassady & Greenwood, 2002; Blischke & Murthy, 2000; Birolini, 2007) for the operational costs in the field creates an urgent need for the manufacturer to consider reliability throughout the PLC and take it already into account early during product development. As Gandy, Jäger, Bertsche & Jensen (2007) identify: ensuring product reliability has to be worked on as soon as possible.

In order to more clearly indicate how reliability influences the LCC, first, the following cost-breakdown structure is introduced (see Figure 1):

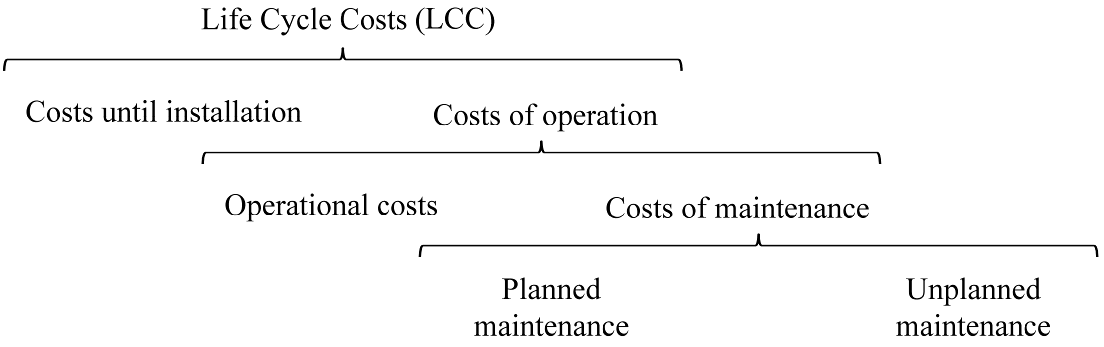


Figure 1: Cost breakdown structure for LCC

From all costs that are shown in the cost-breakdown structure, the costs of maintenance are fully dependent on reliability. In order to keep a system at a certain level of reliability, maintenance has to take place. The level of reliability therefore is dependent both on the reliability that is designed into the system, and the maintenance that is performed to keep a system at a certain level of reliability. The maintenance can be either planned maintenance (e.g. periodical inspection) or unplanned maintenance (e.g. if a component has failed and has to be replaced).

The significance of this fact can be illustrated by the example that is given by Marinai et al. (2004), who take the civil aircraft industry as an example. In their paper (referring to Rupp, 2002), they state that more than one quarter (26%) of the direct operating costs (DOC) for a civil airplane are directly related to the airplane's engine. Furthermore, 31% of the costs related to the engine are related to its maintenance and overhaul. Because of this large percentage of costs that is related to the engine and to its maintenance, it becomes clear that the reliability of the engine plays a very important role, i.e.: if the engine's reliability can be increased, the percentage of the maintenance costs as a part of the costs related to maintenance and overhaul would drop. This would result in a direct decrease of the DOC for a civil airplane, providing the same functionality (reliability), with less maintenance costs. In this way, it is made clear how the 'Power by the Hour™' (trade mark held by Rolls-Royce) type of contracts work, i.e.: the service provider (OEM) provides functionality to the customer, who pays for the functionality per time unit. By increasing the reliability, the OEM provides the same functionality at less (maintenance) cost, leading to higher profits (as long as the increase in costs for improving reliability is smaller than the decrease in costs for maintenance). As a result, in order to address the costs of operation, the reliability of the system has to be addressed.

In particular the maintenance costs that are incurred in the field are strongly dependent on reliability, and present a large percentage of the total costs for many systems (Blanchard & Fabrycky, 2006).

In order to address and manage the LCC (which consist of the costs made during design and production, the costs made in the field and the costs for product retirement and disposal), reliability has to be addressed early in design, since a large part of the LCC of the system is determined already in the early stages of system design (as estimated by Blanchard & Fabrycky (2006)). Reliability is further discussed in section 1.3.

1.1.2. The effect of product complexity and increasing customer demands

Magniez (2007) gives an elaborate discussion on the increasing complexity of products. He identifies a number of reasons for the increase of complexity in products, mainly due to the addition of new functions. The causes for the addition of these functions are:

- The customer may explicitly require new functionality (market pull)
- The OEM introduces new functionality in order to improve product performance, or to make the product cheaper
- In the market, it is important to have a competitive advantage. One way to obtain a competitive advantage, is for the OEM to introduce new technology (technology push)
- In order to stay competitive, the OEM has to offer at least the same functionality as other OEMs that provide a similar product, which may lead to offering functionality that a competing OEM already offers.

One of the consequences of the large increase in functionality that is provided by products is the increase in complexity. This is because the number of interactions increases on a number of levels (Magniez, 2007):

- Between components or subsystems within the product
- Between the user and the product
- Between the product and other products with which it may connect

The increasing complexity leads to an increase in the number of potential failure modes (Magniez, 2007), and makes reliability management more complex and time consuming (Brombacher et al., 2005).

1.1.3. The effect of time to market (TTM) pressure

Time as a business driver has been identified by many (Wheelwright & Clark, 1992; Condra, 2001; Cooper 2001; Meeker & Hamada, 1995). Moreover, Cohen, Eliashberg & Ho (1996) observe that TTM and product performance (including reliability) together define a product's success, but that, at the same time, a trade-off has to be made between these two.
A shorter TTM would give the manufacturer the advantage of being first on the market and create a financial advantage, being able to increase sales, thereby increasing the revenues of the products sold. However, the manufacturer would lose this financial benefit if the performance (amongst others, the reliability) of the product would be insufficient, since this would cause the operational costs to rise.

The increasing TTM pressure creates the need to start addressing reliability already in the early stages of the product development process (PDP), when the system is defined and the technical design constraints are identified (Bedford, Quigley & Walls, 2006) .

The strong impact of time on reliability is also identified by Brombacher et al. (2005) in the context of consumer electronics. They identify that one of the key problems in current PDPs, is the difference between the time required to develop a product, and the time that is needed to collect feedback from the field: whereas the development time has rapidly decreased over time, feedback time has decreased less rapidly (see Figure 2). An important reason for the strong decrease is the rapid influx of new technology, presenting a need to introduce the new technology rapidly in the market. Because the feedback time has not decreased accordingly, a large source of uncertainty is created.

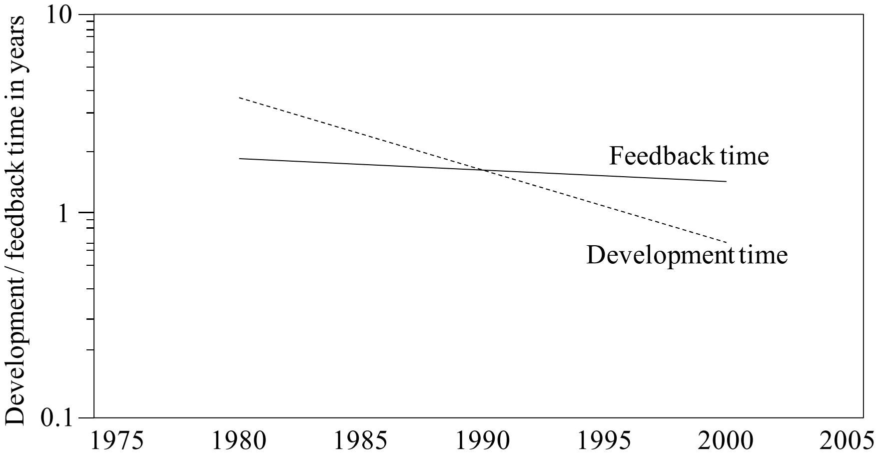


Figure 2: Development time versus feedback time for high-volume consumer electronics (Brombacher et al., 2005)

Improving the feedback loop by increasing the speed and efficiency of the generation of reliability data is identified by Brombacher et al. (2005) as a potential research area, in order to decrease the uncertainty that is related to the lack of feedback. Petkova (2003) identifies that there is a need for the feedback on reliability to be generated faster and more reliable. In this way, the uncertainty that exists because of the increasing gap between the speed of development and the speed of field feedback can be addressed.

Currently, trends in the capital goods industry like real-time monitoring of systems enable the OEM to collect data real-time from the system and its performance in the field. However, although the speed of feedback has increased enormously, the usability and adequacy of these data still leaves a lot to be desired.

1.1.4. The effect of Globalization

The effect of globalization on reliability goes through the outsourcing of efforts in the PDP and through seeking external suppliers of components for products (Magniez, 2007). In the case of effective integration of the suppliers and external parties in the PDP, outsourcing can have a strong positive impact on resulting product quality, cost, and cycle time (Ragatz, Handfield & Scannell 1997). As they state: “Effective integration of suppliers into the value/supply chain will be a key factor for some manufacturers in achieving the improvements necessary to remain competitive”. This also implies that an ineffective integration of the supply chain may negatively impact product quality, thereby negatively influencing product reliability.

1.2. Reliability Management in the PDP

Product support for reliability (in the form of planned maintenance) plays an important role in the context of LCC, since the maintenance costs are an important part of the LCC, as shown in subsection 1.1.1. Therefore, both the reliability that is designed into the product (inherent reliability) and maintenance have to be considered, taking into account that the maintenance that has to be performed is dependent on the reliability that is designed into the product (inherent reliability). Consequently, the performance targets as they should be taken up in the service contract, relate to reliability as a combination of the planned maintenance and the inherent reliability.

Since the focus in this thesis is on creating a certain level of reliability that is a result of the PDP, reliability management has to focus on the PDP rather than the use phase of the product. In this way, the time at which reliability affects the LCC (during use) and the time at which reliability is addressed (through reliability management in the PDP) are separated in this thesis. To get a clear view on this situation, the PLC (consisting of both the PDP and the use phase) will be discussed, addressing the PDP in more detail.

Many different PLC and product development models are available in literature. Using the PLC as described in Blanchard & Fabrycky (2006), the position of the PDP within the PLC will be made clear.

Blanchard & Fabrycky (2006) describe the PLC as consisting of four general stages, i.e.: conceptual/preliminary design; detail design and development; production and construction;

and finally product use, support, phase-out and disposal. The PLC is initiated by a certain “need” that is identified, and for which a product is to be developed. That need is translated into a concept, after which a preliminary design is further developed into a detail design. Finally, the product is ready to be produced and/or constructed, after which it enters the use phase. This phase consists of product use, support, phase-out, and finally, disposal. Blanchard and Fabrycky take the perspective of the customer, rather than the manufacturer, making a distinction between the acquisition phase and the utilization phase. The former consists of the first three stages mentioned above, the latter representing the fourth stage mentioned above. Summarized, this leads to the following representation of the PLC as described by Blanchard & Fabrycky (2006) (see Figure 3):

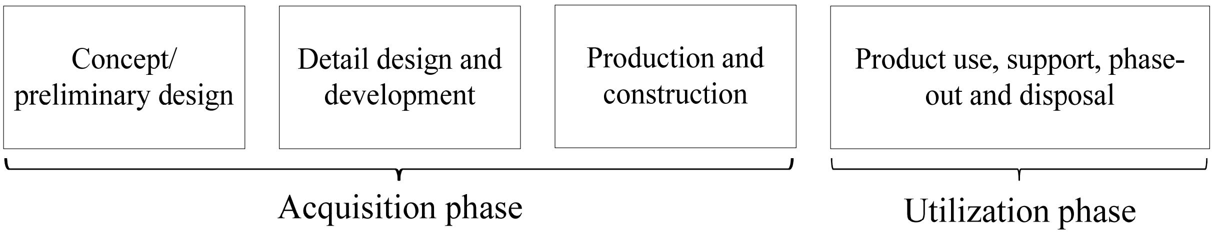


Figure 3: PLC according to Blanchard & Fabrycky (2006)

The discernment between the acquisition phase and the utilization phase represents the distinction between the activities associated with the producer, and the activities associated with the customer, respectively. The first three stages of the PLC as described by Blanchard & Fabrycky (2006) may be denoted as the PDP. Also the PDP as defined in this thesis is defined by these three stages.

Blischke & Murthy (2000) make a distinction between factors prior to the sale of a product (Acquisition phase; factors that can be reasonably controlled by the manufacturer) and factors during use (Utilization phase; factors over which the manufacturer has little (or no) control, but which are influenced by the user). Murthy, Rausand & Østeras (2008), describe the acquisition phase as a product development model. They also present a number of alternative models for representing the PDP; taking different perspectives (see Murthy, et al. (2008) for an overview). From the perspective of the manufacturer, the most interesting stages of the PLC are the stages that represent the PDP, since the manufacturer is able to reasonably control factors influencing the products performance (and hence, reliability) in these stages.

In this thesis, the terminology and definition of the PDP is used that is found in international standards, (IEC 60300 series) as described and referred to by Bedford et al. (2006):

- Concept and definition (C&D)
- Design and development (D&D)
- Manufacturing and installation (M&I)

Concept and Definition (C&D):

In this stage, the technical constraints of the product are defined, based on customer wishes, taking the feasibility of the design into account. Trade-off studies take place with respect to cost-effectiveness and feasibility.

Design and Development (D&D):

In this stage, the system is fully built, tested and refined. If needed, the systems specifications may be adjusted. Functionality of the integrated system (validation) and conformance of subsystems interfaces (verification) is also warranted.

Manufacturing and Installation (M&I):

In this stage, the product is being replicated, and the focus is on controlling the M&I process. More validation and verification takes place here.

Ultimately, inherent product reliability as studied in this thesis – reliability as far as it can be determined by the manufacturer before the product is used by the customer – is defined in these three stages. This is confirmed by Blischke & Murthy (2000), who state that reliability from a manufacturer's perspective can be managed throughout the first three stages (C&D, D&D, M&I) of the PLC.

1.3. Reliability Prediction for Reliability Management

It is stated in literature that reliability improvement (reliability management) is more important than reliability prediction (Davis, 1998), but also that reliability improvement and reliability prediction are very closely related. As Østeras, Rausand & Murthy (2008) state: "the predicted (reliability, red.) performance forms the basis for decisions during the different phases of the product life cycle". In order to discuss the topic of reliability management, first the topic of reliability prediction has to be discussed. After that, the concept of reliability will be elaborated on, in order to identify a suitable definition for reliability in the current context.

1.3.1. Reliability prediction

Reliability prediction "deals with evaluation of a design prior to actual construction of the system" (Blischke & Murthy, 2000). Although it is often related in literature (see e.g. Blanchard & Fabrycky, 2006) to the quantitative measure of reliability, reliability prediction is applicable to both the qualitative and the quantitative measure of reliability. In this thesis, reliability prediction is not limited to quantitative reliability prediction.

Reliability prediction is very important in many different ways (Blischke & Murthy, 2000), especially for large, complex systems, such as capital goods. In literature, many purposes for reliability prediction can be found that are closely related to reliability improvement (and management) (see e.g. Blischke & Murthy, 2000; Healy, Jain & Bennett 1997; Murthy et al., 2008; Yadav, Singh, Goel & Itabashi-Campbell, 2003). These include (but are not limited to):

- performing trade-off studies
- planning for design improvements
- cost analyses

Looking at the information from the first chapter, it becomes clear that prediction of reliability already in the early stages of the PDP is a necessity to manage the reliability and the associated costs. In general, two objectives are related to reliability predictions:

1. Predictions are made in the early stage of the PDP to ensure that the product reliability is good enough
2. Predictions are made in order to plan for product support for reliability in the use stage of the product.

In order to make a clear distinction between the prediction of product reliability taking information from the use phase (as well as product maintenance) into account, and the prediction of reliability in the early stages of the PDP not taking information from the use phase into account, the former will be referred to as *reliability forecasting*, whereas the latter will be referred to as *reliability prediction*. The estimation of reliability will also be referred to as *reliability prediction*, if it encompasses both types of reliability estimation methods. Whereas reliability prediction – as it applies in the early stages of the PDP – reasons from the lack of objective data, reliability forecasting is also based on objective (feedback) data. In this thesis, the focal point will be reliability prediction for reliability management, since the focus lies on reliability management in the early stages of the PDP using qualitative estimates of reliability. It is important to note that reliability prediction in this context does not concern quantitative estimates. Rather, it concerns predicting how the resulting reliability will change, when the values of variables that affect the reliability change. In this way it is directly related to reliability management.

1.3.2. Reliability: definition

The definition of reliability has been debated much in literature. In general, two different definitions of reliability can be given, focusing either on the producer's perspective, or on the user's perspective. From the producer's perspective, reliability is defined as a characteristic of an item (amongst others Lewis (1996), Birolini (2007)), i.e.: "the probability that a system will perform its intended function for a specified period of time under a given set of conditions" (Lewis, 1996). Implicit in this definition is the assumption that unreliability is caused by a product failure, and that product failure is an unambiguous concept, since the required function, the conditions and the time interval are explicitly mentioned and defined prior to use.

However, recent research has shown that the users of a product not only complain about technical failures, but also about non-technical failures (Brombacher et al. 2005, Den Ouden, 2006). Their research underpins the idea of Meeker & Escobar (2004), who propose to change "given conditions" into "encountered use conditions" as a more appropriate definition for reliability, hereby introducing the user's perspective on reliability.

In this thesis, reliability is looked at from the manufacturer's point of view, and the way in which reliability is determined in the early stages of the PDP. In reliability prediction, as it was defined in the previous subsection, the information from the use phase of the system was not taken into account. In this thesis, the focus on reliability lies on reliability as a result of the PDP, rather than on reliability as a characteristic of the technical product design. Rather than rigidly defining reliability, reliability is addressed through the aspects that define reliability, i.e.:

- Dependability
- Successful operation
- Absence of breakdowns and failures.

1.4. Stakeholders involved in the Process and Decision Support

In order to provide support for reliability management, the approach towards reliability management has to be able to support decision making. In order to identify the type of decisions that has to be supported, the stakeholders in the reliability management process have to be identified. In general, three levels of decision making can be identified: strategic, tactical and operational decision making (Eom, Lee, Kim & Somarajan, 1998). At a strategic level (strategic planning (Gorry & Morton, 1971)) decision making is concerned with broad policies and goals for the organization. Strategic decision making focuses on decision making at a high level in the company. In this context, the organizations relationship with the environment is very important. In contrast, on operational level, decision making relates to the day-to-day activities. Operational decision making focuses on decision making at a low level in the company. The related levels of decision making are represented in Figure 4.

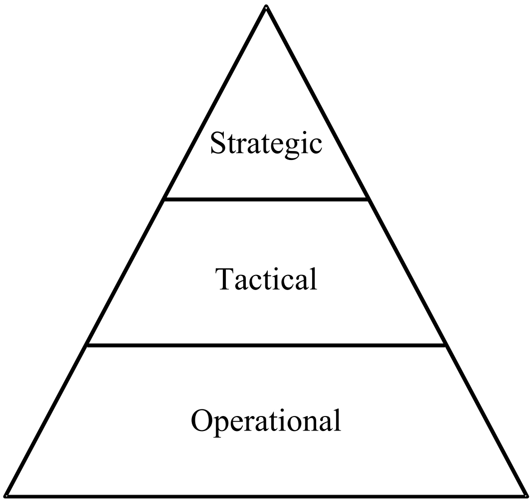


Figure 4: Three different management levels at which decision making takes place

The focus in this thesis lies on the way in which reliability is created throughout the PDP, rather than on reliability as a product characteristic. Therefore, decision support does not focus on decisions regarding the company policies and strategies on the long term, nor does it focus on the decisions at an operational level (decisions made on a day-to-day basis). Rather, it focuses on decision support on the tactical level. The relations between the different levels of decision making can be presented in an example from logistics planning, taken from Ballou (1999). On the different levels of decision making, a decision that is taken regarding purchasing could be the following:

- Strategic: What the policy is regarding purchasing
- Tactical: Vendor selection, and contractual agreements
- Operational: Order releasing

The stakeholders, i.e.: the people that perform decision making on a tactical level, have to be provided with a model as a sound basis for decision support. This model should give insights in the way reliability is ‘created’ throughout the PDP. In this way, they can relate the decisions that are made throughout the PDP to the reliability of the end-result.

At the same time, in order to construct a model for reliability prediction, information is needed from a lower level (operational level) in the organization. Such a model provides input for the stakeholders within the company at a tactical level in two ways:

1. The model provides insights in the processes that ‘create’ reliability and their interactions
2. The model provides a prediction of reliability, gaining insights in the importance of the different factors that affect reliability.

As a result, the model focuses on aspects of the processes that create reliability, and looks beyond the technical, physical aspects of a single system. In this way, such a model provides a means for decision makers on a tactical level to perform reliability management.

1.5. Problem Definition

There are four issues that lie at the root of the problem that is identified in this chapter. These are:

1. *The LCC approach that is becoming increasingly important for OEMs.*
   The current focus on availability and LCC in the capital goods industry introduces both an opportunity (see e.g. Wise & Baumgartner, 1999) and a risk for OEMs, and creates a need to predict and manage reliability from the early stages of the PDP onward.
2. *The increasing product complexity and increasing customer demands.*
   As Magniez (2007) states: products contain an increasing number of interfaces, which makes prediction and management of reliability more complex.
3. *The increasing TTM pressure.*
   The TTM pressure creates a need to take reliability into account already in the early stages of the PDP, during which adequate data for reliability prediction and management are not available. This results in a large uncertainty early in the PDP and creates a need to take information sources outside objective, empirical data into account.
4. *The increasing globalization, market-wise as well as production-wise.*
   The effect of globalization on reliability goes through the outsourcing of efforts in the PDP. Outsourcing can play an important role in determining reliability, as it has a strong impact on resulting product quality, cost, and cycle time

Because of these issues, the problem definition is formulated as follows:

Problem definition

The problem in reliability management arises, because there is no reliability prediction method available that is able to predict reliability early in the PDP, and support decision making at a higher (tactical) management level, when no adequate data are available, and when complexity of the problem context is high, due to the large number of factors affecting reliability and the large number of interactions and their diversity (product related, process related or a combination of both) that are present in the PDP.

Note in this problem definition that reliability prediction also addresses reliability management and that reliability management in this sense is not possible without reliability prediction.

In the next chapter, this problem definition will be further discussed, resulting in research objectives and the related research questions are presented. However, first, the outline of this thesis is presented underneath.

1.6. Outline of the Thesis

In this section, the outline of the thesis is discussed. By looking at the contents of each subsequent chapter, the thread of the thesis is presented.

In Chapter 2, an approach towards reliability prediction is identified, which enables reliability management. Furthermore, a number of criteria for reliability prediction are identified, which lead to the definition of two research objectives. In this chapter, both traditional and recently developed reliability methods will be discussed, leading to the identification of a suitable modelling approach for reliability prediction, as well as related research questions.

In the third chapter, the general modelling approach that is chosen for reliability prediction is discussed, consisting of three steps. For all steps, a discussion is presented on literature describing the way in which the steps can be performed, as well as the way in which the steps may be validated.

Chapter 3– 6 describe the model construction process in more detail. Moreover, the proposed solutions for the difficulties related to this process are elaborated. Based on the challenges identified in chapter 3, both the application and the result of the proposed solution are discussed, together with the validation of the approach and of the resulting model. At the end of chapter 6, the algorithm for Bayesian Belief Network (BBN) construction is presented that is used in the case study that is performed in this research.

In chapter 7, a discussion on the way in which BBNs can be used through Bayesian inference, and results of the method can be evaluated is presented. The use of Bayesian inference for reliability prediction and decision support is discussed, as well as validation of the model behaviour through a focus group meeting.

In chapter 8, generalization of the research is elaborated, addressing the research objectives and related research questions. Moreover, the conclusions and research contribution will be presented and discussed. Finally the approach towards BBN construction in a reliability management context will be reflected upon, and recommendations for future research will be presented.

2. Reliability Prediction Methods: Discussion and Choice

This chapter discusses different reliability prediction methods and related issues, focusing on the problem definition as described in the previous chapter, and addressing its elements. In order to do this, first, the overall modelling approach that is chosen for reliability management will be presented in section 2.1. At the end of section 2.1, this results in the two research objectives. After that, a number of criteria for reliability prediction methods are identified, based on the overall modelling approach and the industry trends. A number of existing reliability prediction methods are then discussed and evaluated with respect to these criteria in sections 2.3 and 2.4. This results in the identification of the proposed modelling approach and the research questions in sections 2.5 and 2.6 respectively.

2.1. Reliability Prediction through Systems Theory

Reliability prediction in this thesis is directly related to reliability management. Therefore, a modelling approach has to be chosen that is able to address the management aspect of reliability prediction.

In this light, a systems thinking approach is of direct relevance to problem solvers and decision makers, since systemic thinking is, as stated by Flood & Jackson (1991), a “vehicle for creative and organized thought about problem situations”. This shows the value of systems thinking in the context of reliability prediction and management, i.e.: systems thinking has been successful as a tool for problem management (Flood & Carson, 1990). As identified in e.g. Blair & Whitston, (1971) (referred to by Checkland, 1989) there are three general systems approaches towards modelling:

- Natural systems: systems that are created by nature
- Designed systems: systems that are by Man
- Human activity systems: a system in which humans try to take purposeful action

Typically, management involves taking actions in order to control the output of the process, based on the insights of the decision makers, and can be typified as a human activity system. This means that management is typically based on the insights, views, and perceptions of the real world, of the people that are involved in the process of decision making. The views of the people that are involved are unique and particular for each individual that looks at the management process. This is represented by the meta-view that is provided by Soft Systems Methodology (SSM).

2.1.1. Soft Systems Methodology (SSM)

The observation that SSM is well applicable in a management context (when problem boundaries and problem structure are less clear) is endorsed by (amongst others) Checkland (2000), Jackson (1991), Checkland & Scholes (1990). As Checkland & Holwell (1998) state:

(SSM) is “a methodological approach to tackling real-world problems”. The choice for SSM as the basis for our approach to reliability management is strengthened by the statement made by Munro and Mingers (2002) (referred to by Mingers (2006)): “Soft Systems Methodology (SSM) the most well-known and successful systems methodology available.” Application of the SSM approach will be discussed in more detail in the next subsection.

The first developed form of SSM was developed by Checkland (Checkland, 2000), in his book “Systems Thinking, Systems Practice” (Checkland, 1986). In this book, Checkland introduces a first form of SSM as a seven step process. In this process, first the real world is studied to get a good view on the problem situation, and to express the problem situation elaborately (steps 1 and 2). Then, the problem is looked at from a more abstract point of view, expressing the problem as a system and modelling it, leading to a “systems model”. This includes the identification of system boundaries, and system elements (steps 3 and 4). In the last three steps, the model is compared to the problem situation as described in the first two steps, after which actions are first identified, and then applied, to change (improve) the situation. As a result of the actions, the situation changes, and the SSM process can be applied again. Depending on the results of the actions, renewed application of SSM may result in small or substantial changes in the system model.

The seven stage process for applying SSM is the basic, most well-known process model for applying SSM. Unfortunately, it does not fully represent the nature of SSM (Checkland & Scholes, 1990). This is because the model implies that the process of applying SSM is a simple sequential process, resulting in actions to change the situation into a more desirable one, whereas in reality, the application process is not necessarily purely sequential. Nevertheless, the seven steps procedure as described in Checkland (1986) (see Figure 5 in the next subsection) is used as reference model. The seven steps of the conventional SSM procedure will be described in more detail underneath, again using Checkland (1986) as reference. Also, the way in which the procedure can be used in the context of reliability prediction early in the PDP is described. It is important to note that the SSM approach describes a number of steps that have to be taken to build a model. The way in which these steps are completed is not described, and is left as a choice for the model builder.

2.1.2. Application of SSM

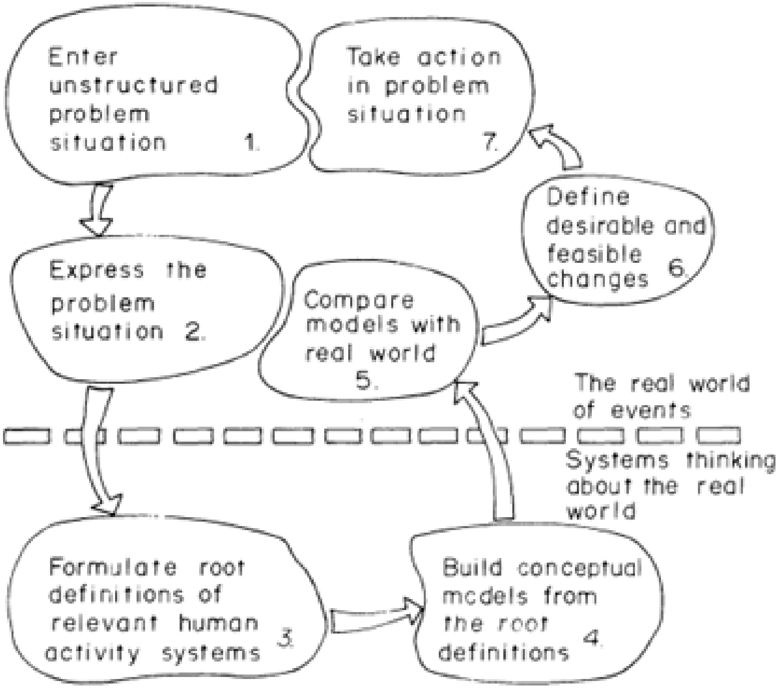


Figure 5: SSM as a learning methodology (Checkland, 1986)

Steps 1 and 2: Expressing the problem situation

In the first two steps, the problem and the problem situation have to be studied in detail. A very important aspect of the first two steps in the 7 step SSM procedure is the fact that the problem situation is studied extensively, without imposing a structure, or pre-defining the boundaries. It is important during these steps, to keep some distance between the researcher and the problem situation. The importance of the systems (an entity, representing e.g. an organization, an activity or a set of activities) that are identified in these two steps should be addressed as well, since it plays an important role in the continuation of the 7 step procedure.

In this case of reliability prediction in the PDP, the problem situation pertains to the 'creation' of reliability in the PDP. It is therefore important to get a broad and comprehensive view on the PDP, from the perspectives of participants, i.e. the different stakeholders and from the people from the operational level that provide the information.

Steps 3 and 4: Translating the problem situation into a systems model

In step 3 and 4, the extensively studied problem situation from step 1 and 2 is modelled. It is important here, that the systems that are identified to be important in the first two steps are well defined in the third step. Not only the definition of the systems is important, but also the nature (a physical entity, a sequence of activities) of the systems that are defined. Defining the systems in step 3 cannot be seen independently from building the conceptual model in step 4, and, later in the application of SSM (steps 6 and 7) the proposal and performance of actions directed to achieve feasible and desirable changes. This is because the actions pertaining to the feasible and desirable changes are directly related to the systems. Therefore, the systems have to be defined such that they are changeable in order to achieve changes at a higher level. In the case of reliability prediction and management, the systems can be described as the processes related to the "development" of product reliability.

Step 4 concerns the actual building of a conceptual model. Checkland identifies two generic ways of describing a system. The first possibility is that the system is described through its related elements, relationships and conditions. This is often the case when describing a physical system. The second way of describing a system is by describing it as a system that transforms inputs into outputs. The latter way of describing a system relates often to a human activity system. Since the model is based on the input of the participants, the model is based on the perceived problem situation, i.e.: the model reflects the perception of these people. Therefore, the model does not necessarily represent real physical processes, although it is able to do so.

It can be seen that the PDP itself looks like a combination of both a physical system, and a system consisting of activities: it can be looked at as being a process as such with certain characteristics, and it can be looked at as a series of activities that transform a conceptual idea into a product with certain reliability. Since describing the system as an entity with inputs and outputs provides the opportunity to incorporate also physical aspects of the system (Checkland, 1986), the soft systems approach, describing the system as an entity transforming inputs into outputs is the most appropriate description of the PDP (since it contains both physical aspects and activities). Furthermore, the goal of the model to be developed in this thesis is reliability management. Since this goal cannot be very clearly defined, a hard systems approach is inappropriate.

Step 5: Comparing the systems model with the problem situation

In this step, the constructed systems model is compared to the real problem situation. This has to be done in the problem situation, together with the participants. Not only does this approach give an opportunity to partly validate the model, but the discussion can also be used to identify possible beneficial changes. Since the model is based on perceptions of people, validation of the model is possible in a subjective sense, comparing the systems model with the perceived problem situation.

In the case of reliability prediction, the PDP is studied from the perspective of the people involved in the process. This creates the opportunity to create a model that is based on and reflects the perception/view of the people involved.

The comparison of the model to reality could be seen as a certain type of validation, although not in the classical sense (it is not a "proper comparison of like with like" (Checkland, 1986, referring to Anderton & Checkland, 1977). The comparison is based on discussion of the model builders with the participants on the possible feasible and beneficial changes that may take place. In this sense, the comparison of the model with the real world seems comparable to face validation (McCall & Lombardo, 1982): people commenting on the realism of simulation outcomes.

Steps 6 and 7: Proposing actions towards desirable changes and implementing actions

In these two last steps, changes to the system are proposed that are both feasible and desirable, in order to improve the problem situation. According to Checkland (1986), three types of changes can be applied:

- changes in system structure
- changes in procedures
- changes in 'attitude'

Structural changes are changes in phenomena that do not change within the short term. Procedural changes are related to procedures in the process; the dynamic aspects of the system. Finally, changes in 'attitude' are the changes that are related to the more intangible characteristics of the system.

With respect to the PDP, the different types of changes all play a role. Although the PDP itself is a process, and therefore might be more susceptible to procedural changes, the PDP is embedded in a company, and performed by people. Hence structural changes (related to the way in which the PDP is embedded in the organization) and 'attitudinal' changes (changes in the behaviour of the people involved in the PDP) may also play a role.

Iteration

After step 7, the problem situation has changed because of the changes implemented. This initiates a repetition of the 7 step SSM procedure. The changes implemented in step 7 define the extent to which the modelled system will change: whether the whole model has to change, or only parts of the system model have to change to represent the new situation.

Application of SSM: issues and observations

A number of observations can be made with respect to the application of SSM. SSM offers a lot of possibilities and is very flexible. One of the characteristics of SSM that is related to its flexibility, is that it is able to incorporate both tangible (physical) and intangible elements. However, at the same time, regarding application of SSM, it is not possible beforehand to identify and define the system and its elements and to express it as such (the first four steps of the application of SSM). Although guidelines are given in literature, it is also not clear how these steps can be implemented.

Although SSM clearly focuses on human activities, it is also able to incorporate physical elements. This leads to difficulties regarding the identification of system boundaries and elements. The identification of system boundaries and the identification of the variables and their relations that are to be included in the model have to be taken into account when looking at the implementation process of SSM in the context of reliability prediction and management.

2.1.3. Research objectives and research questions

As a result of the problem definition, and using SSM as modelling approach, two main research objectives are formulated, together with their corresponding research questions:

Research objectives:
1. To find or develop a reliability prediction and management method which is able to address the criteria that result from the increasing attention that is paid to LCC, as well as the increase in complexity, customer demands, TTM pressure, and globalization
2. To provide a systematic approach for applying the reliability prediction method developed in industry

Because of the current trends in the capital goods industry, methods that may be used in the context of reliability prediction and management have to meet a number of criteria. These criteria first have to be identified from the trends. Therefore, the first research question is as follows:

Research questions 1.1:
Which criteria can be identified for reliability prediction and management methods that relate to the increasing attention for LCC, increasing complexity, customer demands, TTM pressure and globalization?

After that, a method has to be sought that is able to meet these criteria, and at the same time, the method has to be usable for reliability prediction and management. Therefore, the second and third research questions are:

Research questions 1.2 and 1.3:
Is there a modelling method that can meet these criteria?
Can this method be applied for reliability prediction and management?

To address the second research objective, the first research objective has to be addressed first, i.e.: a reliability prediction method has to be found or developed which is able to address the criteria that result from the increasing attention that is paid to LCC, the increasing complexity, and the increasing TTM pressure. The research questions that are related to the second research objective are related to the choice of SSM as modelling approach.

The first research question deals with identifying and defining the boundaries of the model which specify the scope of the model regarding the defined problem. Therefore research question 2.1 is:

Research questions 2.1:
How can the system boundaries of a structural model for reliability prediction be identified and defined systematically, when no information regarding the boundaries is available beforehand?

The variables (elements) of the system that are to be included in the model have to be identified and defined as well. Therefore, research question 2.2 is formulated as follows:

Research questions 2.2:
How can the variables that are to be included in the structural model for reliability prediction be identified and defined in a systematic way?

Finally, the relationships between the system elements have to be identified and defined, leading to research question 2.3:

Research questions 2.3:
How can the structure of the model for reliability prediction be identified and defined in a systematic way?

Research objective 1 will be addressed in this chapter, leading to the identification of a suitable reliability prediction and management method. The application of this approach is addressed in the following chapters.

2.2. Criteria for Reliability Prediction Methods

Based on the previous chapter, a number of criteria can be identified that reliability prediction methods have to meet. These criteria will be discussed in this section. The criteria are related to the three issues that lie at the root of the problem definition (increasing attention for LCC, increasing complexity, customer demands, TTM pressure and globalization). The criteria that can be distilled from these three issues are the following:

1. The method has to give insights in order to support decision making
2. The method has to take non-technical factors into account
3. The method should be usable throughout the PDP
4. The method should be able to incorporate uncertainty

For every criterion, the way in which the criterion is related to the trends in the capital goods industry will be discussed in the next four subsections. Next to this, the consequences of the criteria for the reliability prediction model are discussed.

2.2.1. Insights for decision support

In order to manage LCC, the reliability prediction method should support reliability management. Therefore, the method should be able to give insights to support decision making.

This indicates that the model should be structural rather than black box (Mitchell, 1993), since insights in the cause and effects relations are more important for decision support than an exact quantitative reliability prediction per se.

2.2.2. Non-technical nature

Because the complexity of the product context is high, reliability problems that occur are often of a non-technical nature. Therefore a purely technical and physical approach to reliability modelling does not suffice for predicting and managing reliability. This leads to the conclusion that a reliability prediction method should not only take technical factors into account, but should also include non-technical factors related to the PDP.

Since every organization organizes their PDP in their own way, models that are created are relative rather than absolute (Mitchell, 1993). This means that the models are valid for the PDP for which they are developed, but that they are not valid in general.

2.2.3. Usability throughout the PDP

Because of the TTM pressure, reliability has to be managed from the early stages of the PDP onwards. This means that reliability management has to start already before product design. It furthermore implies that reliability management has to be based on more than only technical system specifications and objective, empirical data, and that it should use other sources of information, like expert judgment, as well (based on availability of information).

It should be possible to include information that emerges throughout the PDP in the reliability prediction as it becomes available. Since the nature of the information that is available early in the PDP varies, it should be possible to deal with diverse types of information.

2.2.4. Information unavailability: epistemic uncertainty

As identified in the previous subsection, reliability has to be managed already in the early stages of the PDP. Epistemic uncertainty is a direct consequence of the lack of availability of adequate data in these early stages of the PDP (Ayyub & Klir 2006; Hora, 1996; Bedford & Cooke, 2001). The reliability prediction and management method has to take this uncertainty into account.
Next to epistemic uncertainty, also aleatory uncertainty can be distinguished, which is inherent randomness that arises through natural variability of the system (and therefore does not decrease when information becomes available). The difference between epistemic and aleatory uncertainty is mainly “for the purposes of a particular model” (Bedford & Cooke, 2001). Whereas epistemic uncertainty is explicitly included in the model, aleatory uncertainty (“uncertainty about which we cannot or choose not to learn” (Bedford & Cooke, 2001)) is included in a model implicitly through the degree by which reality is simplified.

The most widely used means by which uncertainty is modelled is probability theory (see e.g. Schmitt, 1969; Lindley, 1975; Winkler, 1996; Cowell, Dawid & Lauritzen, 1999; Oberkampf, Helton, Joslyn, Wojtkiewicz & Ferson, 2004; Chapman & Ward, 2000, Bedford & Cooke, 2001). Therefore, the discussion on modelling uncertainty includes a discussion of probability theory on a basic level.

Throughout the years, many different viewpoints on probability have appeared. In the book of Neapolitan (2004) a major distinction is made between the 'Relative Frequency approach' and the 'Subjective/Bayesian approach'. The two approaches can be summarized as follows:
The frequentist approach, (as the relative frequency approach is also called) is based on an interpretation of probability as being the proportion that is obtained in an infinite sequence of experiments (Von Mises, 1939, as referred to by Cowell et al., 1999).
In the subjective (Bayesian) approach, probability is represented as a "degree of belief", also called a subjective probability (Neapolitan, 2004) It represents your belief in a certain outcome happening. The reason why the subjective probability is called Bayesian is because it uses Bayes' theorem to infer unknown probabilities from known probabilities.

Although the frequentist view and the subjectivist view are presented differently, Cooke (1991) argues that they are not different from each other. Moreover, he states that taking a frequentist view does not mean that the probabilities that are used are somehow more "objective". In many cases, subjectivist probabilities are based on experiences in the past, which indicates a relative frequency approach. Cowell et al. (1999) also identify this phenomenon, stating that a subjective probability "must be relative to that person's degree of belief or background knowledge". Moreover, Winkler (1996) states: "…in assessing subjective probabilities, an individual can certainly utilize, as appropriate, any relative frequency data…" In this thesis, the comprehensive view on probability is adopted, combining both the frequentist and subjectivist views.

In order to answer the second and third research question (whether there is a modelling method that can meet these criteria and whether it can be applied for reliability prediction and management), the suitability of traditional and recent reliability prediction methods in the light of these criteria will be discussed in the next sections.

2.3. Traditional Methods for Reliability Prediction

Numerous methods are discussed in literature for addressing reliability. When studying the literature, it becomes clear that authors disagree on the way in which different methods should be typified. A good example of this is found by comparing Blanchard & Fabrycky (2006) and Blischke & Murthy (2000). Blanchard & Fabrycky introduce reliability prediction as a quantitative tool that can be used to support the process of system design. Additionally, they discuss three other tools, i.e.: failure mode, effects, and criticality analysis (FMECA), fault tree analysis (FTA), and stress-strength analysis, labelling them as reliability analysis methods. Blischke & Murthy (2000) however, label the last three methods as (qualitative) prediction methods. In this thesis, the categorization of Blischke & Murthy is used, labelling both qualitative and quantitative techniques as reliability prediction methods.

Literature discusses many qualitative as well as quantitative methods for reliability prediction. However, in this thesis, the focus does not lie in providing an extensive description of the different methods, nor does it lie in the exhaustive identification of the available methods. Therefore, this section only discusses a limited number of reliability prediction methods. Because methods like physics of failure, finite elements analysis, Markov models and reliability growth testing demonstrate their added value rather late in the PDP, when a product or its prototype has already been constructed, and the focus in this thesis is on reliability prediction in the early stages of the PDP, these will not be included in the discussion.

The descriptions of the methods include a short description of the goal of the method, of the way in which it is applied as well as a discussion on the way in which the methods deals with the issues regarding reliability prediction. The techniques that are included in the discussion are a number of widely-used techniques. These are: failure mode and effect analysis (FMEA), fault tree analysis (FTA) and database methods, based on parts stress analysis (e.g. MIL-HDBK-217F, PRISM©).

2.3.1. FMEA

FMEA (sometimes extended to FMECA – failure mode, effects and criticality analysis) is a tool that was already used by NASA in the 60's (Gilchrist, 1993), and has become a well known tool in the design and manufacturing of (new) products. Its purpose is the prevention of failures from happening (McDermott, Mikulak & Beauregard, 2009). In short, an FMEA "provides a systematic method of examining all the ways in which a failure can occur. For each identified potential failure, an estimate is made of its effect on the total system, design, process, or service, of its seriousness, of its occurrence (frequency) and its detection." (Stamatis, 2003) In general, two types of FMEA can be identified: design FMEA and process FMEA. Design FMEA focuses on the product design, and makes sure that the product design conforms to all demands (specifications, regulations, etc.). Process FMEA on the other hand deals with the manufacturing and assembly processes after the product design is set. As such, design FMEA and process FMEA can be assigned to the D&D stage, respectively, the M&I stage of the PDP.

In general, an FMEA consists of a number of steps. First, a team has to be formed that has to make an inventory of potential modes of failure and the faults that are the consequence of these. The team composition is essential to the success of the FMEA process and should include customers, manufacturing engineers, test engineers, quality engineers, reliability engineers, production engineers, and sales engineers (Teng & Ho, 1996). The team has to identify potential failure modes and their effects (the faults that may occur). Furthermore, they have to estimate the following three characteristics of the faults:

- the probability that the fault will occur
- the probability that the fault will be detected
- the severity of the consequences of the fault

Next to these three characteristics, the possible action to reduce failure rate or effects (Lewis, 1996) should be identified, making FMEA a useful management tool.

In order to estimate the three characteristics of the faults, a scale is used (usually) containing the numbers 1-10. These numbers are used to score each characteristic of the faults, where 10 is the highest and 1 the lowest probability, respectively severity.

The scoring of these fault characteristics results in a measure of risk (risk priority number; RPN), that can be calculated by multiplying the three scores of the characteristics. As an example, a fault that has a moderate chance of occurrence (1/200; $S_f = 6$), which has a moderate chance of not being detected ($S_d = 4$), and gives customer dissatisfaction ($S = 4$), will have a RPN of 6 x 4 x 4 = 96.

The next step in the process is the ranking of the different faults, based on the RPN, i.e.: faults with a higher RPN will be ranked higher than faults with a low RPN. Starting at the top of the list, countermeasures will then be defined for the high ranking faults, in terms of preventive actions. Furthermore, the effectiveness of these countermeasures will be checked, by revising the measure of risk. This is (again) done by the FMEA team.

Process FMEAs basically are the same as design FMEAs. However, whereas design FMEAs look into the design of the product itself, process FMEAs look into the manufacturing process of the product, and how the manufacturing or assembly process may introduce potential product failures. Although both design FMEAs and process FMEAs use occurrence, detection rankings and severity rankings, the definition of both ranking scales may be different: organizations may customize ranking scales (McDermott et al., 2009).

Although FMEA is very successful and useful, it has to be said that FMEA as a tool has most value as a management tool (rather than as a technical prediction tool), the effectiveness of which is dependent on the extent to which it is carried out.

First of all, the composition of the team that has to identify all potential failure modes and effects, determines very much the extent to which all potential failure modes and faults will be identified. Since the identification of the failure modes as well as their RPN is based on experience and imagination, it is important to get many different viewpoints in the first step of the FMEA process (hence, it is recommended to include many different participants in the FMEA process).

Secondly, the identification of the RPN for the different faults is only one part of the FMEA process. At least as important as the identification of the RPNs, is defining and executing the actions that has to be taken in order to reduce the RPN of high-ranking faults.

An important critique of the FMEA is on the meaning of the RPN. Although it is a measure of the risk of a failure mode or fault, it is the product of three rankings, and as such, it has no meaning as a number: it is only useful in comparison. Additionally, the number of products that are designed and/or produced with this RPN is not taken into account (Gilchrist, 1993).

Furthermore, the FMEA process identifies a large number of failure modes and faults, providing insight in the way in which reliability can be improved. However, it does not provide insights in the relationships between the processes of which the PDP consist, and the reliability of the product as a result of these processes. This is with the exception of the processes in the M&I phase, that are addressed through process FMEAs.

Finally, interaction effects or cause and effect chains do not become evident in the FMEA since every cause is linked to only one effect. Causes that have more than one effect or effects being causes in themselves are not identified in an FMEA.

2.3.2. FTA

FTA is to some extent similar to FMEA. FTA is a deductive technique (Rausand & Høyland, 2004), starting from unsafe and undesired states of the system (TOP-event), and then proceeding to identifying events that are possible causes for those states (first level causes). From this point, possible causes for these first level causes (second level causes) are deduced. In this way, a causal tree is deduced, which has the undesired system state as top-event. This leads to the fact that FTA does not necessarily go through all possible fault modes. The development of the tree stops when a desired level of detail is reached. In general, there are five steps that have to be carried out in order to perform an FTA (Rausand & Høyland, 2004):

1. Definition of the problem and its boundaries
2. Construction of the fault tree
3. Identification of minimal cut and/or path sets
4. Qualitative analysis of the fault tree
5. Quantitative analysis of the fault tree

The definition of the problem consists of describing the TOP-event in terms of ‘what’, ‘where’ and ‘when’. In this way, a clear definition of the TOP-event is generated. The boundaries of the problem should include a description of the physical boundaries of the system, the state of the system at the time of analysis, the level of detail that is required for the FTA to contain, and which external stresses are included and which not.

Construction of the fault tree then takes place in the way described above: the TOP-event is identified, as well as multiple level causes and the way in which the combination of these multiple level causes lead to the TOP-event.

The identification of cut sets basically is the identification of the possible combinations of events that cause the TOP-event to occur. Minimal cut sets are cut sets, where no events can be left out to reduce the set of events while not losing the status as cut set.

Quantitative FTA is related to the cut-sets. The cut-sets can be ordered, by identifying the number of events that have to take place for the TOP-event to occur, and/or by identifying the type of event(s) that has (have) to occur for the TOP-event to occur. Also, quantitative analysis of fault trees is possible (although it is not necessary). In order to perform quantitative FTA, the probabilities of the events have to be introduced. In this way, the occurrence of the TOP-event can be quantified in terms of probabilities. A set of events, whose simultaneous occurrence leads to the occurrence of a TOP-event, is called a cut set. At the same time, a set of events whose simultaneous non-occurrence leads to the non-occurrence of a TOP-event is called a path set (Rausand & Høyland, 2004).

2.3.3. Database methods

A very well-known method for reliability prediction, which is based on parts stress, is the MIL-HDBK 217F (US-MIL-HDBK 217F, 1992). This handbook has been used for the prediction of the reliability of electronic equipment since 1962 (Blischke & Murthy, 2000). A prediction of the system reliability is made by summing up the individual failure rates of the different components of the system. The component failure rates are based on a nominal component failure rate, and a correction of this failure rate for the environmental circumstances under which it is used.

The reliability prediction that is done using the MIL-HDBK 217F is based on two very important assumptions (Condra, 2001):

1. The failure rate of a system is the sum of the failure rates of its components
2. The failure rate of the components is assumed to be constant (the reliability is exponentially distributed)

Although the MIL-HDBK 217F is currently obsolete and no longer actively supported by the Department of Defence (RIAC-HDBK-217Plus, 2006), the parts stress methodology as such is continued to being used. An example of the continued use of part stress analysis is the PRISM® method (Goel & Graves 2006), which, in addition to component failures, also takes failures into account that are due to noncomponent deficiencies (e.g. system mismanagement, design deficiencies (Goel & Graves 2006)).

2.3.4. Testing traditional methods against the criteria

1. *Insights for decision support*

   FMEA gives insights in cause and effect relations, but the insights are limited, i.e.: only one cause is linked to only one effect. Interaction effects are not identified, since causes that have more than one effect are not identified as such. Furthermore, cause and effect chains do not become evident in the FMEA, since every cause is linked to only one effect. Whether or not effects are causes in themselves is not identified in an FMEA.

   FTA and database method (Walls, et al., 2006) aim to estimate reliability as a function of the logic relating the component parts. As a result, these methods only give insights to support decision making on component level.
2. *Technical/non-technical nature*

   All presented methods provide a technical approach to reliability modelling, focusing on system failure. They do not focus on the PDP and on the way in which reliability is ‘created’.
3. *Usability throughout the PDP*

   FMEAs, FTAs and database methods can only be used when there the system design is finalized. Therefore, these methods are useful only in the later stages of the PDP. FMEAs use input from experts, and FTAs may use expert judgment as input. Database methods however, are fully based on existing data, and do not use input from experts. In FMEAs and FTAs, new insights that become available throughout the PDP may replace initial expert judgments.

4. *Incorporating uncertainty*
   Uncertainty is incorporated in FMEA and in FTA in the form of probability. However, the probability is related to a single failure mode (FMEA), or to the probability of a fault occurring FTA). It does not provide any insights in the total uncertainty that is related to the system reliability.

2.4. Recent Applications of Methods for Reliability Prediction

In the recent literature, a number of papers can be found in which not only the reliability prediction method is discussed, but also explicitly the application of such models. In the following subsections, both the models and their application (the total modelling approach) will be discussed, based on the indicated literature. The applied modelling approaches that will be discussed are:

1. *REMM (Reliability Enhancement Methodology and Modelling)*
   (Marshall, Walls & Jones, 2002; Marshall, Balderstone, Davies & Lumbard, 2003; Jones, Marshall, Aulak & Newman, 2003; Walls et al., 2006; Marshall & Jones, 2006)
2. *PREDICT (Performance and Reliability Evaluation with Diverse Information Combination and Tracking)*
   (Booker, Bement, Meyer & Kerscher III, 2003; Kerscher III, Booker & Meyer, 2001; Kerscher III, Booker, Meyer & Smith, 2003)
3. *TRACS (Transport Reliability Assessment and Calculation System)*
   (Neil, Fenton, Forey & Harris, 2001; Neil, Fenton, Forey & Harris, 2003)

2.4.1. REMM

The model that is used in REMM is a learning curve model. Although this approach is not new as such (Duane, 1964), the way in which the learning curve model is applied in the context of REMM is new. Whereas traditional learning curve models use test data in order to predict field reliability of systems, REMM also uses information that is obtained through expert judgment. In this way, it can give insights in the system under development already before test data become available. The method is focused on the development of a system, and does not focus on the process that 'creates' reliability.

The data gathering process for potential failure modes is extensively described, and is based on action research, using semi-structured interviews. For these interviews, appropriate experts are selected, based on their knowledge domains. The involved experts are briefed and trained in the failure mode estimation process. After this, concerns of the experts are elicited, leading to a map of potential faults, root causes, failure modes, classification and chance of occurrence. The outcome of the process is then fed back to the experts, and consensus about the resulting concerns is reached. The outcomes of the process can be used as input for the learning curve model. When test data become available, these data can then be used as input for the learning curve model as well, using a Bayesian updating scheme. Bayesian updating combines existing information (in the form of a probability distribution function (pdf) – the prior distribution) with new information (in the form of a pdf – the likelihood) through Bayes' theorem, resulting in a new pdf, i.e.: the posterior. In this way, REMM provides insights in the growth of reliability over time. For an elaborate discussion on Bayesian statistics, see e.g. (O'Hagan & Forster, 2004; Badoux, 1991; Smith, 1991).

Decision support is also provided by REMM. This decision support is directed at the elimination of classes of concerns as they are identified by the experts. The decision support relates to the physical system, the concerns being eliminated through design modifications.

Testing REMM against the criteria

1. *Insights for decision support*
   REMM provides decision support in the sense that it identifies the class of concerns that is most critical to reliability. Decision support may then be provided in the sense that classes of concerns may be prioritized (so that they may be eliminated). REMM does not give insights in the way in which the PDP 'creates' reliability, but focuses on the identification of concerns that can be eliminated through redesign..
2. *Technical/non-technical nature*
   REMM takes a technical approach to reliability modelling; focusing on reliability growth modelling based on (potential) system failure, and proposes physical modifications as failure elimination mechanism.
3. *Usability throughout the PDP*
   REMM uses expert judgment to create a first estimation of system reliability. Since elicitation is based on failure modes within the design, some system characteristics have to be known. Therefore, REMM cannot be used before a system concept has been defined. REMM is able, at later stages in the PDP, to update the reliability prediction, using new insights (e.g. test data that become available)
4. *Incorporating uncertainty*
   Uncertainty is incorporated in REMM in the form of probability. Furthermore, an estimation of the likelihood of the elicited concerns is asked, to obtain an uncertainty indication for the occurrence of these concerns.

2.4.2. PREDICT

PREDICT is a reliability prediction tool that explicitly combines influences on reliability from technical system design and from the manufacturing process. This is done through a reliability success tree, which consists of a system design part (the contribution of the technical system design to system reliability) and a process part (the contribution of the manufacturing process to system reliability). In order to be able to model system reliability in this way, a full understanding of the system is needed: "The system – all its parts, pieces, components, processes, activities, failure mechanisms, workings, environments, conditions, etc. – must be diagrammed or structured according to all these aspects affecting performance and in ways familiar to the community." (Booker et al., 2003).

The reliability contribution of the system design part is modelled as the sum of the reliability contribution of its components. Component (and system design) reliability is here defined by a two-parameter Weibull distribution. This distribution is constructed based on an estimation of the number of failures per unit time (failure rate) of a component.

The reliability contribution of the process part is also defined by a two-parameter Weibull distribution. This distribution is constructed based on the expected failure rate of the system that is a result of deficiencies in the manufacturing process alone.
The reliability estimation of the system in this way is based on a contribution of the product to system reliability and a contribution of the process to system reliability. In order to use PREDICT as modelling tool, it has to be possible to identify the contribution to the system failure rate separately for the system design part and for the process part.

New (additional) information that becomes available throughout time can be combined with existing information using Bayesian statistics.

Testing PREDICT against the criteria

1. *Insights for decision support*
   The PREDICT tool provides insights in the sense that the most important determinant of reliability (the component or process with the highest failure rate) may be identified. Furthermore, what-if scenarios may be performed, so that the effects of alternative components or process characteristics on reliability can be determined.
2. *Technical/non-technical nature*
   The PREDICT tool provides a combination of technical and non-technical input, using reliability as a result of the product design, as well as reliability as a result of the manufacturing process. The way in which both inputs are combined, is of a technical nature, defining the contribution to reliability in terms of failure rates.
3. *Usability throughout the PDP*
   The PREDICT tool uses expert judgment and historical information to create a first estimation of system reliability. Since elicitation is based on failure modes within the product design and within the process, a full understanding of the system and the manufacturing process is needed early in the PLC. In the beginning, reliability prediction would be based on expert judgment. Like REMM, PREDICT is able, at later stages in the PDP, to update this reliability prediction, using new insights.
4. *Incorporating uncertainty*
   Uncertainty is incorporated in the PREDICT tool, as it is incorporated in REMM, through pdfs. The estimation of the failure rates (contribution of product design and processes) is in the form of a pdf, i.e.: a failure rate is estimated, the uncertainty being represented by estimations of the failure rates related to best and worst case scenarios.

2.4.3. TRACS

The TRACS tool is a tool for reliability prediction in the context of military vehicles that uses BBNs as reliability prediction technique. The TRACS tool uses both hard evidence in the form of failure data, and soft evidence in the form of design and process capability evidence. In this way, the model provides a method for system reliability prediction already early in the PLC. The TRACS tool proposes a combination of the hard and soft evidence through the use of Bayesian probability, and more specifically, BBNs. As also identified in the previous methods, Bayesian probability is able to combine statistical (hard) evidence and subjective evidence. An important difference between the methods identified in the previous subsections is the fact that BBNs represent uncertain variables and their mutual causal and influential relationships also graphically (not only through the use of a statistical model). This provides additional insights in the way in which model variables influence reliability.

The description of the information gathering processes and of the discussions on which variables to include in the model is limited. The development of TRACS is described as an iterative group negotiation process between model developers and engineers and staff of the OEM for which the model was developed. In this way, group consensus was reached on the model. This contrasts to the extensive discussion on the data gathering process for REMM.

Although the TRACS tool is able to improve the accuracy of reliability prediction, this is not the only benefit that is provided by the TRACS tool. Rather, the insights in potential areas of improvement are identified as the main benefit of the tool.

Testing TRACS against the criteria

1. *Insights for decision support*
   As clearly identified in the description of TRACS, the tool provides insights in the way in which reliability is created, including aspects of products as well as processes. In this way, it provides insights that can support decision making at an operational and at a tactical level.
2. *Technical/non-technical nature*
   The TRACS tool provides an approach towards reliability prediction that combines both technical and non-technical data. The tool combines information and insights in the system as well as in the processes that create the system. In this way, it provides a technical reliability prediction, in the form of an estimation of the system failure rate (in the form of a pdf). This implies that the influence of the model variables on reliability has to be formulated technically as an influence on the technical reliability parameters of a system (like in the PREDICT tool).
3. *Incorporating uncertainty*
   As in the REMM and the PREDICT tool, uncertainty is incorporated in the TRACS tool in the form of probability, i.e.: the resulting reliability estimate has the form of a pdf.
4. *Usability throughout the PDP*
   The TRACS tool consists of a number of sub-models that provide insights in the separate stages of the PLC. In this way, information can be used as input for the model as soon as it becomes available. Furthermore, the model can be used throughout the PDP, and data collection focuses on data collection for the different variables included in the model.

2.5. Proposed Modelling Approach for Reliability Prediction

In section 2.1, the two objectives of this research have been introduced, i.e.:

Research objectives

1. To find or develop a reliability prediction and management method which is able to address the criteria that result from the increasing attention that is paid to LCC, as well as the increase in complexity, customer demands, TTM pressure, and globalization.
2. To provide a systematic approach for applying the reliability prediction method developed in industry

In order to address the first of the research objectives, the criteria corresponding to reliability prediction in the capital goods industry early in the PDP have been identified, and a number of existing methods (both traditional methods and recently developed methods) have been discussed.

As a result of the discussion on traditional methods in section 2.3, it can be concluded that traditional methods are not able to meet the criteria as they are identified in the research objective.
From the discussion on the application of recently developed and applied methods for reliability prediction in the previous section, it can be inferred that recent reliability prediction methods are able to meet the criteria, to a smaller or larger extent.

Recent reliability prediction approaches use Bayesian statistics as basis to ensure usability throughout the PDP, since Bayesian statistics enable the methods to update the reliability prediction based on new information that becomes available later in the PLC. Furthermore, the use of probability in all reliability prediction methods enables the incorporation of uncertainty. Inclusion of both technical and non-technical factors is not covered by REMM (the method looks at failure modes specifically related to the technical product design), whereas both PREDICT and TRACS also incorporate non-technical (e.g. process related) variables. Regarding the usability of the method for decision support, it can be said that the REMM tool and the PREDICT tool provide such support, but that they are limited to providing suggestions for technical improvement, e.g. failure mode elimination (REMM) or failure rate reduction (PREDICT). Although they provide reliability improvement opportunities, these are present at an operational level, whereas decision support is asked for at a tactical management level. For this, as mentioned earlier, insights in the processes that ‘create’ reliability are needed, i.e. the influence of the PDP on reliability has to be made explicit. TRACS, using BBNs as modelling base, provides such an opportunity.

As a result of the discussion on recent reliability prediction approaches, BBNs are selected as a potential modelling method, in this way addressing the first research objective. The discussion of BBNs in the context of reliability prediction has so far been limited and in the literature, an extensive discussion on the implementation process of BBNs is missing. Therefore further discussion on BBN implementation and application is needed in order to address the second research objective.

The most important problem with the current literature on BBNs is the fact that only a limited amount of literature can be found on the practical construction of these networks. This is in spite of the fact that BBNs are becoming more and more popular, as can be witnessed by the growing number of papers regarding BBNs (amongst others: Nadkarni & Shenoy, 2004; Marquez, Neil & Fenton (2007); Fenton, Neil, Marsh, Hearty, Marquez, Krause & Mishra 2007) and, more specifically, regarding the use of BBNs in the field of reliability (Langseth & Portinale, 2007). In this literature, results of the application of BBNs in practice are shown, and the opportunities seem very promising.

A general description of the steps that have to be taken in the process of building and using a BBN is given by Sigurdsson, Walls & Quigley (2001) (see Figure 6). However, no literature has been found that describes a concrete, methodical approach towards constructing and using BBNs for reliability prediction, in the context in which this research takes place. Papers that describe the way in which BBNs (like e.g. TRACS) are constructed in practice (Neil et al., 2001), often only briefly discuss the consecutive steps needed to construct a BBN.

In order to provide a more elaborate discussion on BBN construction, more general discussions are presented on the identification of model variables and mutual relations (section 4.1) and on the specification of probabilities (section 6.1).

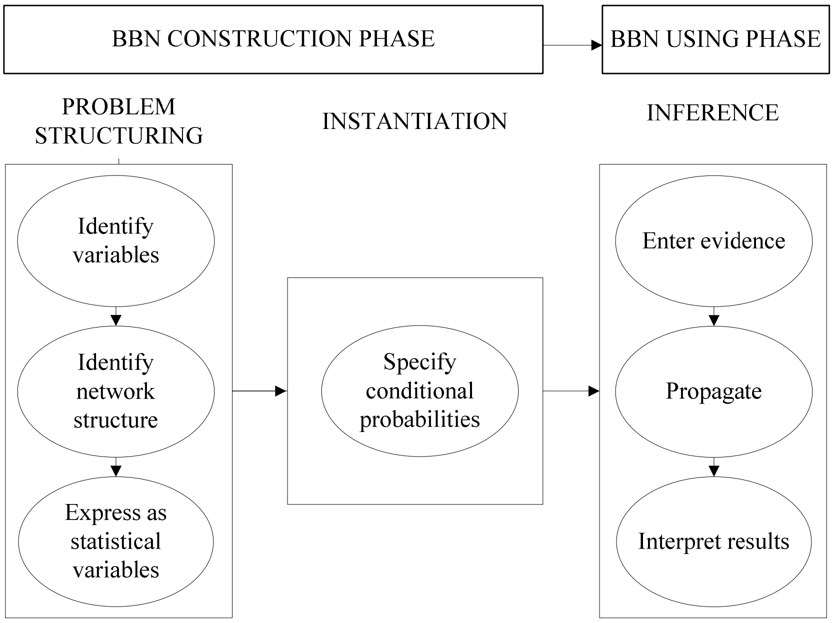


Figure 6: Process of building and using a BBN (adapted from Sigurdsson et al. (2001))

2.6. Research Questions and Research Strategy

Addressing the second research objective will be done through relating the research questions identified in section 2.1 to BBN construction in practice.

2.6.1. Relating research questions to BBN construction

In literature, the scope and boundaries of the BBN are often given or predefined. However, in practice, the scope and boundaries have to be identified. This is in line with the application of SSM. The boundaries of the BBN, which specify the scope of the BBN regarding the defined problem, have to be identified and defined. This leads to the first relevant research question.

Research questions 2.1:
How can the system boundaries of a structural model for reliability prediction be identified and defined systematically, when no information regarding the boundaries is available beforehand?

The variables that are to be included in the BBN represent the causes and the effects that are modelled. Research question 2.2 is related to the identification (related to the identification of the causes and effects that will be modelled (step 1 in Figure 6)) and definition (related to expressing them as statistical variables (step 3 in Figure 6)) of the variables in the BBN.

Research questions 2.2:
How can the variables that are to be included in the structural model for reliability prediction be identified and defined in a systematic way?

A BBN structure represents the relationships between the different variables in the model. Research question 2.3 is related to identifying and defining the network structure. In order to identify the corresponding BBN structure, the dependencies between variables have to be identified (step 2 in Figure 6). The definition of the BBN structure relates to the definition and probabilistic quantification of the relationships (step 4 in Figure 6).

Research questions 2.3:
How can the structure of the model for reliability prediction be identified and defined in a systematic way?

In order to answer these research questions, a research strategy has to be defined. This is done using Creswell (2003) and Yin (2003)

2.6.2. Research strategies

Creswell (2003) identifies two research paradigms, i.e.: the qualitative paradigm and the quantitative paradigm. Both are discussed in detail in (Creswell, 1994; Creswell, 2003). Underneath, a number of relevant issues related to each approach are summarized.

1. *The quantitative approach*
   The quantitative approach assumes that the variables that have to be included in the model are known, as well as the model itself. Furthermore, the quantitative approach focuses on numeric data rather than text or image data. Finally, quantitative research is mainly used in order to verify theories or explanation rather than build these.

2. *The qualitative approach*
   In the qualitative paradigm, the assumption is made that variables are largely unknown. Moreover, qualitative research makes use of textual data, focusing on a single concept or phenomenon. Finally, since theory is not available or incomplete, qualitative study cannot be guided by theory (in contrast to the quantitative approach).

In order to construct a BBN, factors have to be identified that influence reliability, which can be included as nodes in the model. As already indicated, the problem context is complex, and the information that is available is limited, which means that the model cannot be based on readily available information. Since the research focuses on theory building (model building) rather than hypothesis testing (model testing), the research approach that is adopted, is the qualitative approach.

According to Creswell (2003), there are five important strategies associated with the qualitative approach that can be used. These are (note: the references, except Yin (2003), are taken from Creswell, 2003):

1. *Ethnography (Creswell, 1998; Lecompte & Schensul, 1999)*
   Purpose: learn about culture-sharing behaviour of individuals or groups
   The researcher uses mainly observational data to study a group in its natural setting, over a longer period of time. The research process is flexible, and it evolves according to the encountered field settings, which define the context.
2. *Grounded Theory (GT; Strauss & Corbin, 1990; Strauss & Corbin, 1998)*
   Purpose: exploring processes, activities and events
   The researcher uses the views of participants in a study, to generalize a theory of a process, action or interaction. GT involves multiple stages of data collection and data analysis. GT has two important characteristics:
   - Constant comparison of data with findings
   - Theoretical sampling of different groups
3. *Case studies (Stake, 1995, Yin, 2003)*
   Purpose: exploring processes, activities and events
   The researcher explores a single topic in depth (this may be e.g. a program, an event, an activity), gathering detailed data through a variety of methods. Cases are bound by time and activity.
4. *Phenomenological research (Moustakas, 1994; Nieswiadomy, 1993)*
   Purpose: studying individuals
   The researcher tries to find the "essence" of human experiences concerning a phenomenon. The study includes a small number of participants, who are studied extensively, in order to develop patterns and relationships of meaning. Experiences of participants are set against experiences of the researcher, in order to provide more insights.

5. *Narrative research (Clandinin & Connelly, 2000)*
   Purpose: studying individuals
   The researcher tells a story of the life of one or more individuals in a narrative, in a chronological way. In the end, the narrative is combined with views of the researcher's life, resulting in a collaborative narrative.

Since the purpose of this research is to build a model with which reliability can be predicted and managed throughout the PDP, the focus lies on studying the processes and activities that are related to the PDP. Hence, GT and case studies seem to be suitable strategies for the research. The suitability of case studies for the research objectives lies in the fact that case studies enable in-depth study of a process (in this case, the PDP). The added value of GT lies in the fact that it provides a systematic basis for research and model building through the multiple stages of data collection and analysis. GT furthermore provides a process that is explicitly aimed at identifying and constructing theories, as well as obtaining a certain level of completeness. GT is discussed in more detail in chapters 4 and 5.

The research design, related to the second research objective, looks on the whole as follows:

1. The development of a structured, methodical approach towards the construction of BBNs, making use of a qualitative research approach, including:
   - The identification of variables, their relations, and model boundaries
   - The definition (quantification) of the identified variables and relations
2. The application of the developed approach

The context for the case study in which the methodical approach is applied is discussed in more detail in the next subsection.

2.6.3. Case study: context, environment and description

According to Yin (2003), basically, four types of case study designs can be identified:

- The holistic (single unit of analysis), single-case design
- The embedded (multiple units of analysis), single-case design
- The holistic (single unit of analysis), multiple-case design
- The embedded (multiple units of analysis), multiple-case design

The central issue of the research questions is the way in which BBNs can be constructed and applied in practice. Identifying and solving difficulties with the construction and application of BBNs in practice is therefore the central subject of the case study research. As a result, both single-case and multiple-case design is possible. The choice has been made not to perform broad research on the general application of BBNs in multiple cases, but to spend available time on in-depth research on the construction and application of a BBN in a single case study.

The case study will be performed in the context of reliability prediction. The environment, in which the case study will take place, will be the capital goods industry. More specifically, the study will be performed in a company that develops and produces medical scanning equipment, in the business unit (BU) that focuses on the cardio-vascular division. In this BU, a BBN will be constructed and applied.

It is important to note that the research focuses on one case study. Transferability of the model beyond the specified BU of the medical scanning equipment manufacturer is not a goal of this research.

In order to warrant the quality of the research and research output, both aspects of the research, i.e.: model construction and model application, will be validated. For the model construction, validation will take place according to the qualitative paradigm (Creswell, 1994; Creswell, 2003)). This includes internal validation, external validation, and reliability. All three aspects will be further discussed in the next chapter.
For validation of application of the model, the model behaviour is validated in a focus group meeting. Since the validation focuses on model behaviour, other aspects play a role than in the case of model identification. The validation aspects that are taken into account relate to validation of decision aiding models. This is because the BBN that is to be developed should provide support for decision making. Model validation in the context of decision support systems (DSSs) mainly concerns with two aspects of validity (Finlay, 1995; Finlay, Forsey & Wilson, 1988), i.e.: analytical validity and synoptic validity. These aspects of validation will also be discussed in the next chapter.

2.7. Summary and Conclusions

The way in which reliability prediction will be approached, through SSM, is presented in section 2.1. Following the discussion of this meta-methodology, the research objectives have been defined, i.e.:

Research objectives:

1. To find or develop a reliability prediction and management method which is able to address the criteria that result from the increasing attention that is paid to LCC, as well as the increase in complexity, customer demands, TTM pressure, and globalization.

2. To provide a systematic approach for applying the reliability prediction method developed in industry

In order to answer the first research objective, three research questions have been identified and answered in this chapter, i.e.:

1.1. Which criteria can be identified for reliability prediction and management methods that relate to the increasing attention for LCC, increasing complexity, customer demands, TTM pressure and globalization?

1.2. Is there a modelling method that can meet these criteria?

1.3. Can this method be applied for reliability prediction and management?

The criteria related to reliability prediction and management in the environment of the capital goods industry are identified in section 2.2. A discussion of different methods for reliability prediction and management (widely used, traditional methods as well as recently developed methods, see section 2.3, respectively 2.4)) leads to the proposal of a modelling approach for reliability prediction and management, making use of BBNs.

The application of BBNs has not been described in literature in a detailed, methodical way, leading to a number of research questions related to the second research objective:

2.1. How can the system boundaries of a structural model for reliability prediction be identified and defined systematically, when no information regarding the boundaries is available beforehand?
2.2. How can the variables that will be included in the structural model for reliability prediction be identified and defined in a systematic way?
2.3. How can the structure of the structural model for reliability prediction be identified and defined in a systematic way?

In the next chapter, the literature study will be elaborated on, discussing the topics of BBN construction in general, as well as presenting background literature for the development of a structured, methodical approach for BBN construction. This approach consists of an approach for identification of the BBN problem structure, which is described, applied, and reflected upon, in chapters 4 and 5. Validation of the model is presented in section 5.4. Chapter 6 discusses the application of BBN quantification, and reflects on it.
Validation of the instantiation is done in section 7.6 by validating the model behaviour. This results in the presentation of the collective approach, including both problem structuring and instantiation in section 6.7. Validation of the outcomes of the BBN construction process will be done by validating the behaviour of the model in section 7.6. The structure of the research and the way in which it is included in this thesis is presented in Figure 7.

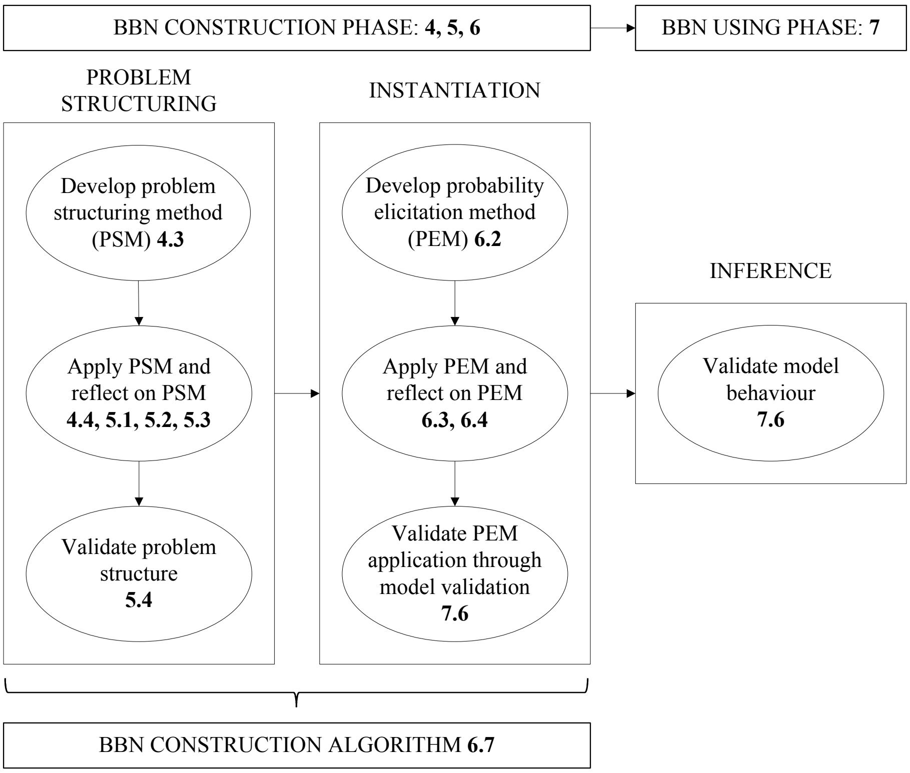


Figure 7: overview of the thesis contents, together with the chapter and section numbers (in bold)

3. Application of BBNs

As already identified in the previous chapter, the application of BBNs has not been treated in literature in a complete, methodical way. Most examples that include BBNs start with a model that is given, or discuss the BBN construction process only superficially. Also, model boundaries are often pre-defined.
In this chapter, the application of BBNs will be discussed in more detail, identifying the opportunities and challenges that lie within the application of BBNs (see section 3.3), and BBN model building in particular (see sections 3.1, 3.2). The chapter will conclude with a summary of the challenges and opportunities that lie in the building and use of BBNs that will be addressed in more detail in chapters 4– 7.

3.1. BBN Application: The Art and Science of Problem Structuring

The application of BBNs in practice, as stated by Sigurdsson et al. (2001), is a difficult process. It mainly consists of three different stages, i.e.: ‘problem structuring’ (qualitative stage), ‘instantiation’ (quantitative stage), and ‘inference’. These three stages are represented in Figure 8 by the three connected boxes (adapted from (Sigurdsson et al., 2001)). Sigurdsson et al. (2001) indicate that difficulties in the process of constructing a BBN are particularly related to BBN construction, i.e. problem structuring and instantiation.

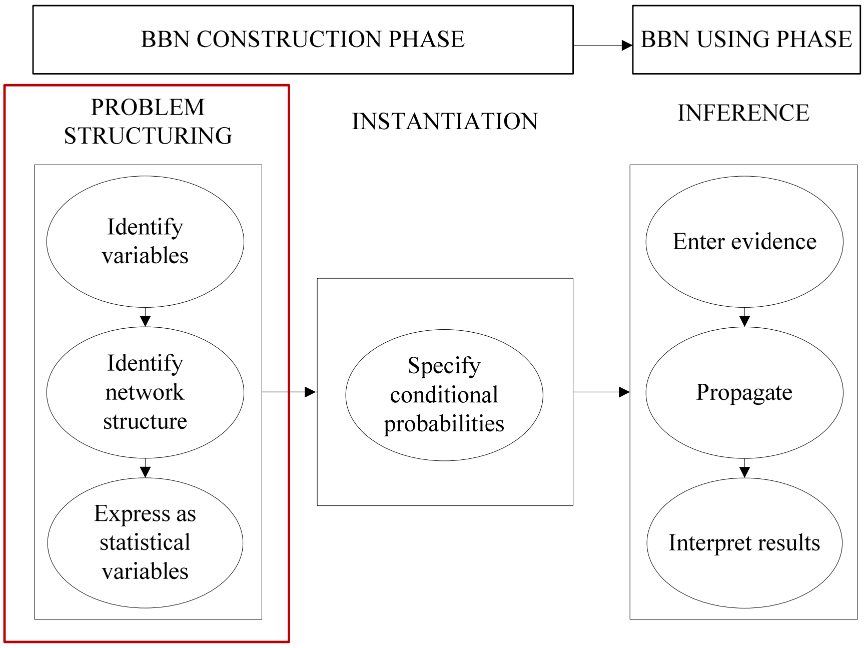


Figure 8: Process of building and using a BBN (adapted from Sigurdsson, et al. (2001))

The main problem with BBN construction is that a clearly defined process is missing. As Sigurdsson et al. (2001) state: "it remains that elicitation of judgement from experts to both structure and to instantiate the network is challenging."

Although Kjaerulff & Madsen (2008) identify several points of attention when identifying the problem structure, and many separate, general techniques can be identified that can be useful in the process of BBN construction (knowledge acquisition tools, see e.g. Carrico, Girard & Jones (1989); Cooke (1994); McGraw & Harbison-Briggs (1989) and Studer, Benjamins & Fessel (1998)), a step by step approach to problem structuring is not available.

In literature, a distinction can be made between building a network structure through automatic learning from data – machine learning in artificial intelligence – and building a network structure based on expert knowledge (Oniśko, 2008; Kjaerulff & Madsen, 2008) and we can clearly recognize both types of BBN construction in literature. For BBN construction based on data, see e.g. (Cooper & Herskovits, 1992; Singh & Valtorta, 1995; Larrañaga, Kuijpers, Murga & Yurramendi, 1996; Cheng, Greiner, Kelly, Bell & Liu, 2002). For BBN construction based on expert knowledge, see e.g. (Hodge, Evans, Marshall, Quigley & Walls, 2001; Langseth & Portinale, 2007; Nadkarni & Shenoy, 2004; Neil et al., 2001; Neil, Fenton & Tailor, 2005; Norrington, Quigley, Russell & Van der Meer, 2008; Oniśko, 2008. Because there is no hard data available in the early stages of the PDP, and because an SSM approach is chosen, taking the perceptions and mental models of participants into account, the focus in this thesis lies on BBN construction using expert knowledge. Since experts provide the knowledge that is used to construct the BBN individually, providing a BBN for every expert would be most true to the opinion and perception of the individual expert. However, although this would be a better reflection of the individuals' viewpoint, it would not be possible to use all BBNs individually for decision support, since this requires consensus. The ways in which different authors have approached the problem structuring stage using experts' knowledge is discussed further in the next subsection.

3.1.1. Problem structuring through knowledge elicitation

Neil et al. (2001) discuss the problem structuring stage only briefly in their paper. They emphasize that iterative negotiation was used, in order to come to a problem structure that was acceptable for all parties involved. Langseth & Portinale (2007) discuss the BBN modelling process in some more detail. They confirm Neil et al. (2001), presenting the problem structuring as an iterative process that takes place involving domain experts and BBN experts. Because it is a group process, the possibility of problems occurring in the problem structure identification is present: e.g. individuals dominating group discussions, new ideas being discouraged, conforming to group's view, etc. (Winkler & Clemen, 1999; Meyer & Booker, 1991).

Nadkarni & Shenoy (2004) give an example of interviewing a single expert to identify a complete BBN structure. Although they spend a little more attention to the interview, they imply that the expert is known, and that the scope of the network is known. Furthermore, for applicability of their approach, it has to be possible to construct the full BBN by interviewing only one single expert. This limits the possibilities of their approach; since in this case it is required that one individual expert has complete understanding of the problem under study, which is certainly not the case in our problem context. Norrington et al. (2008) describe an interview process with 2 interviewees, using iterative discussion to construct a BBN. In the elicitation process discussed by Oniśko (2008), three domain experts are involved in the identification of a model for the diagnosis of liver disorders. The limited number of experts involved is due to the specialized area of application of the network. Finally, Hodge et al. (2001) describe a process in which experts, related to the field of expertise that is studied, are first interviewed individually. After that, the results are evaluated in group sessions, aggregating the knowledge of different experts on concerns relating to the subject under study.

Neil, Fenton & Nielson (2000) describe a structured approach to BBN construction. However, this is in fact a general approach to model building, and does not provide guidelines on how to identify the relevant variables and their relationships that are to be included in such a network.

When reviewing the literature, we see that problem structuring in BBNs is recognized, but rarely described extensively or treated systematically. Furthermore, we find that the problem/system boundary is treated in papers as if it is known beforehand. Hence, we identify two main challenges with respect to the problem structuring stage of BBN building:

1. How can we identify, in a systematic way, the variables to include (nodes in the BBN) and the structure of the BBN?
2. How can we construct a BBN, when the problem boundaries are not pre-defined (i.e. when the domain experts are not known beforehand)?

In order to address these challenges, we want to be able to systematically identify the random variables to be included in the model as well as the structure of the BBN. Also, we want to establish a structured method which enables us to infer the boundaries of the problem, thereby escaping the limitations of current model building practice, in which the problem context and boundaries are known beforehand or predefined. For this purpose, problem structuring will be elaborated on in the next chapter. As already indicated in subsection 2.6.2, GT can provide a systematic basis for research and model building through the multiple stages of data collection and analysis (Houben, Sonnemans & Newby, 2010). Therefore, GT procedures and techniques will also be discussed in the next chapter, together with other methods for problem structuring.

3.1.2. Validation

Important aspects that contribute to the quality of qualitative research are validity (both internal and external) and reliability (Creswell, 2003). Since the first part of the research focuses on the construction of a BBN, internal and external validation focus on the constructed BBN, whereas the reliability focuses more on the BBN construction process. Creswell (2003), referring to Merriam (1998) addresses internal validity, external validity, and reliability, as representations of the trustworthiness of qualitative research. All representations will be presented underneath, so that they can be addressed in the next chapter.

Internal validity (Truth value/credibility):

Internal validity concerns with the accuracy of information and the extent to which it matches reality. Creswell (2003) suggest the following strategies that have a positive effect on internal validity (in order of decreasing ease of implementation):

- Triangulation of data: data collection may take place through multiple sources and methods.
- Member checks: findings can be taken back to the interviewees, so that the accuracy of findings can be checked against the real experiences of the interviewees.
- Rich, thick description: a rich, thick, and detailed description of the performed research may be given , so that readers may experience a shared experience in the discussion of the research.
- Researcher's bias: Clarification of the role of the researcher at the outset of the research.
- Presentation of negative, discrepant information: not all information will be in account with the findings, nor does all information have to be in account.
- Spending a prolonged time in the field: observations/data collection can take place (repeatedly) over a longer period of time.
- Peer debriefing: Inclusion of peer examiners that review finding of the research and ask questions about it.
- Using an external auditor: involving an external auditor, who is new to the project and who can comment on the research process.

External validity (Transferability)

External validity is concerned with the limited transferability of the research. External validity can be increased using the following strategies:

- Rich, thick description: a rich, thick, and detailed description of the performed research.
- Typicality of cases: a description of the extent to which the case is typical compared to other cases in the same category.
- Multi-site design: investigating several sites, situations and cases.

Reliability (dependability)

Reliability is related to the limited repeatability of the research. Three different techniques can be used to ensure reliability:

- Researcher's position: providing a detailed description of the research focus, the researcher's role, the interviewee's position, the basis for their selection and the context from which the data will be gathered.
- Triangulation: using multiple methods for data collection and analysis.
- Audit trial: Reporting data collection and analysis strategies in detail in order to provide a clear and accurate picture of the methods used in the study, all phases of the research being under scrutiny of an experienced researcher from the field of social sciences.

3.1.3. Summarizing the challenges in the problem structuring stage

Summarizing, two challenges can be identified in the problem structuring stage:

1. Identifying the variables to be included in the BBN and the network structure
2. Identifying the problem boundaries of the BBN

These challenges are directly related to research questions 2.1, 2.2, and 2.3, introduced in section 2.6. The first challenge relates to the identification and definition of the variables that are to be included in the BBN (research question 2.2), and part of research question 2.3: the identification of the BBN structure. The second challenge directly relates to the research question 2.1, which addresses the boundaries of a BBN.

The way in which both challenges are addressed is discussed in chapters 4 and 5, together with validation of the problem structure.

3.2. BBN Application: Instantiation

Where Sigurdsson et al. (2001) already identified that elicitation of expert knowledge to build BBNs is challenging, Renooij (2001) states that thee instantiation stage (second stage of the BBN construction process, see Figure 9) is even more difficult. She states: “the construction of the qualitative part with the help of domain experts is feasible; the elicitation of the large number of probabilities required, however, is a far harder task.”

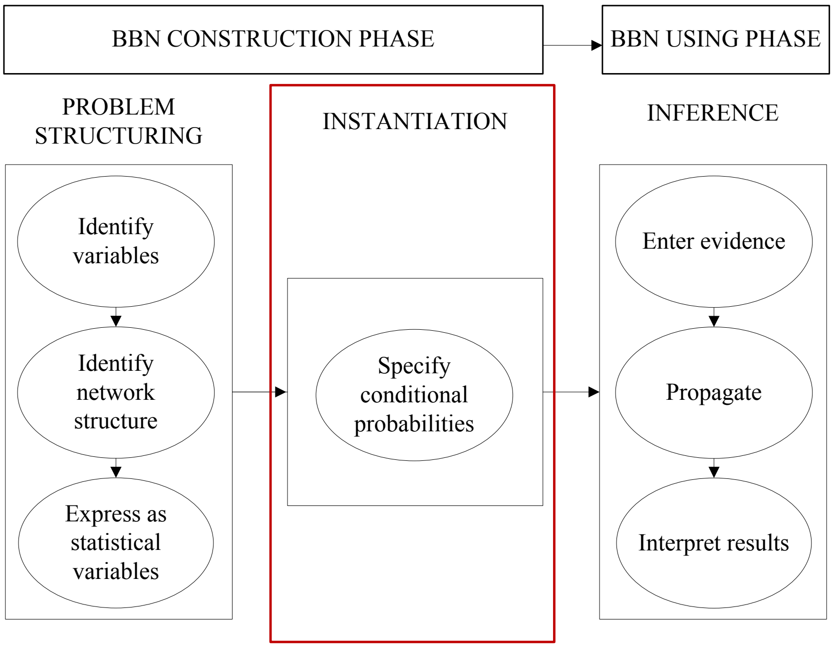


Figure 9: Process of building and using a BBN (adapted from Sigurdsson et al. (2001))

In this section, different ways for acquiring the probabilities for the CPTs are identified and described. Also, an argumentation is presented for the choice of one of these methods. Finally, literature is discussed in order to identify the important points of attention with respect to the application of this method. The way in which the application of the method takes place will be discussed in chapter 6.

3.2.1. Acquisition of the probabilities

Kjaerulff & Madsen (2008) identify three possible ways to identify the probabilities to use in the BBN in order to instantiate the network: the CPTs (as introduced in the example model discussed in section A) can be elicited (Kjearulff & Madsen, 2008):

- By establishing the probabilities using a mathematical model
- By eliciting the probabilities from experts
- By retrieving the probabilities from databases

If a clear mathematical relationship can be identified between two different variables, it becomes easy to specify the relationship. Kjaerulff & Madsen (2008) identify a number of possible mathematical relationships.

Generally, mathematical expressions can make use of probability distributions (e.g. the normal distribution), logical relationships (e.g. 'if-then-else' or 'and'), standard mathematical functions (e.g. exponential), and relations (e.g. equals). An important advantage of this method is that the reusability and maintenance of mathematical models is much easier (Kjaerulff & Madsen, 2008). Two important disadvantages are that there has to be a mathematical relationship between two variables (which is not always the case), and that this relationship has to be known.

A second way of expressing the relations in a BBN is through the elicitation of the conditional probability tables (CPTs; these express the relations in the BBN quantitatively) from experts. This is a topic that is much discussed in literature. An important advantage of eliciting the probabilities from experts is the availability of the information. Especially when the BBN structure has been elicited from the experts (which is the case in the context of a systems approach), direct relevant knowledge is readily available from experts. An important disadvantage of probability elicitation is the fact that humans have difficulties reasoning with probabilities (Cooke, 1991) and are generally not good at it. Also, humans feel awkward in expressing their domain knowledge in terms of conditional probabilities (Kjaerulff & Madsen, 2008). Therefore, although it provides an important advantage because of the availability of the information, as an information source for probabilities, expert knowledge should be treated very carefully and much effort has to be put into the elicitation process in order to get appropriate results.

A third way of expressing the relations in a BBN is through the analysis of existing knowledge (Bedford et al., 2006). In the case that relationships are quantified on the basis of data, normally, the problem structure is also based on data. In order to do this, a database has to be available of different cases, on which the BBN structure can be based. If enough data are available, the different probabilities that together form the CPT can be calculated from the data. As an example (see appendix A), the model from Figure 24 (page 163) is taken. Looking at the node 'Jack on time at work? (JOT)' in the corresponding CPT (Table 24, page 168), the probability can be found that Jack is not on time at work, given that there is no train delay. This probability can be deduced (if the data are available), by taking the total number of times that there was no train delay, and the total number of times that Jack was not on time, in these cases (so only taking into account the cases in which there was no train delay), and divide the latter by the former number. This would result in the probability of Jack not being on time at his work, while there is no train delay being 0.20 (e.g. because, of 100 times that there was no train delay, Jack came late 20 times). Although this is a straightforward way to identify the probabilities, this method can only be used if a large amount of data is available with respect to the model elements, which is a disadvantage if this is not the case.

In the situation of reliability prediction, a systems approach is chosen. As a result, experts have become strongly involved and the model building process is heavily dependent on the mental models of the involved experts. The only source that is available for identifying the probabilities that are related to the mental models, are the experts themselves. Hence, probability elicitation from experts is essential in the model building process.

Although much literature can be found on the topic of probability elicitation, only a small amount of literature is focused on probability elicitation in the context of BBNs, see e.g. Renooij (2001). An important challenge that can be identified in this context, is that experts are often reluctant to express probabilities numerically (Renooij & Witteman, 1999), whereas often a large number of conditional probabilities has to be elicited (Van der Gaag, Renooij, Witteman, Aleman & Taal, 2002). It is therefore important to control (and possibly limit) the number of probabilities that has to be elicited.

The elicitation process that is developed for the elicitation of the CPTs has to take all problems and difficulties that are related to probability elicitation with experts into account. The development of a suitable method for acquiring the probabilities that constitute the model, and that is able to deal with the challenges mentioned in this section is discussed in chapter 6, together with its application.

3.2.2. Validation

Validation of the probabilities that are entered in the BBN is approached through validation of the model behaviour, which diverges from validation of the problem structure. As stated by Renooij (2001): “An indication of the validity of the [probability] assessments can also be obtained by entering observations into the belief network and computing the effect of the observations on the probabilities for certain variables of interest. The outcomes for these variables can then be checked against available data or presented to the expert.” Therefore, the validation of the probability assessments is incorporated in the model validation process that is described in subsection 3.3.3.

3.2.3. Summarizing the challenges in the instantiation stage

One main challenge can be identified in the instantiation stage, i.e.: acquiring the probabilities for the BBN. With respect to this challenge, two important issues that have to be taken into account, i.e.:

1. Experts are the only source of information
2. The number of probabilities to be identified is large

The challenge of acquiring the probabilities for the BBN is directly related to a part of research question 3 (see section 2.6): the definition of the BBN structure. Because the BBN structure is defined by the node probability tables (NPTs), the acquired probabilities define the network structure.

Next to addressing the topic of probability acquisition, the topic of validation needs to be addressed as well (as is the case in the problem structuring stage). This is done in section 7.6.

3.3. BBN in Practice: Inference

Inference in BBNs represents the stage in which the network(s) are used (see Figure 10).

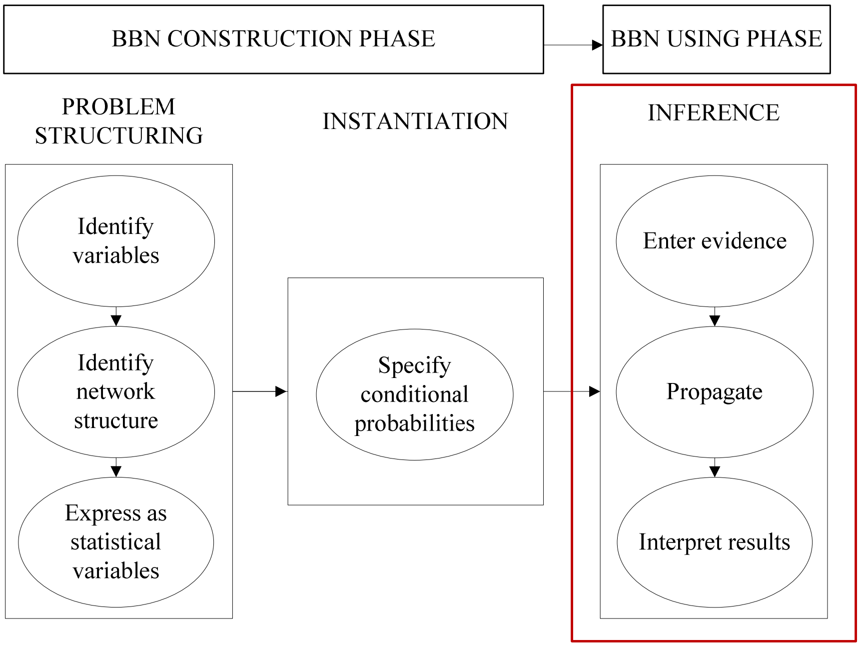


Figure 10: Process of building and using a BBN (adapted from Sigurdsson et al. (2001))

Bayesian inference in BBNs is based on Bayes' theorem, the use of which is shown in appendix A. Application of BBNs on a larger scale has been strongly dependent on the availability of efficient algorithms and the availability of computing power (as indicated by Neil et al. 2001). Although relevant literature on algorithms for Bayesian inference is available (see e.g. Pearl, 1988; D'Ambrosio, 1999), this discussion is left out of the thesis. This is because this thesis focuses on the application of BBNs in practice for reliability prediction and management, and not on the effectiveness or efficiency of the calculation process of BBNs. For the execution of the inference, the choice has been made to use the general purpose BBN software tool AgenaRisk (Agena Ltd., 2009).

In practice, next to for reliability prediction (see subsection 3.3.1), Bayesian inference through BBNs can also be used for supporting decision making (Sigurdsson et al., 2001; Fenton & Neil, 2001). In this context it is important to note that BBNs often only focus on a single attribute (in the case of this thesis, reliability) (Fenton & Neil, 2001). Since more criteria play a role (such as e.g. costs and effort), decision making is only supported to a limited extent. Furthermore, BBNs (in contrast to influence diagrams (IDs; Shachter, 1988;

Howard & Matheson, 2005; Zitrou, 2006) are not able to directly compare alternative actions on the basis of their ‘value’, i.e.: they do not represent a ‘preference function’, as IDs do. Moreover, regular BBNs are time-independent, not able to take time into account. The way in which decisions can be supported by BBNs is discussed in more detail in subsection 3.3.2.

Next to for reliability prediction and decision support, Bayesian inference can also be used for validation purposes. As already discussed, BBNs are models that are based on subjective information, which makes empirical validation impossible. Therefore, validation closely involves the evaluation of the behaviour of the model with the experts that provide the information. An extended discussion on the way in which validation can take place through inference is presented in subsection 3.3.3.

3.3.1. Using Bayesian inference for reliability prediction

The use of BBNs for reliability prediction is the most straightforward use of the BBN. In order to make an estimation of the reliability, the outcome of the model (the leaf node, being reliability) has to be calculated, based on the inputs for the model in the sense of quantitative values for the input variables (the root nodes). Note that the values of the nodes in a BBN are expressed in terms of marginal probability distributions (MPDs)

As stated, BBNs may also be used for supporting decision making (Neil et al., 2003; Fenton, Marsh, Neil, Cates, Forey & Tailor, 2004). This will be discussed in the next subsection.

3.3.2. Using Bayesian inference for decision support

Scenario analysis through BBNs may be used for decision support in two generic ways:

1. The outcome (value of the leaf/child node) can be calculated as a result of different inputs to the model (values of the parent/root nodes)
2. The likelihood of the input (values of the parent/root nodes) that corresponds to a certain outcome (value of the child/leaf node) can be calculated

Both ways of decision support can be combined. A situation can be characterized in terms of the inputs (pre-defined values of input variables). In the other direction, the most likely values of inputs that correspond to particular outcomes can be calculated.

Next to using scenario analysis for decision support, also sensitivity analysis can be used to compare the effect of the different parent nodes on their child nodes, and order them with respect to the strength of the influence. In this way, the most effective way of influencing the target variable can be identified, enabling decision support. Because changes in the value of the output variable may be small with respect to changes in the input variables – the value of the model output (MPD) may change from ‘0.20 – 0.80’ to ‘0.18 – 0.82’, when the value of a model input changes from ‘0.05 – 0.95’ to ‘0.25 – 0.75’ – the probabilistic values per se should not be used to base decisions on. A disadvantage of sensitivity analysis is that it does not give any insights into the effects of this change and therefore does not provide a very good basis for decision making.

It has to be stated that the support of decision making can only be in terms of the model, i.e. the support that is provided by the specific model that is developed in this thesis focuses solely on reliability and therefore cannot take any monetary or other considerations into account.

A more elaborate discussion on the use of Bayesian inference for decision support (including both scenario analysis and sensitivity analysis) will be presented in chapter 7.

3.3.3. Validation of model behaviour

Validation of the model behaviour is different from validation of the problem structure. As indicated in the research questions, structural models have to be used in the context of reliability prediction early in the PDP, rather than models based on objective, empirical data only. These models generally do not represent physical relationships that can be tested, but typically provide insights through representing cause and effect relations. The validation of such a model is different than the validation of a physical model. In models that are built using a subjective, but rational approach (a systematic approach, based on expert knowledge), evaluation should focus on showing that the relationships that are included in the model are appropriate. This means that evaluation should aim at identifying the extent to which the relationships correspond to the viewpoint of the user. This is confirmed by Zitrou (2006), who states that validation “involves ensuring that the model is a meaningful representation of the particular situation”, hereby referring to the meaningfulness for the experts that are involved in the model building process.

In this thesis, BBN validation is discussed, incorporating the considerations presented by Finlay (1995), Finlay et al., (1988), O’Leary, Goul, Moffitt & Radwan (1990) and Borenstein (1998). Although their considerations are made in the context of DSSs, the nature of DSS corresponds to the nature of models built using a subjective, rational approach. Hence, this discussion can be used in the context of BBN model behaviour validation. Finlay (1995) and Finlay et al., (1988) present a very elementary discussion on validation. They distinguish two types of validation:

1. Analytical validation (i.e. checking out each part)
2. Synoptic validation (i.e. checking that an acceptable output is reached for each of a set of inputs, and total model performance is checked).

Analytical validation

Analytical validation focuses on the model structure and consists of four steps (see Table 2). Analytical validation ends with a summation of the assumptions that underlie the model. As can be seen, steps 1 and 2 in the analytical validation process relate to the definition of the variables that are incorporated in the model, respectively to their statistical definition, whereas steps 3 and 4 relate to the quantitative definition of the relationships. The first two steps of the analytical validation process are addressed in the problem structuring stage (in the steps ‘Identify variables’, and ‘Express as statistical variables’, see Figure 8). The third and the fourth step of the analytical validation process are relevant after the BBN has been quantified (after the instantiation phase of the BBN construction process). Analytical validation is very similar to 'face-validity' as described by Borenstein (1998) and O'Leary et al., (1990). They refer to face validation of the model as the certification that the model includes all essential concepts, and correctly represents the problem and expert knowledge.

Table 2: Four steps constituting the analytical validation process

1. Checking the definitions of the variables: The meaning of the variables that are included in the model should be clear to all stakeholders
2. Checking the consistency (both with respect to dimensions and with respect to units) Both the dimensions (e.g. length or time) and the units (e.g. kilometres, miles, hours, or seconds) of the variables should be consistent.
3. Checking on the feasibility of the relation A relation can be checked on its feasibility by the stakeholders that have provided the information for the relation. Two simple checks on relations are on the behaviour of the relation when the concerned variables have extreme values, and possible 'strange' behaviour of the relation when the concerned variables take on a certain value.
4. Checking that the range of application of the relation fits the intended range of application, also specifying the areas where the two do not match The range of the concerned variables that are involved in a relation should be specified

Synoptic validation

Next to analytical validation, also synoptic validation is identified by Finlay (1995). Synoptic validation concerns the working of the model, by looking at the acceptability of output of the model when applying sets of inputs. In the case of 'hard systems', i.e. physical, measurable systems, synoptic validation can be done by measuring the output and comparing it to the model output. When historical output of the system corresponds to the model output, there is so-called 'replicative validity'; when output of the system corresponds to earlier generated output of the model, there is so-called 'predictive validity' (Finlay, 1995). Stakeholders of the model need to be involved in model validation in the sense that they have to 'explore' the sensitivity of the system. System sensitivity can be studied in two ways:

- Reflecting on changes in the behaviour of the model based on inclusion and exclusion of variables (this is related to the problem structuring stage)
- Calibrating the system output to the output that is expected by the stakeholders

In case of soft systems, it is important to stress that the model is based on the mental models of the stakeholders. Therefore, the involvement of stakeholders in the model building process is the keystone of validation, since the stakeholders have to validate the model output.

Synoptic validation as described by Finlay (1995), can take place when the model generates outputs for a certain set of inputs (corresponding to certain situation(s); scenario analysis). If these outputs are acceptable for the stakeholders, then the model is synoptically validated. Borenstein (1998) (referring to the model assessment of Glass (1983) in the context of decision making) identifies a similar approach to validating model output; 'user assessment', i.e.: “the process by which interested parties can determine, with some level of confidence, whether or not the model’s results can be used in decision-making”.

To enable the stakeholders to validate the model output, the behaviour of the model has to be validated. This may be done by evaluating the effects of changing the value of different input variables on the model output. By performing design of experiments (DoE), only a limited number of values for input variables are needed for analyzing the effects of the input variables. Therefore, analysis of the model behaviour is performed using sensitivity analysis through DoE. The topic of DoE will be further addressed discussed in chapter 7.

In conclusion, it can be said that validation of the model in the traditional sense is only partly possible:

- Analytical validation is possible, checking the definitions of the model variables, checking their consistency (with respect to dimensions and units), checking the feasibility of the relation and checking that the range of application of the relation fits the intended range of application.
- As a result of the involvement of experts in the model construction process, synoptic validation of the model has to take place by involving the participants of the model construction stages in the validation process. Since the model is built based on the collective knowledge of a number of experts, the validation also has to take place collectively, i.e.: in the form of a group session.

A more detailed description of the validation process is given in section 7.6.

3.3.4. Summarizing the opportunities in the inference stage

Three opportunities for the use of Bayesian inference can be identified in the inference stage:

1. Using Bayesian inference for reliability prediction
2. Using Bayesian inference for decision support
3. Using Bayesian inference for BBN model behaviour evaluation

The way in which Bayesian inference plays a role with respect to these three opportunities is further elaborated on in chapter 7.

3.4. Summary and Conclusions

As a conclusion to this chapter, the identified challenges with respect to the two stages of the BBN construction and the identified opportunities related to the use stage together with possible approaches are presented in Table 3.

Together, the challenges and opportunities address the research questions formulated in section 2.6, regarding the systematic application of a technique for reliability prediction that is able to address the following criteria:

- The method has to give insights in order to support decision making
- The method has to take non-technical factors into account
- The method should be usable throughout the PDP
- The method should be able to incorporate uncertainty

All stages of the BBN application process as described in Figure 6 are applied in a single case study, which is discussed throughout chapters 5, 6 and 7. Only one case study is performed at one BU in a company. Since transferability of the model that results from the research is not a goal of this research, this does not present a problem. Still, further discussion on transferability of the research and the research results is included in this thesis in chapter 8.

Table 3: Challenges, opportunities and approaches in building and using a BBN

Problem structuring stage (Chapters 4 and 5)	
Challenges	*Approach(es)*
Identifying BBN problem structure (variables and network structure), and identifying BBN problem boundaries *Related to research questions 1, 2 and 3 (section 2.1)*	Adapted GT approach
Validation	Survey, filled in by experts (*section 5.4*)
Instantiation stage (Chapter 6)	
Challenges	*Approach(es)*
Acquiring a large number of probabilities, with only expert knowledge available as information source *Related to research question 3 (section 2.1)*	Probability elicitation
Validation	Bayesian inference (*section 7.6*)
Inference stage (Chapter 7)	
Opportunities	*Approach(es)*
Using Bayesian inference for reliability prediction *Related to the research objectives (section 2.1)*	Scenario analysis
Using Bayesian inference for decision support *Related to the research objectives (section 2.1)*	Scenario analysis Sensitivity analysis
Using Bayesian inference for synoptic model validation	Feedback from experts using scenario analysis and sensitivity analysis (*section 7.6*)

4. BBN Construction: Problem Structuring

In this chapter, the application of GT in the context of BBN problem structuring is discussed. First, in section 4.1, Mitchell's (1993) 7 model dimensions are discussed, together with a number of methods for problem structuring – amongst which GT – and the way in which they give an interpretation to the dimensions. After this, GT will be discussed in more detail, addressing the method as well as its canons and procedures. Then, in section 4.3, the way in which GT is adapted for use in BBN construction is presented. Section 4.4 discusses how the case study is approached. The chapter concludes with the summary and conclusions.

4.1. Methods for Problem Structuring

In literature, no systematic, coherent, and structured approach is found for problem structuring in the context of BBNs. Often, literature describes the problem structuring phase only superficially, and does not go into detail on the identification of variables, their mutual relations, or their state-space. Therefore, model building in general is looked at and studied, to find and develop a suitable approach towards problem structuring for BBN construction.

In this section, two topics regarding model building will be discussed. First, the 7 model dimensions that are introduced by Mitchell (1993) will be discussed. These model dimensions will also be directly related to modelling for reliability prediction and reliability management. After that, a number of procedures and techniques are presented with which variables and their mutual relations may be identified.

4.1.1. Mitchell's 7 model dimensions

The 7 model dimensions that are identified by Mitchell (1993) are discussed here, relating them to the model for reliability prediction and management early in the PDP.

1. *Actuality – abstract*

The first model dimension relates to whether the model models real life directly (like e.g. a physical scale model), or whether the model only represents the logic. In the context of reliability prediction and management, it is not interesting to replicate the way in which the PDP 'creates' reliability, i.e.: an abstract model is more useful than an actual model.

2. *a. Black box – structural and related dimensions*

The second model dimension relates to the question whether cause-effect relations are implicitly or explicitly taken into account. Mitchell relates a number of other (sub)dimensions to this dimension. In the criteria for the reliability prediction method it was already made clear that there is a strong need to make cause-effect relations explicit, in order to provide decision support.

2. *b. Predictive – exploratory*

As a reliability prediction and management model, the model should be able to predict reliability as well as allow some exploring of possible realities in order to provide a basis for decision support.

2. *c. Instrumental – realistic*

The model is realistic in the sense that the model relations are understandable, i.e.: people are able to explain and understand the relations between the variables in the model.

2. *d. Micro model – macro model*

The reliability prediction and management model looks at how the different aspects of the PDP are related to reliability. As such, it looks at the system as a whole (the PDP at a macro level), rather than at a limited part of the PDP in a high level of detail (micro level).

3. *Standard – purpose built*

The model is a purpose built model, being built for a single BU, in a specific company in a specific case study.

4. *Absolute – relative*

The model is not generally valid. The model can be considered valid, when the people that are involved in the model building process agree upon the model.

5. *Passive, normative, behavioural, interactive*

The model has a passive nature. This means that the model shows the way in which input and output are related, i.e.: how reliability changes if the factors that influence reliability change.

6. *Private – public*

Since the model is specifically built for a BU in a specific company in a case study, the model is private to the company, rather than publically available.

7. *Part – whole*

The model that is to be constructed looks at a number of factors that affect reliability throughout the PDP. It is clear that the number of factors that is incorporated in the model is limited, and that the model only reflects a part of the PDP. The model construction process should cover a broad area of the PDP however, so that the model can be built, taking a broad spectrum of information into account. The external influences on the model and the resulting uncertainty are taken into account through the use of probabilistic modelling.

4.1.2. Problem structuring procedures and techniques

In literature, a number of problem structuring procedures and techniques may be identified. The aim in this thesis is not to be exhaustive in the discussion of problem structuring techniques and procedures. Rather, the aim is to discuss these techniques and procedures, as well as the procedures and techniques that are used in the research strategy GT (subsection 2.6.2) for the same purpose. In this way, the techniques and procedures may be compared, and the suitability of GT for the purpose of problem structuring in BBN construction may be better addressed.

Process Mapping

Process mapping as defined by Anjard (1996) is identifying, documenting, analyzing and developing an improved process. Through process mapping (Soliman, 1998), "it is easier to determine where and how to improve the process". Soliman (1998) identifies three steps in which process mapping takes place:

1. Identifying products, services and their related processes. The starting and finishing points of processes are identified
2. Data gathering and preparation

3. Transforming the data into a visual representation in order to identify the bottlenecks, wasted activities and duplicated efforts

Biazzo (2002) provides 2 main criticisms on the ways in which process-mapping techniques are used:

1. Process mapping techniques in general spend only little attention to social and human aspects of the processes. Rather, process mapping techniques focus on the physical and technical aspects that play a role in transforming input into output.
2. Process mapping provides a simplified image of the processes of a firm, leaving out the complexity that is inherent to these processes.
   This does not comply with the suggested macro focus of the model dimensions (dimension 2d).

Process mapping techniques are especially useful in identifying opportunities to improve the processes that play a role in the company. However, they take a strongly technical view, ignoring social and human influences to a considerable extent. This does not comply with the suggested approach towards model dimension 7: looking at the PDP as a whole, taking a broad spectrum of information into account.

Cognitive Maps (SODA)

Cognitive maps are pictures that represent the way in which different concepts are related in terms of the descriptions of the concepts in words, and the relations in terms of connecting lines. They provide an instrument that can be used to structure the thoughts of people, and they are mainly used in the context of decision support, in SODA (Strategic Options Development and Analysis). Cognitive maps are used in a situation where there is a problem that needs to be solved, and is aimed at making choices (Mackenzie, Pidd, Rooksby, Sommerville, Warren & Westcoombe, 2006). As a result, they are often used in consultancy. The use of cognitive maps in the context of decision making in early research was, according to Eden (1988), focused on problem solving with individuals. This has changed into problem solving with teams of 6-10 people (Eden, 1988). The ways in which the teams contribute to the development of the cognitive maps can be diverse. Mackenzie et al., (2006) identify the development of cognitive maps in two ways, as part of SODA I, and as a part of SODA II. The cognitive maps are based on interviews with individuals in SODA I. In SODA II, the basis for cognitive maps lies in a group session, where all group members brainstorms, and contributes concepts to the session through a network of connected computers.

The focus of cognitive maps on action and problem solving makes it a useful technique for problem structuring (Eden, 2004). Eden & Simpson (1989) provide a detailed case study of SODA, using cognitive mapping. The case study clearly shows the strong influence that the problem owners have on the resulting map, the problem owners also being the decision makers, and their providing the input on which the model is built. As a result, the model is relative (dimension 4) to the limited extent of the view of the problem owners (also limiting the spectrum of information that is taken into account; dimension 7).

Grounded Theory (GT)

GT looks at theory building (Strauss & Corbin, 1990), based on observations and data collection in practice. GT is able to use a wide spectrum of data and information that can be collected in the field. As stated by Locke (2001), GT is empirically based. Rather than testing a pre-defined view, it provides an observation of reality. Strauss & Corbin (1990) state that GT "uses a systematic set of procedures to develop an inductively derived grounded theory about a phenomenon", again stressing the inductive character of the methodology. An important characteristic of GT is that it does not use any prior information, and that it builds theory only based on information that is obtained throughout the research, making it very suitable in the context of model building without any prior information. In the context of problem structuring procedures and techniques, the focus within GT has to be on the systematic set of procedures that is referred to by Strauss and Corbin.

The focus of the research questions that are related to GT lies in general on actions and processes, making it suitable for decision support in the PDP. An important aspect of GT as it is presented by Corbin & Strauss (1990) is that it needs an amount of procedural rigor, while at the same time allowing for some flexibility. When using GT it is therefore important to discuss the extent to which procedures and canons that are related to GT by Corbin & Strauss (1990).

4.1.3. Choice for GT as problem structuring method

Based on the discussion of the model dimensions, the problem structuring methods have been evaluated. Process mapping is not suitable for problem structuring in the context of BBN construction, since it focuses solely on the technical aspects of processes (Biazzo, 2002).
Cognitive mapping can be very useful for problem structuring in practice. Its strong roots in reality makes cognitive mapping suitable for problem structuring. However, a strong disadvantage of cognitive mapping is the fact that the decision makers and problem owners also have to provide the input for the model, and that the number of people that are involved is limited (as indicated by Eden, 1988 and Mackenzie et al., 2006) Furthermore, cognitive mapping focuses on making a decision or solving a problem (involving a short-term viewpoint).
In contrast to cognitive mapping, GT originally focuses on building a theory, involving a long-term point of view. This, together with the possibility that GT can use any type of information, makes GT very flexible and suitable for use in the context of BBN construction.
It has to be taken into account that applying GT does not automatically lead to a good BBN. In order to obtain a BBN that is suitable for reliability prediction and management, GT should be applied in the light of the criteria that are related to reliability prediction. This may lead to adaptations of the GT approach.

Although only a very limited number of problem structuring techniques and procedures has been discussed, the procedures and techniques that are provided by GT are chosen for problem structuring in the context of BBN construction based on the evaluation in the previous paragraph. GT and its procedures and techniques will be discussed more in detail in the next sections.

4.2. Grounded Theory: Discussion and Evaluation

As identified in the previous section, GT is a research methodology that seems suitable for use in the context of BBN construction, although GT in itself focuses on the development of theory, and not on model construction. In this section, first the theory on GT is discussed in more detail. Since GT is not primarily developed for model construction, the GT method will not be followed to the full extent in the context of BBN construction for reliability prediction and management. The extent to which GT will be applied is also discussed (sections 4.3 and 1 and chapter 5), in order to address criticisms by Bryant (2002; referring to Robrecht, 1995 and Babchuk, 1996), that often the GT method is claimed, but without clear indication on how it was applied.

4.2.1. GT: Discussion of the Text-Book Method

As already indicated by Creswell (2003), GT involves multiple stages of data collection and analysis. One of the essential characteristics of GT is that it is an inductive approach, as opposed to a deductive approach (Strauss & Corbin, 1990; Easterby-Smith, Thorpe & Lowe, 1991; Locke, 2001). This means that theory building takes place based on data that are observed, rather than that a pre-defined theory is tested through data collection.

GT uses a process of continuous comparison in order to ‘update’ a theory, based on new insights from data analysis. The process of data collection and analysis is elaborated on next.

The goal of GT in general is the grounding of theory on data that is collected, mainly being observations and interviews. Data collection is guided by the data analysis process, i.e.: through data analysis, insights are gathered in the area that is researched. These insights are then used to give direction to the data gathering process (Strauss & Corbin, 1990).

In the data analysis process, first of all, ‘open coding’ takes place. This is followed by ‘axial coding’, and finally followed by ‘selective coding’. Although the different stages of the coding process are usually presented in a sequential order, the different coding phases typically take place iteratively. For an orderly discussion, the different stages are briefly discussed sequentially. A more detailed description can be found in (Strauss & Corbin, 1990).

Open coding (Strauss & Corbin 1990) is a structured way of analyzing data, leading to the identification of ‘concepts’. These concepts are abstract terms for ‘things’ (e.g. phenomena, causes, effects, properties, circumstances, etc.) that you observe in the data that you analyze. These concepts are related to each other in some way. During open coding, concepts that are related to each other are described as being a ‘category’. Through enhancement of the theoretical sensitivity (Strauss & Corbin 1990), the concepts will be given more depth by answering questions about the concepts (e.g. how, why, what, where, etc.) In this way, the concepts get more meaning, and it becomes possible to put them in the right context. This can be done through further analysis of existing information, or by collecting extra information. In this thesis, the identification of concepts will be referred to as the ‘first part’ of open coding and giving more meaning to the concepts as the ‘second part’ of open coding. Note that both parts are not chronologically related, as it is an iterative process.

Axial coding and selective coding (Strauss & Corbin 1990) are both related to the paradigm model, i.e. the central model in GT. The paradigm model is used in GT to analyse data and categories, in order to link different concepts and/or categories to each other. Whereas axial coding focuses on the direct relationships between concepts and categories, selective coding looks at a higher level of abstraction, and is focused on integrating all concepts, categories, and their relations.

The links that are identified through the paradigm model depict the different relationships between concepts and/or categories. The paradigm model describes a number of ways in which categories can be related, by predefining a number of conceptual labels that can be given to a relationship. There are six different conceptual labels (Strauss & Corbin, 1990), i.e.:

- Causal conditions (events or incidents leading to the occurrence or development of a phenomenon)
- Phenomena (central idea, event, happening to which a set of concepts is related)
- Context (the set of properties pertaining to a phenomenon)
- Action/interactional strategies (directed at managing, handling, carrying out, responding to a phenomenon as it exists in a context or under perceived conditions)
- Intervening conditions (broader structural context pertaining to the phenomenon; broad and general conditions bearing upon action/interactional strategies)
- Consequences (outcomes or consequences of action/interaction)

Next to the paradigm model, the conditional matrix also plays an important role in GT. The conditional matrix is represented by Strauss & Corbin (1990), as a set of concentric circles, depicting 8 levels that all correspond to different aspects of the world. The circle that is farthest away from the centre e.g. represents the international level, whereas the centremost circle represents the direct actions pertaining to a phenomenon. The circles are named as follows (starting from the circle farthest away from the centre; Strauss & Corbin, 1990):

- International
- National
- Community
- Organizational and institutional
- Sub-organizational, sub-institutional
- Group individual, collective
- Interaction
- Action pertaining to a phenomenon

A more detailed discussion on the paradigm model and the conditional matrix can be found in Strauss & Corbin (1990).

Axial coding focuses on the direct relationships between the different individual concepts that were found during open coding. This leads to an insight in the way in which all the different concepts that are identified through the open coding process are related and how they interact. Selective coding basically does the same, but on a higher level of abstraction (Strauss & Corbin, 1990): it identifies relationships between categories (groups of concepts) rather than relationships between individual concepts, creating an insight in these relationships and interactions.

Validation of the theory that is developed using this procedure lies in the grounding of the theory through the data that is gathered. Relations between different categories, as one initially identified from the data, are checked and validated throughout the data analysis process.

Strauss & Corbin (1990), identify three major components of which GT is composed, which are also presented by Partington (2000) on a somewhat higher level of abstraction (see Table 4), i.e.:

Table 4: Three major components of GT

Strauss & Corbin (1990)	**Partington (2000)**
Data	The nature of the data
Analytic or interpretive procedures	The analytical framework, and the basis of the theory that is built
Written and verbal reports	The form of the theory

According to Partington (2000), the major components of the original GT approach as proposed by Strauss & Corbin (1990) are interpreted as follows:

- Mainly observational and interview data
- Paradigm model and conditional matrix
- Emphasis on rich description

In order to make use of GT for the problem structuring stage of BBN construction, the procedure of GT first has to be adapted to match the elements of BBNs. For this purpose, the three components of GT need to be defined in terms of BBNs. The way in which GT has been adjusted and applied for use in building a BBN is discussed in the next section (section 4.3). Before discussing the adaptation of GT for application in BBN construction, the research guidelines and the canons of GT are discussed in the next subsection in order to be able to take the canons and procedures into account in adapting the original GT approach.

4.2.2. A critique of GT canons and procedures for BBN construction

In his paper, Bryant (2002) clearly identifies the wide environment in which GT is claimed to be used. Referring to Babchuk (1996), he identifies the fact that the GT method has been used as an 'umbrella term' by a large number of researchers. Babchuk addresses the flexibility of GT, and poses the question whether it is strength or weakness. In his study, he notes that some authors claim to have used the GT method, when they only apply one element, whereas others, who claim to have used the GT method, make careful efforts to follow all rules and dictates. Robrecht (1995; referred to by Bryant, 2002) also finds that, in management literature, many authors claim to have applied GT, without providing an indication that the method has been followed.

At the same time, referring to Orlikowski (1993), Bryant (2002) identifies that the method that is underlying the GT method offers a distinctive basis for research, and is "applicable to many other forms (next to sociology, red.) of research and in particular to anything focused on people's actions and interpretations in organizational and other social contexts."

As a result of the discussion on the use of GT in literature, it is very important to state that the GT method forms the basis for BBN construction. With respect to the elements of GT, concept generation will be used, as well as the identification of interactions. However, GT will not be adapted for the purpose of forming a theory. To stress the fact that GT is not applied as proposed by Strauss & Corbin (1990) and Corbin & Strauss (1990), the procedures and canons as introduced by Corbin & Strauss (1990) are summarized in Table 5. The extent to which these procedures and canons are met in this research is discussed in the next chapter.

Table 5: Procedures and canons of GT (Corbin & Strauss, 1990)

1. Data collection and analysis are interrelated procedures *The findings that are obtained by analysis have to guide the data collection. Concepts turn from provisional concepts into grounded concepts through repetition.*
2. Concepts are the basic units of analysis *The raw data has to be "translated" into concepts in order to build theory. The abstraction of the concepts increases as the process continues.*
3. Categories must be developed and related. *Categories can be named as names for groups of concepts. The groups of concepts have to be related between each other, to make them "real" categories: real categories are not just groups of concepts, but are related groups of concepts.*
4. Sampling in grounded theory proceeds on theoretical grounds *Sampling takes place based on concepts, not based on samples from specific groups.*
5. Analysis makes use of constant comparison *Concepts should be continuously compared for differences, similarities, etc. In this way, they can be grouped into categories.*
6. Patterns and variations must be accounted for *Patterns (structural similarities in the data), and variations (structural differences in the data) both should be identified.*
7. Process must be built into the theory *Process can mean that a concept is broken down into different steps/stages/procedures, but process may also mean that changes occur under different/prevailing conditions.*
8. Writing theoretical memo's is an integral part of grounded theory *Theoretical memo's give more detail to the conceptual analysis, because they contain the ideas and thoughts of the coder throughout the coding process.*
9. Hypotheses about relationships among categories should be developed and researched as much as possible during the research process *As hypotheses are developed, they have to be taken back into the field to be verified and validated. Throughout the coding process, hypotheses should be continuously revised.*
10. A grounded theorist need not work alone *Working together with others, may help to overcome bias, may lead to new insights and theoretical sensitivity.*
11. Broader structural conditions must be analyzed, however microscopic the research. *The analysis of the setting should also take a larger context into account (e.g. using the conditional matrix (Strauss & Corbin, 1990)), not only focus on the context which is the ground of study.*

Especially unbiased data collection, and continuous reflection on the findings are important aspects with respect to performing GT research.

It is important to consider that inductive research (and thus, GT research) takes the data as starting point, rather than existing literature, "rejecting *a-priori* theorizing" (Locke, 2001). It is therefore very important for the researcher to be unbiased in data collection, as well as in data analysis (Mansourian, 2006). This is identified by Hall & Callery (2001) as well, who state that it is important to take into account that data gathering is a process, in which the data gatherer (the researcher) plays a role, and therefore, influences the process. They stress the necessity of reflexivity, i.e.: examining one's effect as a researcher on the research process.

Although GT rejects a priori theorizing, this does not require the researcher to ignore all other knowledge (Sarker, Lau & Sahay, 2001). Rather, it warns for being driven by pre-conceptions, and focuses the attention on inducing theory from data rather than using prior knowledge for theory building.

The extent to which the canons and procedures of GT are followed is discussed in the next chapter, after applying the GT approach for BBN construction, as it is adapted from original GT.

4.3. Adapting Grounded Theory for BBN Construction

The way in which GT is adapted to suit the purposes of BBN construction is elaborated on through discussing the three components that determine the approach to GT as they are identified by Partington (2000) (nature of the data; analytical framework and the basis of the theory that is built; form of the theory; see Table 4). For all components, first the component is discussed as it appears in the original form of GT. After that, the adaptation of the component for use in the context of constructing BBNs for reliability prediction and management is presented.

4.3.1. Nature of the data

Original GT

Strauss & Corbin (1990) state that data can come from various sources, including e.g. interviews, observations, documentation (both formal and informal) and literature. As main sources (also identified by Partington (2000)), they use observational data and interviews.

Regarding the data gathering process Corbin & Strauss (1990) clearly state that "Data Collection and Analysis are Interrelated Processes". As such, in GT, the data collection process is guided by the data analysis process.

Adapted GT

The context of BBN construction is reliability prediction and reliability management in the capital goods industry, including non-technical variables, and non-physical variables. Because of the non-technical nature of factors that are taken into account in the reliability prediction process, it is more difficult to collect and use observational data for the purpose of model building for reliability prediction. Therefore, interviews are the main source of information in the context of this case study. In the context of management, this is also proposed by Partington (2000)

Whereas data collection in GT is driven by data analysis (Easterby-Smith et al., 1991; Corbin & Strauss, 1990), this is not the case in the context of BBN construction for reliability prediction and management. This is because the object of study is already known (the way in which reliability is 'created' throughout the PDP), rather than that the object of study emerges from data, as is the case in original GT. Therefore, the choice is to incorporate information from many different experts that are all involved in different disciplines related to reliability creation throughout the PLC. This means that data collection is primarily guided by covering a diverse range of disciplines and viewpoints. This also implies that there is less emphasis on constant comparison of the findings as driver for data collection.

4.3.2. The analytical framework, and the basis of the theory that is built

Original GT

As identified in the previous section, the original GT process consists of three general phases of data analysis, i.e.: open coding, axial coding and selective coding. Within the open coding phase, concepts are identified, and their theoretical sensitivity is enhanced. In axial and selective coding, GT uses the paradigm model and the conditional matrix. Using the conditional matrix, GT is able to identify the context of the theory. Using the paradigm model, GT enables the identification of the different types of relationships and interactions between concepts and between categories that play a role in the theory. However, in a BBN context, not all types of relationships of the paradigm model are relevant. Therefore, the paradigm model is adapted for use in the context of BBNs and more specifically, in the context of BBN construction for reliability prediction and management.

Adapted GT

Also in adapted GT, coding has to take place. In the case of adapted GT in the context of reliability prediction and management, open coding should be aimed at identifying the different factors (concepts) that affect reliability.

It is important to take an additional aspect of the enhancement of theoretical sensitivity (which is part of original GT), into account in adapted GT. It is important to enhance theoretical sensitivity in such a way, that it is possible to define the concepts as statistical variables (identify their state-space).

Axial and selective coding can then be used to identify the relations between the different concepts that are identified during open coding. However, rather than using the paradigm model for this purpose in axial coding, the BBN idioms are used that are introduced by Neil et al. (2000). These idioms define different types of relations that can occur in a BBN, and are used to describe the different types of relations between categories (i.e. nodes in a BBN). The idioms as defined by Neil et al. (2000) are:

1. Definitional/synthesis idiom (depicts the relationship when one node is defined by other nodes. In the BBN context, this idiom enables the prevention of the explosion of the number of conditional probabilities)
2. Cause-consequence idiom (depicts relationships between causes and consequences)
3. Measurement idiom (depicts relationships between observed and true variables)
4. Induction idiom (depicts statistical inference: relationship between historical observations and a future population)
5. Reconciliation idiom (depicts the comparison of two or more values for one variable, being outcomes of different models using different sources of evidence)

These idioms provide questions to ask at a high level of abstraction (i.e.: which idiom describes the relationship), but do not directly provide tools to answer these questions. Nevertheless, GT provides guidelines to answer these questions using the background of the relationship (idiom). For a more detailed discussion on the idioms, see Neil et al. (2000). In the context of BBN construction for reliability prediction and management, the most important idioms to be used are the cause-consequence idiom because it provides insights in the way in which causes and effects are related, and the definitional/synthesis idiom because it can help to control the number of probabilities that has to be elicited.

4.3.3. The form of the theory

Original GT

In the form of the theory as it is proposed originally, the focus is on rich description (Partington, 2000). As Strauss & Corbin (1990) identify, there are four criteria that can be used for judging the applicability of a constructed GT to a phenomenon. These are:

1. Fit: the theory should be faithful to everyday reality and carefully induced from the collected data.
2. Understanding: the theory should be understandable both for persons who practice in the area of research, and people who were studied.
3. Generality: the theory should be abstract enough and include sufficient variation to be applicable to a wider variety of contexts.
4. Control: the theory should be able to provide control regarding actions toward the phenomenon under study.

Adapted GT

The form of the resulting theory from the adapted GT approach is defined by the context of the research, i.e.: the context of reliability prediction and management using BBNs. The form of theory will be a BBN, which is based on factors that affect reliability throughout the PDP. In this way, it provides a prediction of reliability early in the PDP, and decision support for reliability management throughout the PDP. Of the four criteria that can be used for applicability of a constructed GT, three of these can be directly related to the research objectives, i.e.:

- Fit is directly related to the fact that the BBN has to be usable in everyday reality.
- Understanding is related to the fact that it should provide insights for people who use the model, and the model should be built on the information that is gathered from the people involved in everyday practice.
- Control is directly related to the fact that the model should provide insights for supporting decision making.

Since transferability of the BBN that is to be developed is not the goal of this research, generality is not considered in this thesis with respect to the adaptation of GT for BBN construction.

4.3.4. Problem Structuring in BBNs using adapted GT

The approach towards problem structuring in BBN construction using GT, can be related to the three steps for BBN problem structuring that are identified by Sigurdsson et al. (2001); (see Figure 6). These three steps (identify variables; identify network structure; express as statistical variables), as well as the way they are addressed through the adapted GT approach, are discussed underneath:

Identify variables

Identify the factors (i.e. nodes) that influence the topic by analyzing and coding the interviews.

To identify the nodes to include in the network, open coding has to take place. This basically consists of two parts (these parts are typically executed in an iterative way). First, the concepts are identified that have a bearing on the subject under study. Secondly, related concepts can be grouped into categories. Both concepts and categories can represent nodes.

The concepts and categories that are to be included in the model can be identified by counting the number of repetitions of a concept or category throughout all information sources. This is in line with the grounding of concepts as described by Corbin & Strauss (1990), which is also done based on repetition.

Identify network structure

Characterize the relationships between the different nodes using the idioms introduced by Neil et al. (2000) through analysis and coding of the interviews.

During the second part of open coding and during axial/selective coding, relations between nodes in the network are identified. The relations are represented by directed arcs in the BBN, and can be typified by the idioms introduced by Neil et al., 2000. The nodes identified in step 2, together with the arcs, represent the network structure.

In the construction of the network for reliability prediction and management, it is important that the cause-effect relations are taken into account, using the cause-consequence idiom.

The definitional idiom is very important in BBN construction for the prevention of exploding NPTs. After discussion and testing with experts, the maximum size of the NPT that they were prepared to fill in was an NPT containing 81 conditional probabilities. In order to control the elicitation burden, the number of input variables that is included in the model has to be controlled. Therefore, if more than 3 input variables should be taken up in the model, first a dummy node has to be created that combines (a number of) these nodes, using the definitional/synthesis idiom. The subject of probability elicitation is discussed in chapter 6.

Express as statistical variables

Identify the different possible states (state-space) of the variables through coding.
In the open coding process, the different nodes were identified. Further analysis of the related information may be used to identify related concepts that have a bearing on the state space for these nodes and enable us to represent them as statistical variables.
Integrating data collection and data analysis leads to a proposal for a BBN problem structuring algorithm, which is presented in Table 6.

Table 6: Proposed BBN problem structuring algorithm

1. Gather data, based on diversity of information sources (e.g. experts from different disciplines in the field)
2. Perform open coding on the gathered data to:
   - Identify concepts and categories (first part of open coding)
   - Express the variables as statistical variables (second part of open coding; enhancing theoretical sensitivity)
3. Perform axial and selective coding on the gathered data to:
   - Relate concepts and categories to each other using the cause-consequence idiom (Neil et al., 2000)
   - Control the number of probabilities to elicit, using the definitional/synthesis idiom (Neil et al., 2000)

4.4. Application process of Adapted GT for BBN problem structuring

In this section, the way in which the adapted GT approach is applied in practice is described. This is done by discussing a number of steps that are part of the qualitative research procedure, as presented by Creswell (1994, 2003). This section will elaborate on:

- The researcher’s role (subsection 4.4.1)
- Study boundaries and considerations (subsection 4.4.2)
- Data collection (subsection 4.4.3)
- Data recording (subsection 4.4.4)
- Data analysis procedures (subsection 4.4.5)
- Verification of the research (subsection 4.4.6)

As identified, there are two important reasons for applying adapted GT in the research setting. Firstly, no prior information and knowledge is available on the ‘creation of reliability’ through the PDP in the BU of the company that is studied. Therefore, a research procedure had to be used that can deal with a lack of prior information regarding the field of study. Secondly, in the setting where no prior information is available, no systematic and methodical approach is available for BBN problem structuring. Therefore, such a method had to be developed.

Combining the unavailability of prior insights and the lack of a systematic approach for BBN problem structuring in a setting where there is no information beforehand, the adapted GT approach is introduced as procedure for problem structuring in the context of BBNs. Following the guidelines of Creswell (2003), the different aspects of the research procedure are discussed in the following subsections.

4.4.1. The researcher's role

Three researchers were involved in this case study, although two of the researchers were only involved in the data collection phase. All researchers possessed general knowledge of PDPs and reliability, but did not have any detailed prior information regarding reliability and the PDP in the BU of the company under study. It is very important to note that only one of the researchers analysed the data, and that he was not present during all interviews. The involvement of more than one researcher made it more difficult to ensure consistency in data collection. In that way, it was a disadvantage. However, involvement of more than one researcher also provided the advantage of having more than one data source (i.e.: more than one data collection source), providing triangulation possibilities.

The researchers had access to the company under study through the co-called 'gatekeeper', i.e.: one contact person within the company who provides the facilities to perform the research at the company. The benefit for the gatekeeper was twofold. Firstly, he was involved in the study to gain insights in the way reliability is affected throughout the PDP. Secondly, he was involved in the study to obtain a model that enables him to 'steer' reliability, through factors that can be managed by the stakeholders in the company that are involved in the PDP.

4.4.2. Study boundaries and considerations

Setting

The study was conducted in a company that develops and produces medical scanning equipment. More specifically, the research was conducted in the BU of this company that focuses on cardio-vascular systems.

Actors

The "gate-keeper" was a principal scientist in the company. He provided facilities and access to informants as needed.

In order to build a model that can be used to predict and manage reliability, insights had to be gained in the way in which the PDP 'creates' reliability. This was done by involving people from the operational level (operational level in Figure 4) of the company. By involving people from a broad range of disciplines, a large variety of viewpoints on the PDP was collected. Since there was no prior information on the way in which reliability is 'created' in the PDP, no prior distinction with respect to level of expertise could be made between different experts. Therefore, the information that was gathered from the different experts was treated equally, i.e.: no distinction was made between high quality and low quality data.

Processes

Using adapted GT, the focus was on the PDP, and the processes and decisions that play a role throughout the PDP. Since there was no prior information regarding the way in which reliability is created throughout the PDP, all information that was collected regarding the PDP and factors that influence reliability were treated equally. This means that prior to the data collection stage, no difference was made between information that was obtained from different experts. All data that were collected in the data collection stage were treated equally, implying that both high and low quality data were used in the research. This is because there was no way of distinguishing between high and low quality of information due to a lack of prior insights.

Considerations

In order to collect data as honest as possible, the reports of the interviews with the participants in the research were treated anonymously. Moreover, the interview transcripts and reports were fed back to the participants after data collection.

4.4.3. Data collection

According to Creswell (2003), four types of information are available:

1. Observations
2. Interviews
3. Documents
4. Audiovisual material

In this study, the main source of data was face-to-face interviewing. Furthermore, some documentation was available. Since observation of the contribution of the participants to the ‘creation’ of reliability through the PDP was not possible directly and audiovisual material was also not available, these two forms of data were not usable.

One of the reasons that expert knowledge was used as main data source is that there was no historical data available on the non-technical factors affecting reliability throughout the PDP (which were incorporated in the reliability prediction method). It is important also to note that there are a number of limitations of interviewing, as identified by Creswell (2003). Firstly, interviewing is an indirect, rather than a direct information source. Secondly, information collection takes place in a different place than in the natural field setting. Thirdly, the researcher being present may cause bias in the results, and finally, not all people are equally capable of providing useful information.

Data collection was done through face-to-face interviews with a large number of experts from different disciplines. By involving many different disciplines in the study, the information base was kept broad. In this way, the influence of possible prejudices regarding model building or beforehand definition of the problem area was minimized. Due to the lack of prior knowledge, the first interviewees (from a variety of disciplines) were selected by the gatekeeper. Since the researchers have no knowledge about the company, this could not be avoided.

For the purpose of directing the research within a period of information collection, notes were used that were made throughout the interviews. This was done, rather than that a diary and memos were kept and used for this purpose (as proposed in original GT). Furthermore, throughout the research and between periods of information collection, also the periodically analyzed information from the interviews was used to guide the research.

4.4.4. Data recording

Throughout the interviews, data was recorded in two ways, i.e.:

1. During the interviews, notes were taken
2. A large number of interviews were recorded

The collected data was transformed into four types of information:

1. Reports of the points of interest that are mentioned during interviews
2. Full transcripts of the interviews
3. Point by point summaries of factors that affect reliability
4. Summaries of interviews

Next to these types of information, also the interview notes remained available as data source. Since more than one researcher could be present during data collection, the notes that were taken during the interviews were not only taken directly by the analyst, but also by other researchers. These notes provided secondary material to the analyst.

It is important that, because of the lack of prior knowledge on the subject under study in the company, the nature of the interviews was unstructured and open-ended. Regarding the interview protocol that is recommended by Creswell (2003), only two subjects of the interview were predefined:

1. A short introduction to the research and the purpose of the interview, i.e.:
   - Collecting information through interviews in order to create a model to predict reliability early in the PDP, and to use the model to manage reliability.
   - To get insights in which factors affect reliability throughout the PDP.
2. The opening of the interview, involving a question that aims at obtaining information on the PDP, and on how different factors affect reliability throughout the PDP.

4.4.5. Data analysis procedures

Using the adapted GT process, the data that were collected were periodically analyzed, i.e.: after each interviewing period, the collected data were analyzed.

Analysis took place through the qualitative data analysis computer software NVivo®. With this software, it is possible to perform coding throughout the information sources, and keep track of codes. It is possible to directly enter information, and if needed, codes can be changed.

4.4.6. Verification of the research

Important aspects that contribute to the quality of qualitative research are validity (both internal and external) and reliability (Creswell, 1994). Since the first part of the research focused on the construction of a BBN, internal and external validation focus on the constructed BBN, whereas the reliability focuses more on the BBN construction process. Creswell (2003), referring to Merriam (1998) addresses internal validity, external validity, and reliability, as representations of the trustworthiness of qualitative research. All representations are presented underneath. Next to a discussion on internal and external validity, as well as reliability, also the strategies to address these characteristics of qualitative research are elaborated on. Finally, a discussion is presented whether or not a particular strategy is applied in the case study presented in this thesis.

Internal validity (Truth value/credibility):

Internal validity concerns with the accuracy of information and the extent to which it matches reality. Creswell (2003) suggest the following strategies that have a positive effect on internal validity (in order of decreasing ease of implementation):

- Triangulation of data: data collection may take place through multiple sources and methods.

  Data was collected by three researchers, enabling data collection through multiple sources. Only one method of data collection (interviewing) was used, making triangulation through data collection methods impracticable.
- Member checks: findings can be taken back to the interviewees, so that the accuracy of findings can be checked against the real experiences of the interviewees.

  The incorporated variables were checked against a larger number of concepts that were identified through GT, but not included in the model. This validation effort is discussed in more detail chapter 5.
- Rich, thick description: a rich, thick, and detailed description of the performed research may be given so that readers may experience a shared experience in the discussion of the research.

  The application of adapted GT is described in detail, thereby providing a rich and thick description of the performed research.
- Researcher's bias: Clarification of the role of the researcher at the outset of the research.

  The role of the researcher has been clearly identified in the beginning of the research.
- Presentation of negative, discrepant information: not all information will be in account with the findings, nor does all information have to be in account.

  All information is disclosed. Discrepancies and deviations are also reported.
- Spending a prolonged time in the field: observations/data collection can take place (repeatedly) over a longer period of time.

  Data collection took place from February 2007 until September 2008 (19 months), comprehending four interviewing periods.

- Peer debriefing: Inclusion of peer examiners that review finding of the research and ask questions about it.

  Although the data collection process took place with three researchers, these researchers were not all able to also review the research findings. Therefore, it was not possible to perform peer debriefing to improve the internal validity of the research.
- Using an external auditor: involving an external auditor, who is new to the project and who can comment on the research process.

  No external auditor was available to comment on the research process, and in that way improve the research process.

External validity (Transferability)

External validity is concerned with the limited transferability of the research. External validity can be increased using the following strategies:

- Rich, thick description: a rich, thick, and detailed description of the performed research.

  The application of adapted GT is described in detail, thereby providing a rich and thick description of the performed research.
- Typicality of cases: a description of the extent to which the case is typical compared to other cases in the same category.

  Only one case study was performed. Therefore, the research could not be generalized. Generalization was not the goal of the research. Rather, the use of the adapted GT approach for BBN problem structuring was the research objective.
- Multi-site design: investigating several sites, situations and cases.

  Only one case was studied, so multi-site design does not apply.

Reliability (dependability)

Reliability is related to the limited repeatability of the research. Three different techniques can be used to ensure reliability:

- Researcher’s position: providing a detailed description of the research focus, the researcher’s role, the interviewee’s position, the basis for their selection and the context from which the data will be gathered.

  The role of the researcher has been clearly identified in the beginning of the research, i.e.: he collects data through interviews with people in the company in order to collect information for constructing a structural reliability prediction and reliability management model. Next to the researcher’s position, also the position of all participants has indicated. The way in which the position of the participants is indicated, is generic. This is because many people are suitable to participate in the research, since all people that are involved in the PDP in some way are suitable, provided that they are involved in diverse disciplines.
- Triangulation: using multiple methods for data collection and analysis.

  Data was collected by three researchers, enabling data collection through multiple sources. However, only one method of data collection (interviewing) was used, making triangulation through data collection methods impracticable.

- Audit trial: Reporting data collection and analysis strategies in detail in order to provide a clear and accurate picture of the methods used in the study, all phases of the research being under scrutiny of an experienced researcher from the field of social sciences.

  No experienced researcher from the field of social sciences was available, so audit trial was not performed.

To complete the construction of the BBN-model, additional quantitative data have to be collected about the (conditional) probabilities for each variable to be in their different states. This next phase, called instantiation as introduced in section 3.2, is not addressed in this chapter, but will be discussed in chapter 6, since the focus in this chapter lies only on the problem structuring part of BBN-building.

4.5. Summary and Conclusions

In this chapter, problem structuring is discussed as the first stage in the BBN construction process. A number of problem structuring methods are discussed, together with the way in which they cover the model dimensions as they are identified by Mitchell (1993). GT was identified (section 2.6) to provide a suitable set of techniques and procedures for application in the context of problem structuring for BBN construction. Therefore, GT is further discussed, analyzed, and finally, adapted for the purpose of BBN problem structuring. The adapted GT approach that is developed for this aim is shown in Table 6.

Table 6: Proposed BBN problem structuring algorithm

1. Gather data, based on diversity of information sources (e.g. experts from different disciplines in the field)
2. Perform open coding on the gathered data to:
   - Identify concepts and categories (first part of open coding)
   - Express the variables as statistical variables (second part of open coding; enhancing theoretical sensitivity)
3. Perform axial and selective coding on the gathered data to:
   - Relate concepts and categories to each other using the cause-consequence idiom (Neil et al., 2000)
   - Control the number of probabilities to elicit, using the definitional/synthesis idiom (Neil et al., 2000)

In order to show the applicability of this adapted GT approach, the next chapter discusses a case study, in which a BBN problem structure is identified for reliability prediction and management.

5. Applying the Adapted GT Approach on Problem Structuring: A Case Study

In the previous chapter, the researcher's role and the study boundaries and considerations were already discussed, as well as the way in which data were collected and recorded. Also, the way in which the validity of the research was secured was discussed (see subsection 4.4.6). In this chapter, a case study is presented, in which the adapted GT approach is applied. The way in which data collection and recording procedures are applied is further elaborated on in the first section. Then, in section 5.2, the application of the adapted GT approach for BBN problem structuring is discussed. Section 5.3 discusses validation of the resulting model. Then, in section 5.3, the way in which the model addresses the challenges regarding BBN construction (see subsection 3.1.1) and the extent to which the procedures and canons of GT (see subsection 4.2.2) were followed are elaborated on. In section 5.5, the effects of using adapted GT for BBN problem structuring on the instantiation stage are presented. Finally, the summary and conclusions are discussed in section 5.6.

5.1. Data Collection and Recording

In this case study, data collection was done through face-to-face interviews. For applying the adapted GT approach, 26 different experts involved in the PDP were interviewed. These experts originated from different disciplines in the field.

The 26 experts were selected based on their possible involvement in (recent) new product development activities. Furthermore, as far as possible, the interviewees came from different disciplines, and had different functions and responsibilities, so that a broad view on the problem area could be obtained from the information from the interviews (Gathering data, based on diversity of information sources as described as step 1 in the BBN problem structuring algorithm provided in Table 6, subsection 4.3.4). Through the inclusion of many different experts from many different disciplines, a large coverage of the problem area was achieved. In appendix B, the details can be found on the different experts that were involved, and their responsibilities within the company. By interviewing 26 different experts, from different disciplines, it is ensured that 'the model builder's view' is not 'imposed' and that the problem is approached from a broad perspective. It also helps to overcome bias that would be the result of a strong focus towards one specific function or discipline. The interviews took place in four stages, i.e.:

1. Over a time period of two months (March 2007 – April 2007), 7 experts were interviewed by two interviewers (interviewer I and interviewer II). In these interviews, issues that were related to reliability were identified, focusing on problems with the system, using a semi-structured interviewing technique. Reports of these interviews were written, based on notes.

2. Over a period of three months (June 2007 – September 2007), 9 experts were interviewed by two interviewers (interviewer I and interviewer III). In these interviews, the experts were asked to elaborate on factors that may affect reliability, using non-structured interviews. By using non-structured interviews, and letting the interviewee determine the flow of the interview, the influence of pre-conceptions of the interviewers was minimized. This led to a minimization of bias that is introduced in this way. Transcripts were made from these interviews, based on tape recordings.
3. In a period of half a month (May 2008), 10 experts were asked to estimate the reliability of a newly developed subsystem by interviewer III. They were asked to explicitly state their reasoning process, providing background information for their estimates (factors that they took into account). The interviews were structured interviews, focusing on obtaining an estimate of the reliability of the newly developed subsystem (see Houben, Sonnemans, Stollman & Newby, 2009). The information on the factors that they took into account in the estimation process was written down in reports, based on notes and recordings.
4. Finally (September 2008), 2 experts that were already interviewed once (in April 2007 and June 2007) were interviewed again by interviewer III. In these interviews (as in the interviews performed by interviewer I and interviewer III), the experts were asked to elaborate on factors that may affect reliability, using non-structured interviews (like in period 2). In his way, the influence of pre-conceptions of the interviewers was again minimized together with the bias that is introduced by guiding the interview. Summaries of these interviews were made based on notes and recordings.

   The reason why the experts were interviewed twice, was that their expertise was directly related to the physical development of new systems (since they were system architects), and therefore they were expected to provide much information. The extra interviews were not regarded as extra information source, but as extra information from the same source.

In total, a set of documents was obtained containing 9 full transcripts, 9 reports, and 10 summaries of factors obtained from a reliability estimation process regarding a newly developed subsystem (Houben et al., 2009). All documentation remains available for further analysis.

The interviews were conducted by three different interviewers (together or individually). The analysis of the information however, was done by only one of these three interviewers (interviewer III). Interviewer III not being present at the first interviewing period may have led to an inconsistency in data collection. Nevertheless, by making sure that there was an overlap in the data collection (interviewer II was present during both the first and the second interviewing period), problems with inconsistency of data collection were minimized. In this case, involvement of interviewer II in the first and second interviewing period provided to a large degree consistency over these two periods. Involvement of interviewer III throughout the second, third and fourth interviewing period provided in a great measure consistency over these three interviewing periods. Hence, to a large extent, consistency was provided throughout all interviewing periods.

As there were no prior indications of different quality of the reports and/or transcripts, all data were treated equally.

For the purpose of directing the research, within a period of information collection, notes were used that were made throughout the interviews. Furthermore, the dates of the interviews were kept. Throughout the research and between periods of information collection, the periodically analyzed information from the interviews was used to guide the research, rather than that a diary and personal memos were used for this purpose.
Furthermore, because of practicality reasons, often a number of interviews were held on the same day, or shortly after each other. Since direct analysis of the gathered information was not possible, the information was gathered and analyzed in batches. This led to periodical comparison, rather than constant comparison.

The adapted GT approach is discussed by demonstrating and discussing the different steps. The results can be found for all stages of the PDP (C&D stage, D&D stage and M&I stage) in appendices C, D and E. A full account of the application of the adapted GT approach is presented for one stage of the PDP only (i.e. the C&D stage) for reasons of conciseness. This section includes a presentation of the model nodes (subsection 5.2.1), their definitions and their possible value ranges (i.e. state-space) (subsection 5.2.1) and their relationships (subsection 5.2.2). In section 5.3, the suitability of the approach for BBN problem structuring is further evaluated.

The coding process of GT (i.e. open, axial and selective coding) is adjusted to the needs of BBN construction. For the first part of the open coding stage, nothing has to be changed, i.e.: interview-data are analyzed in order to identify concepts (which may represent nodes)
The second part of the open coding stage and the following two coding stages however, – axial and selective coding – will be merged and applied only to a limited extent, because the different types of relationships that are identified through GT (connected to the paradigm model) are not all relevant (as indicated in the previous section). Matching the GT approach to the task of BBN building, nodes (related concepts) are searched for that depict a ‘definitional/synthesis’ relation or a ‘cause-consequence’ relation. Relating the GT approach to BBN construction clearly limits the extent to which GT is – and has to be – applied.
In the following subsections, the four steps of the approach are discussed in detail. The discussion includes a description of the way in which the different steps were applied in the case study, and how difficulties were handled.

5.2. Data Analysis

Steps 2 and 3 in the BBN problem structuring algorithm (provided in Table 6, subsection 4.3.4) relate to data analysis. The steps are discussed in subsections 5.2.1 respectively 5.2.2.

5.2.1. Identifying the variables and their state space

The nature of the 28 documents was such, that the coding process was not straightforward. The 18 reports of interviews were not literal transcripts of interviews, but summaries, and it was not possible to obtain the original recordings of the interviews to make such transcripts. So, a first round of "open coding" had already taken place writing the reports of the interviews, since the information that was gathered in the interviews was already condensed into summaries. In this way, information from the interviews was already on a higher level of abstraction. As a consequence, many of the concepts that were identified during the open coding could be typified as nodes to be included in the model.

Another consequence of the high abstraction level of the information was the fact that it was not possible to identify the state space through enhancing theoretical sensitivity in the second step of open coding. Therefore, identification of the state space took place through direct conversations.

Using the open coding process as described in Corbin & Strauss (1990), 339 nodes (random variables) were initially identified. Below, an impression is given of the way in which the coding process was applied to the interview material, by showing small sections of different interviews and identifying the associated nodes (see also Table 8).

> *A: "...Four years are gone and it is expected to start failing in four years. It is not possible to convince management to handle this problem now..."*
> *A: "The interviewee volunteered to participate in the reliability activities next to performing his normal responsibilities."*
> *B: "...Now we have to secure that quality is designed in from the start. So that's another approach..."*

The three pieces of interviews above, originate from two interviews, indicated A and B. In all pieces of text, the words that relate to the same node are underlined. The node that reflects the underlined text is named 'approach towards reliability'.
The first underlined text of the first interviewee (*not possible to convince management*) relates to 'approach towards reliability' since it states that reliability is not important enough for management to react on the problem. The second piece of text from the first interviewee (*participate in the reliability activities next to his normal responsibilities*) confirms this viewpoint. The second interviewee indicates that a change has taken place in the approach towards quality (*that's another approach*). The combination of these individual concepts through their commonality resulted in the node 'approach towards reliability'.

During this open coding process, 339 nodes (i.e. factors that influence the reliability) were identified. In general, it is possible to include such a large number of variables in a BBN, provided that the complexity remains comprehensible and that sufficient data is available to derive all conditional probabilities in the BBN.

During the identification of the variables, it became clear that these 339 nodes affect reliability in different stages of the PDP, leading to the construction of a separate BBN for each stage as a first step to reduce the number of nodes. Bedford et al. (2006) distinguish three stages, i.e. C&D, D&D and M&I. Because of this, the different nodes were allocated to these main stages, using the definitional/synthesis idiom: the effect on reliability of all nodes was defined as the combined effect (synthesis) of the reliability effects in all three stages of the PDP. Although this reduces the complexity of the model greatly, the resulting 'sub-models' still remain very complex.

Because data availability is fully dependent on the experts (in this case study, the only available data comes from experts), whose availability is scarce and costly, the probability elicitation burden had to be limited. As identified in literature elicitation of probabilities can take up to 30 minutes per number (Druzdzel & Van der Gaag, 2000), although more optimistic assessments can be found: Van der Gaag et al. (2002) state that it is possible to elicit 150-175 probabilities per hour. Since in this study the availability per individual expert for probability elicitation was limited to 3 hours, this would lead to a maximum of 450-525 probabilities that could be elicited (using the optimistic estimation). Because the probability elicitation burden is amongst others dependent on the number of nodes included in the BBN, this number had to be limited. In this case, the number of nodes per network was limited to around 10. This number was chosen, as a trade-off between the elicitation burden, together with practicability of the model on the one side, and the insight in the way in which reliability is 'created' in the PDP on the other side. This resulted in the following elicitation burdens for the different stages of the PDP: C&D stage: 10 nodes, 189 probabilities; D&D stage: 14 nodes, 287 probabilities; M&I stage: 8 nodes, 207 probabilities. As indicated, the BBN structures for all different stages of the PDP can be found in appendix A.

In order to select the nodes to include in the BBN, they were first ranked. This was done based on the number of experts that had addressed any of the underlying concepts and/or nodes directly, and on the number of times the concepts and/or nodes were addressed in total (i.e. including repeated mentioning of a concept). The top of the list of concepts for the C&D stage is shown in Table 7.

Table 7: List of most mentioned concepts for the C&D stage, together with the number of people (sources) and number of times (References) that they were mentioned in total

Concept name	Sources	References
setting correct specifications	11	23
system variety	8	16
introduction of new components or subsystems	7	10
focus on TTM (time-to-market)	7	10
focus on reliability	6	14
configuration options	6	10
focus in the development process	5	14
focus on new features	5	7

In addition to using the number of people that mentioned a concept as criterion, also a process of discussion on inclusion or exclusion of each node in the model was performed. This was done, because some concepts were very similar. An example is given by not including both 'focus on reliability' and 'focus on TTM'. Although these concepts were often mentioned (and were therefore high on the list of concepts to include), they were combined into the variable 'Approach towards reliability' (derived from 'focus on reliability'). The reason for this is that, if the focus of the company would lie on TTM, then it automatically would not lie on reliability. The effect of the focus on TTM would then be reflected by a lack of focus on reliability. By including 'Approach towards reliability', implicitly 'focus on TTM' is also included in the BBN.

The nodes were included one by one, starting at the top of the priority list and from there working down the list. In this phase, the variables were thoroughly discussed and if needed, definitions and/or names were revised in order to make them usable in a BBN. As example, the final list of nodes that are to be included in the BBN for the C&D stage, together with a short description, is presented in Table 8.

Table 8: Descriptions of variables that are included in the BBN for the C&D stage

Variable name	Description
Approach towards reliability	The priority level of reliability for the company
Mindset towards reliability	How the people think about reliability throughout the PLC
Attention towards reliability	The attention paid towards reliability, as a combined effect of the variables "approach towards reliability" and "mindset towards reliability"
Requirements	How good the designers are known with the customer wishes and the ability to translate these to system and subsystem level, taking into account what the function are, under which circumstances they has to perform and on what kind of level
Component specifications	The quality of the specifications of the components
Setting specifications	The quality with which specifications are set, as a combined effect of the variables "requirements" and "component specifications"
Introduction of new components	The percentage of the total functions of the product that is new
Configuration variety	Number of different possible configurations per system type
Adding features	Number of features and functions that is added to the specific product, looked at in the light of previous developments.
System diversity	This variable represents the changes in the system that are either related to the configuration variety or to the introduction of new components, taking both functionality and hardware of the system into account

With respect to the ranking criteria, clearly, other criteria could be chosen, which could possibly lead to a different network including different nodes. Moreover, it is possible to add nodes to or remove nodes from the network during model checking and validation. This enables the model builder to introduce more and/or other criteria for nodes to be included in the network, and in this way create a model that meets the model builders' requirements. The full list of concepts and nodes remained available.

By enhancing the theoretical sensitivity (second step of open coding) the different possible states and the related potential values for the nodes could be identified. In this way the state-space could be defined of the statistical variables that are represented by the nodes.
Although it was possible to define a generic state space for each variable, the abstraction level of the information from the interview reports was such that in most cases it was not possible to directly identify an unambiguous state space for the variables. Therefore, the state space for most variables was in terms of 'good – average – bad'; 'high – medium – low' (ordinal scale), or 'yes – no', etc. (nominal scale).

Furthermore, it turned out that identifying the state in an unambiguous way from the documents only, was not always possible. A good example is found in the node FMEA, in the D&D stage. In the transcript of one interview, the states for this concept were identified as "being done", or "not being done". However, in another interview, the states were indicated as being done "well", "average", or "poorly". Although in both interviews, the related concept was FMEA, it was difficult to conclusively assign possible states to the variable in order to satisfy the corresponding criterion for BBN variables, which have to have a "finite set of mutually exclusive states" (Jensen, 2001). This was also pointed out by Kjaerulff & Madsen (2008).

As a consequence, extra information had to be gathered. This information was gathered through a direct discussion about possible values of the variables. This led to the identification of an unambiguous state-space for most variables. In case if the state-space of a variable could not be defined unambiguously through interview analysis and discussion, the experts had to interpret the state-space. Again, the variable 'attention towards reliability' is taken as example. It was left to the interviewees to determine how the combination of these variables determines the value of the synthesis variable 'attention towards reliability' (i.e. what high, medium or low attention towards reliability exactly means). The resulting state-spaces for the C&D stage variables are represented in Table 9.

Table 9: Variables that are included in the BBN for the C&D stage and their state-spaces

Variable name	State space
Approach towards reliability	1. Always on the agenda of the MT-meeting 2. Ad-hoc on the agenda of the MT-meeting (due to problems) 3. Once a year on the agenda of the MT-meeting
Mindset towards reliability	1. Clear activities, deliverables & evidence during the Product Creation Process (PCP) 2. Activities and deliverables are present, but no evidence 3. Ad-hoc (due to problems) determining activities and deliverables

Attention towards reliability	1. High 2. Medium 3. Low
Requirements	1. Customer wish translated to system and sub-system level. And insight in the functions influencing this process. 2. Customer wish translated to system level, partly sub-system level and ad-hoc knowledge with respect to the functions 3. Not able to translate requirements to system level
Component specifications	1. Knowledge on component & the conditions under which it has to perform is present 2. Knowledge on component & NO knowledge on the conditions is present 3. No knowledge on component & no knowledge on the conditions is present
Setting specifications	1. Good 2. Average 3. Bad
Introduction of new components	1. <10% 2. 10-30% 3. >30%
Configuration variety	*1.* Large *2.* Average *3.* Small
Adding features	*1.* Many new features & functions *2.* Average number of new features and functions *3.* Small number of new features & functions
System diversity	1. Large diversity 2. Average diversity 3. Small diversity

5.2.2. Characterizing relationships between variables using BBN idioms

Regarding the characterization of the relationships between variables, both the analysis of the interview information and the discussions about the nodes played a very important role, since the relationships between these variables became clearer. In order to give an example of the way in which relations were identified in the coding process, the following pieces of text are looked at:

> *C:"Maybe the real problem is that the X-Ray tube has different specifications then the requirement the company put for the system."*
> *D: "The tube has to deliver certain power; however there is more power asked during use then the tube is designed for. Reliability issues are occurring in the form of power drop (less X-ray intensity). When more output is required than the tube can deliver the filament of the tube wears out quickly."*

In both pieces of text, reliability problems are related to specifications, as well as requirements. This has led to the identification of the cause-consequence relation between the node 'component specifications' and reliability, and between the node 'requirements' and reliability.

Next to the cause-consequence idiom, also the definitional/synthesis idiom was used often, to control the number of probabilities that had to be elicited. This was done when more than 3 variables affected 1 other variable. An example of the use of the definitional/synthesis idiom can be found again in the synthesis of the nodes 'component specifications' and 'requirements' into one node: 'setting specifications'. Another example of a 'definitional/synthesis' idiom is the node 'attention towards reliability', which is defined by the combination of 'approach towards reliability' and 'mindset towards reliability' after discussion about the nodes. Note that the definitional/synthesis idiom is only usable for the purpose of controlling the number of probabilities to elicit, if the parent nodes that are synthesized are independent of other nodes (that are not synthesized).

Because the interview information was already at a higher level of abstraction, it was not always possible to identify the relations between concepts from the interview information only. Therefore, discussion played an important role: a large number of relations between concepts have been identified during the discussions.

The resulting BBN problem structure for the C&D stage can be seen in Figure 11.

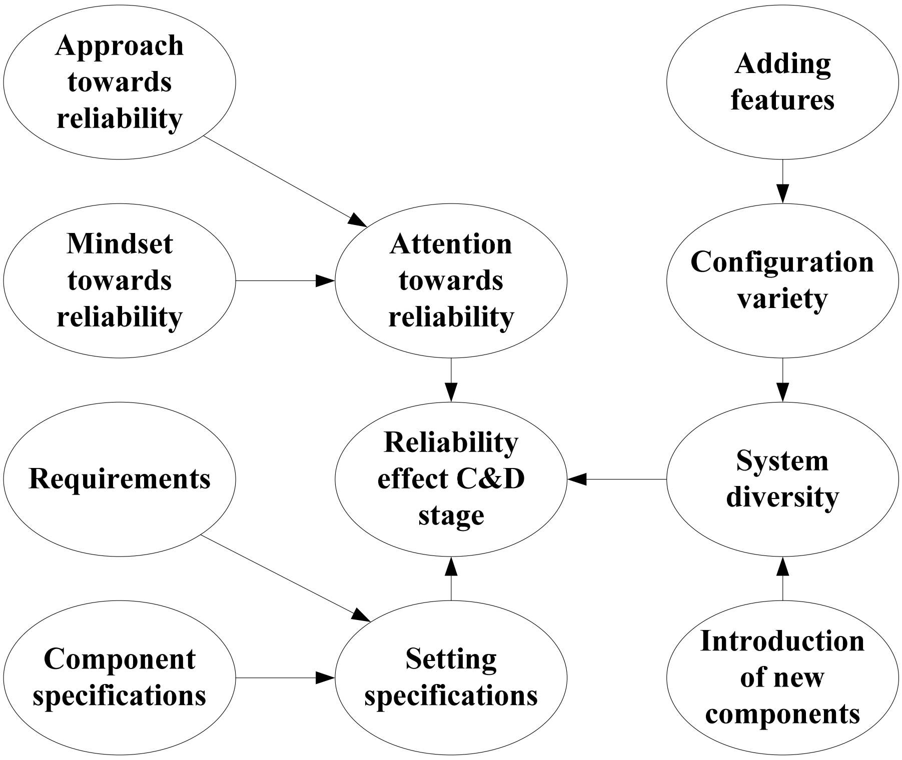


Figure 11: Resulting BBN for the C&D stage

5.3. Reflection on the Approach used for BBN Problem Structuring

In the case study, the step by step approach based on GT was applied. This approach was introduced in section 4.3 (Table 6) to overcome the challenges with respect to BBN problem structuring (see subsection 3.1.1). In this section, first the extent to which this approach deals with these challenges will be discussed in subsection 5.3.1. After that, in subsection 5.3.2, the evaluation of the adapted GT approach will take place based on the extent to which it follows the procedures and canons introduced for original GT, presented in Table 5 (Corbin & Strauss (1990))

5.3.1. Addressing the challenges with adapted GT

The BBN problem structuring algorithm, provided in Table 6, section 4.3.4 proved to enable the identification of problem boundaries as well as the identification of model variables. Through the data collection procedure described in section 5.1, many different data sources were included, covering many different viewpoints from people from a broad range of disciplines. In this way, it was made sure that the boundaries of the problem were not pre-defined.

Furthermore, identification of the variables took place based on the number of times that a concept was repeatedly mentioned throughout data analysis. This criterion for incorporating a variable was based on the criterion provided by GT, to call a concept grounded, based on the repetition of the concept in the analyzed data.

Finally, by limiting the maximum number of variables to include in the BBN, and by basing the inclusion of variables on the number of repetitions, the boundaries of the problem emerged through the process, rather than being pre-defined. In this way, both challenges that were identified in subsection 3.1.1 were overcome.

It is important to note that, with respect to the application of the adapted GT approach, an important problem arose. In the case study it was found that it was not possible to define the state space of the variables solely on the basis of the open coding process. This problem was solved by gathering additional information. This collection of extra information to determine the variables as statistical variables provides an addition to the BBN construction algorithm presented in Table 6.

5.3.2. Evaluating application of the adapted GT approach

In this subsection, the way in which the adapted GT approach was applied is discussed. This discussion is split up into three parts. First, the list with 11 canons and procedures, introduced by Corbin & Strauss (1990) is discussed in Table 10. After that, some additional attention is spent to data gathering and theorizing. Finally, a discussion is presented on the way in which the proposed BBN problem structuring algorithm was applied.

Table 10: Extent to which procedures and canons of GT (Corbin & Strauss, 1990) are followed

1. Data collection and analysis are interrelated procedures *The findings that are obtained by analysis have to guide the data collection. Concepts turn from provisional concepts into grounded concepts through repetition.* Because of the aim to take a broad view on the problem area, and to take many different viewpoints into account from many different disciplines, data collection is not directed, and interviews in one discipline will not be guided by findings in another discipline. The criterion repetition, which is used as a criterion to turn provisional concepts into grounded concepts, is used as criterion to judge whether to include concepts in the BBN.
2. Concepts are the basic units of analysis *The raw data has to be "translated" into concepts in order to build theory. The abstraction of the concepts increases as the process continues.* Abstraction was already on a high level because of the information gathered in the interviews. Further abstraction of the concepts was done, based on common objective (see the example given for 'Approach towards reliability', subsection 5.2.1)
3. Categories must be developed and related. *Categories can be named as names for groups of concepts. The groups of concepts have to be related between each other, to make them "real" categories: real categories are not just groups of concepts, but are related groups of concepts.* Relations ran through the activities within the PDP, and through the different stages of the PDP. As such, they were related to each other, amongst others, over time.
4. Sampling in grounded theory proceeds on theoretical grounds *Sampling happens, based on concepts, not based on samples from specific groups.* We have pre-defined the sample, i.e.: different people throughout the PDP, over the breadth of the company (many different functions and disciplines within the company).
5. Analysis makes use of constant comparison *Concepts should be continuously compared for differences, similarities, etc. In this way, they can be grouped into categories.* Due to the abstract nature of the information from the interviews, comparison of the concepts was possible to a limited level of detail (not to a high level of detail).
6. Patterns and variations must be accounted for *Patterns (structural similarities in the data), and variations (structural differences in the data) both should be identified.* This has not appeared, although it became apparent that different phases of the lifecycle formed the "thread" through the process of data analysis. Although relationships were identified, the definition of the relationships was left for the instantiation stage.
7. Process must be built into the theory *Process can mean that a concept is broken down into different steps/stages/procedures, but process may also mean that changes occur under different/prevailing conditions.* The process is reflected in the PDP and its time relation.

8. Writing theoretical memo's is an integral part of grounded theory *Theoretical memo's give more detail to the conceptual analysis, because they contain the ideas and thoughts of the coder throughout the coding process.* As already indicated in section 5.1, data collection took place systematically, and was dated. The tentative results of the data analysis were used as guideline for further analysis. Throughout the research and between periods of information collection, the periodically analyzed information from the interviews was used to guide the research, rather than that a diary and personal memos were kept and used for this purpose.
9. Hypotheses about relationships among categories should be developed and researched as much as possible during the research process *As hypotheses are developed, they have to be taken back into the field to be verified and validated. Throughout the coding process, hypotheses should be continuously revised.* Hypotheses are suggested in the field through suggestions of BBNs. Verification and validation of the findings is done based on evaluating the BBN structures with interviewees. Also, a survey is used for model validation (and validation of the findings)
10. A grounded theorist need not work alone Data gathering was done with three persons, but data analysis was done individually.
11. Broader structural conditions must be analyzed, however microscopic the research. *The analysis of the setting should also take a larger context into account, not only focus on the context which is the ground of study.* There is awareness of the larger context of the research. However, this research focuses on a single case, and no external influences are claimed and/or identified.

It was indicated earlier that the way in which data were gathered was not necessarily guided by data analysis. Rather, data was gathered through face-to-face interviews, where the selection criterion regarding which experts to include in the data gathering process was based on their area of expertise, i.e.: the discipline in which they were engaged. In this way, the problem of which variables to include in the BBN, was approached from a broad perspective. Although this does not guarantee that the researcher is unbiased in the data collection and analysis (which is important in the context of GT, see Mansourian, 2006), the inclusion of many different perspectives helps to keep an open mind to new information.

Although GT originally focuses on theorizing, it was used for the purpose of model building. The GT approach was chosen for the construction of the BBN because of its inductive nature, i.e.: the chosen approach supports inducing the model from the collected data, rather than basing the model on pre-conceptions (as was warned for by (Sarker et al., 2001)). In this way, imposing the model builders' view throughout model construction can be avoided.

It is important to note that the proposed approach based on GT could not be followed literally. Throughout the problem structuring stage, some deviations were necessary in order to construct a BBN problem structure. These deviations are taken into account when the full algorithm for BBN construction is presented in section 6.7.

5.4. Validation of the Resulting BBN Problem Structure

Validity and reliability considerations regarding the construction of the BBN were discussed already in section 3.1.2. Also, the strategies for obtaining internal and external validity, as well as reliability were discussed there.

In this section, the validation of the resulting constructed BBN is discussed. Since the validation of an existing model (the constructed BBN) is defined as deductive instead of inductive research, other tools can be used to validate the BBN.
Validation of the network took place using a survey method, based on the surveys presented in (Zhang & Pham, 2000; Jacobs & Van Moll, 2007).

As input for the survey, the results of the application of the adapted GT process were used. The survey contained a list of potential model variables that was obtained from the application of the adapted GT process, together with their description (similar to the ones presented in Table 8). The number of variables that was given for evaluation to the participants was twice the number of variables included in the model. The relations in the various models that are definitional/synthesis relations did not have to be validated explicitly.

Before the survey was sent to the experts, it was first evaluated by the ‘gate-keeper’ in the company. This led to changes in the survey regarding the potential model variables that were included in the survey. The first list of model variables contained a number of potential variables that partly overlapped. To the extent possible, these were removed from the list and replaced by others (which did not overlap, or overlapped less). Furthermore, definitions were accentuated, so that differences became clearer.

After this, the survey was presented to all 26 interviewees that were involved in the four rounds of interviews (see appendix B). The survey can be found in appendix G. In the survey, for all potential model variables, it was asked how significant the influence of the variable is on reliability. This was done on a 7-point Likert scale, ranging from ‘no influence’ (1) to ‘very much influence’ (7). Although the expectation is that the experts will not use the first ranking (‘no influence’), because the variables were obtained from the GT analysis, it is included to provide a full coverage of all possible outcomes.

The problem of including the rank ‘no influence’ in the 7-point scale, is that it is does not represent the other end of the spectrum, when comparing it to the last rank (‘very much influence’). If the one end of the spectrum is ‘very much influence’ (rank 7), then the other end of the spectrum would be expected to be ‘very little influence’. In the scale, this other end of the spectrum is now represented by rank 2, although this is not explicitly indicated in the survey. Although the ranking scale that was used in the survey was based on the survey scale as it was represented by Zhang & Pham (2000), this is a weak point in the survey. Unfortunately, since this weak point was identified after the survey was already sent out to the participants, it could not be improved. For the purpose of ranking the variables based on their perceived importance, it is not expected to have much impact. This is because it is clear

that the ranking from 1 – 7 indicates an increasing level of influence of the factor. Therefore, comparison of the values of the variables with each other is still possible.
By using a 7-point scale, the experts were provided with sufficient possibility to distinguish between levels of significance. As was the case in the survey presented in (Zhang & Pham, 2000), also the option "no opinion" was provided. This was done in case an expert considered himself not knowledgeable about the potential influencing factor.

The survey was sent to the experts by e-mail, in which it was indicated that the survey would take about 15 minutes (the time needed as estimated by the gate-keeper). After 28 days, 2 reminder e-mails had been sent to all participants, which resulted in 9 replies. This equals a response rate of 35%. Unfortunately, this response rate is not high enough for model validation. This is because the group that participated in the GT study was a very heterogeneous group, i.e.: people from many different disciplines were included, which means that a response rate near to 100% (Leslie, 1972) would be required for validating the problem structure through a survey.

One of the participants provided a reply after being involved in a group meeting on the adequacy of model as it was developed based on the adapted GT approach (the inclusion of variables being based fully on the number of times that they were mentioned in the interviews). Because inclusion of this survey would bias the results, the input from this survey was not taken into account.

In the analysis, the scores of the identified variables were averaged, and ranked according to their score, a higher score leading to a higher rank. Finally, the reliability of all separate models was determined, by looking whether the variables that ranked highest, were also the variables included in the model. By including twice as many variables in the survey, the measure of reliability (Miles & Huberman, 1994) (the level of overall agreement of the experts with the model variables) can range from 0 (full disagreement) to 1 (full agreement), the measure for reliability being determined as follows:

$$\text{Reliability (Re)} = \frac{\text{Number of agreements (Na)}}{\text{Total number of agreements + disagreements (Na+Nd)}}$$

Since twice as many potential variables were included in the survey, as there were included in the model, the top half of all ranked variables would be included according to the survey. When a variable that is *now* (based on adapted GT) included in the model is *also* included in the model based on its ranking in the survey, there is agreement. When a variable that is *now* (based on adapted GT) included in the model, is *not* included in the model based on its ranking in the survey, then there is disagreement.

For full model validation, this led to a survey that was filled in by the individual experts on the significance of the influence of 42 variables:

- 12 for the C&D stage
- 18 for the D&D stage
- 12 for the M&I stage

The outcomes are presented in Table 11 underneath.

Table 11: Level of agreement regarding the constructed model reliability of all PDP models

PDP stage	**Agreements (Na)**	**Disagreements (Nd)**	**Reliability (Re)**
C&D	2	4	0.33
D&D	6	4	0.6
M&I	4	3	0.57
Overall	12	11	0.52

The results of the analysis, total reliability being 0.52, show that there is no agreement, i.e.: the variables that were ranked highest by the people that filled in the survey were not the same variables that were found through applying the adapted GT approach.

One possible explanation of the lack of agreement is that the number of people that replied to the survey was very limited. This might lead to a tendency toward identifying variables that are typically related to the area of expertise of this limited number of people, leading to a biased view on the importance of potential variables. Furthermore, since the adapted GT approach was developed to prevent the model from being pre-defined, this does not necessarily show whether or not the adapted GT approach is suitable for identifying the most important or most influential variables to include in the reliability prediction and management model.

A possible explanation may lie in the individual nature of the GT approach which is inherent to the strategy that is part of the GT approach. Rather than identifying the variables using a group meeting, the views of single individuals are collected. This approach was chosen because it provides the advantage of collecting a very broad spectrum of knowledge, since the individual is not influenced by ′group thinking′. However, this also means that the individuals are not stimulated by other individuals to think in other directions than their mindset at the moment of the interview, which is a drawback of the approach. The fact that a participant does not mention a potential variable when interviewed individually does not necessarily mean that he judges the variable not to be important, since there may be many reasons for an interviewee not to mention a potential variable.

A clear conclusion regarding the validity of the model cannot be drawn from the results of the survey, i.e.: the model cannot be fully validated, but neither can it be falsified. It can only be said that the variables that are included in the model are all moderately important (although the model does not necessarily include the most important variables of the survey as filled in by the participants).

Next to using the survey purely for validation purposes, the survey presented above could also be used to determine which variables to include and which not (instead of using the number of repetitions of a concept in the data analysis). In this way, another approach is provided for determining which variables to include and which not. It still remains an issue that the limit on the number of nodes that is included in the BBN has to be chosen by the researcher. The researcher will still have to make the trade-off between practicability of the model and the insights that it provides.

5.5. Consequences of the Approach for Problem Structuring for Instantiation

In the next stage of the BBN construction process, the instantiation stage, the CPTs for the BBN have to be determined, and the probabilities have to be obtained. Since no objective data are available, the CPTs in the BBN have to be obtained through probability elicitation, where probabilities have to be estimated by experts. The number of nodes to be included in the BBN has to be limited, because of the exploding elicitation burden, i.e.: the more nodes and relations that are included in the network, the more probabilities have to be elicited ("The number of probabilities to be assessed for each variable in the belief network is therefore, in general, exponential in the number of parents of the variable"; Renooij, 2001). Consequently, in order to limit the elicitation burden, not all potential nodes should be incorporated in the network.

Although the adapted GT approach is able to meet the challenges that are identified with respect to problem structuring, it is also important to limit the number of nodes to be incorporated in the model. This is because the elicitation burden is directly dependent on the number of nodes that is incorporated in the model. The identified trade-off that is to be made between the level of detail on the one hand, and practical manageability on the other hand, affects the instantiation stage.

The use of adapted GT has to trade off the level of saturation against practical manageability of the model and elicitation burden. The risk of applying GT that has to be taken into account is that it may lead to pursuing a too high level of saturation, leading to intractable BBNs, as well as a very large elicitation burden that cannot be carried by the experts, their availability being limited.

Management of the elicitation burden is possible, e.g. through aggregation using the synthesis-definitional idiom. This has been discussed in section 4.3.

5.6. Summary and Conclusions

In this chapter, a case study is presented in which the problem structure of a BBN is constructed using a systematic approach based on GT. The approach introduces systematization into the BBN-construction process: i.e. the iterative application of the different steps in the coding process based on GT.

Using the approach, the first research question could be answered, related to systematically identifying and defining system boundaries without prior insights. By taking a large number of views into account (involving people from different disciplines), prejudice and limitations in BBN building can be kept off, in this way avoiding pre-fixed ideas of model builders in problem structuring.

The second and third research question (pertaining to the systematic identification and definition of variables in the BBN, and the systematic identification and definition of the BBN structure) are partly addressed in the approach applied in the described case study. In the case study, it has been found that the proposed approach based on GT cannot be followed literally, and that some deviations are necessary in order to construct a BBN problem structure. These deviations are taken into account when the full algorithm for BBN construction is presented in section 6.7.

Definition of the of the variables and structure in quantitative terms (through the identification of the values and CPTs related to the network) is not addressed through the adapted GT approach applied in the case study described in this chapter. This topic is treated in the next chapter.

Within the larger context of constructing a BBN for reliability prediction already early in the PDP, it can be seen that the approach that is introduced in this chapter enables the methodical construction of the problem structure for a BBN where only very limited information is available (which is the case early in the PDP).

6. BBN Construction: Instantiation

In this chapter, the way in which the probability elicitation process (which is used to define variables and relations) is designed is discussed.

In order to design a probability elicitation protocol, the first section of this chapter discusses the literature on probability elicitation. Then, in section 6.2, the way in which the probability elicitation process is organized so that it addresses the main challenge regarding probability acquisition is discussed. The practical application of the elicitation process is then presented in section 6.3. The probability elicitation approach as introduced in this chapter is reflected upon in section 6.4. Section 6.5 treats the repeatability of the probability elicitation process. The summary and conclusions are presented section 6.6, and the chapter is closed with the presentation of the BBN construction algorithm, based on adapted GT and the probability elicitation method introduced in this chapter.

6.1. Probability Elicitation: Discussion of Literature

Regarding probability elicitation, there are two important subjects that need to be discussed. These are firstly, the different probability elicitation methods that are available, and secondly, the points of attention that have to be taken into account throughout the probability elicitation process. Different elicitation methods will be discussed in subsection 6.1.1, whereas the points of attention that play a role throughout the probability elicitation process will be discussed in subsections 6.1.2 – 6.1.4.

In general, three types of probability elicitation methods can be identified: direct methods, indirect methods, and parametric methods (Cooke, 1991), the latter one only being related to continuous probability distributions. In the next subsection, examples of direct and indirect methods are discussed. Since the BBNs that are described in this thesis are solely based on discrete probability distributions, parametric methods are not discussed in this thesis.

6.1.1. Probability elicitation methods

In this subsection, two subtypes of direct elicitation methods and two subtypes of indirect elicitation methods are elaborated on.

Direct elicitation method: direct enquiry

One way of asking for a probability estimate, is simply asking directly for the probability. Spetzler & Staël Von Holstein (1975) identify that using an indirect way of elicitation is preferable over direct elicitation. As a reason, they identify that respondents often experience difficulty in providing direct numerical probability estimations. For the respondents that feel capable of giving direct numerical probability estimations, they state that later, it is found that they have little confidence in their estimates. The preference of indirect over direct elicitation methods is supported by Cooke (1991), who states that asking questions that directly address the probability of an event in this way, is the most common way of probability elicitation, but equally so the worst.

Looking at the argumentation of Spetzler & Staël Von Holstein (1975), the preference of indirect elicitation methods over direct elicitation methods is mainly based on the respondents' ability to provide direct numerical probability estimates.

Direct elicitation method: probability scale

In direct elicitation using a probability scale, either a horizontal or a vertical line can be used, together with 'numerical anchors'. An example is represented in Figure 12, where a line is presented together with the numerical anchors, indicating 0%, 25%, 50%, 75% and 100% probabilities.

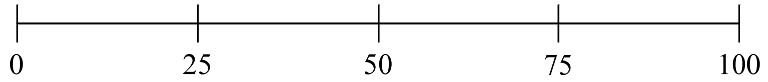


Figure 12: A horizontal numerical probability scale (Renooij, 2001)

Renooij & Witteman (1999), in response to the reluctance with which experts express estimates, have developed a set of 7 verbal expressions of probability, related to a set of 7 numerical expressions of probability. This has lead to a so-called verbal-numerical scale (see Figure 13). The verbal expressions that are introduced at the left hand side of the scale are not direct translations of the numerical values, and are therefore not in one direct line with the numerical expressions on the right hand side.

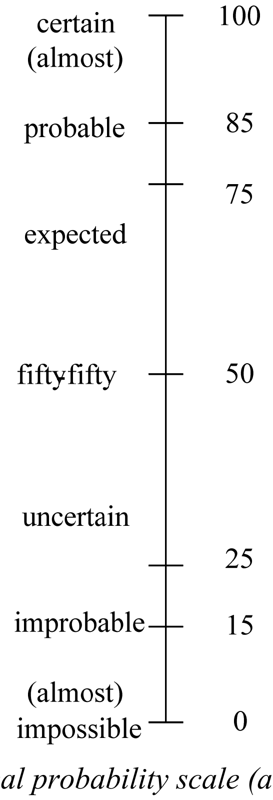


Figure 13: Verbal-numerical probability scale (adapted from Renooij 2001)

In their paper, Witteman & Renooij (2003) conclude that the double scale "leads to accurate assessments and it facilitates the process for the numerically less literate."

Indirect elicitation method: probability wheel

The respondent can be asked to estimate probabilities using the probability wheel. This is a wheel with a surface consisting of two parts with different colours, and it contains a pointer that can spin. The respondent is asked to fix the area of one of the parts, so that the probability of the pointer pointing at that area when the wheel is spun is estimated to be equal to the probability under consideration. As an alternative to the probability wheel, an example of an urn containing 1000 balls of two different colours may be used. The respondent is then asked how many of these balls should have a certain colour, for the chance of drawing a ball with that specific colour to be equal to the probability under consideration.

Indirect elicitation method: betting rate

In this method, a bet is proposed. The bet relates to the event under consideration. The question which is asked is how much money the respondent should get if the event turns out to be “favourable” (i.e. in accordance with the estimate), and how much money he should get, if the bet turns out to be “unfavourable”, under the condition that he would be indifferent which of the bets to take. The ratio between the money of the bet and the total amount of money is then set equal to the probability under consideration.

For further discussion on probability elicitation methods, see e.g. Spetzler & Staël Von Holstein, (1975); Cooke (1991); Morgan & Henrion (1990); Meyer & Booker (1991); Jenkinson (2005); Van der Gaag et al., (2002); Witteman & Renooij (2003); Renooij & Witteman (1999).

An important observation is made by Jenkinson (2005), who addresses the problem that the probability wheel, the betting rate, and probability scales have in common that the use of these methods is restricted to binary events (this goes for a large number of elicitation methods, see Jenkinson, 2005). Since state spaces are in general not binary, direct enquiry is the type of method that is used in this thesis, despite its disadvantages.

Although the probability elicitation process as treated in this chapter mainly focuses on the elicitation of probabilities with experts, it is more than just that. Before discussing the process itself, the participants in the process are discussed. The participants in the process are the decision maker (client), the expert(s), the statistician(s), and the facilitator (not always needed) (Jenkinson, 2005)

The decision maker is the one who needs the resultant of the elicitation process. The experts that are referred to are the substantive experts: people “who are involved in the application and have some knowledge about it” (Jenkinson, 2005). The statistician is a normative expert, i.e. he can give probabilistic training, validate results and provide feedback. Finally, the facilitator facilitates a session, in the case that more experts are involved.

Much literature can be found about the steps that have to be included in the elicitation procedure (see e.g. Meyer & Booker, 1991, Keeney & Von Winterfeldt, 1991, Renooij, 2001, Walls & Quigley, 2001, Spetzler & Staël Von Holstein, 1975). It is clear from literature that the probability elicitation process has to be performed in a structured way. Although authors define different steps for the elicitation process, in general, three main steps are identified. These three steps can be named as follows (Cooke & Goossens, 2000):

- Preparation for the elicitation
- Elicitation
- Post-elicitation

In the next subsections (6.1.2 – 6.1.4), different points of attention that play a role in these three steps are described.

6.1.2. Points of attention: preparation for elicitation

In this subsection, the methodological principles that play a role throughout the whole elicitation process are discussed first. After that, three important points for the preparation step in the elicitation process are discussed.

Methodological principles for elicitation

Next to a description of the elicitation process and possible heuristics and biases, Cooke (1991) also identifies a number of guidelines/conditions which apply to the use of expert opinion in science. He identifies five guidelines/conditions, which are combined by Cooke & Goossens ((2000), (2008)) into four methodological principles (Table 12), i.e.:

Table 12: Methodological principles for elicitation

1. *Scrutability/accountability*
   All data, including experts' names and assessments, and all processing tools are open to peer review and results must be reproducible by competent reviewers.
2. *Empirical Control*
   Quantitative expert assessments are subjected to empirical quality controls. This means that "there are theories and measurements relevant to the issue at hand, but the quantities of interest themselves cannot be or is not measured in practice" and "…necessarily, there will be relevant experiments which can in principle be used to enable empirical control" (Cooke & Goossens, 2000).
3. *Neutrality*
   The method for combining/evaluating expert opinions should encourage experts to state their true opinions, and must not result in bias.
4. *Fairness*
   Experts are not to be pre-judged (i.e. they have to be treated equally), prior to the results of the assessments.

These principles have to be adhered to throughout the elicitation process. The first step in the elicitation process, i.e. preparation, consists mainly of three activities, i.e.: selecting the experts, training them and structuring the questions.

Selecting and motivating experts

In the first activity, the size of the expert pool plays a role, i.e. the number of experts that are available for involvement in the elicitation process. Hora & Von Winterfeldt (1997) give six criteria on which to base whether or not an expert should be included in the elicitation process, i.e.: tangible evidence of expertise; reputation; availability and willingness to participate; understanding of the general problem area; impartiality; lack of an economic or personal stake in the potential findings. In practice, the number of experts is often low (Renooij 2001), and their availability is limited (Fenton, Neil & Caballero, 2007). Furthermore, they have to be involved in the problem structuring stage as well as in the instantiation stage of BBN construction. This means that selection of experts has to be done very carefully. Also a choice has to be made how many experts to involve, and how the experts' input should be combined, i.e.: whether they are asked to give an estimate as a group, or whether they are individually asked to give estimates that can be combined later.

Motivating the expert has as purpose to make sure that the expert is involved in the elicitation process. Motivation is attempted by describing the purpose and format of the elicitation exercise, as well as showing how the results are going to be used. At this stage of the elicitation process, motivational bias should be explicitly discussed with the expert, trying to prevent it ("conscious or subconscious adjustments in the subject's responses motivated by his perceived system of personal rewards for various purposes" (Spetzler & Staël Von Holstein, 1975)) from occurring in the process.

Training the expert

In this second activity of the first step of the elicitation process, the expert is made familiar with the concept of probability. Also, the expert has to become able to express his belief using the elicitation method that is chosen to be used in the process (Jenkinson, 2005). Both aspects of this activity are built into a training session. In this session, the expert is asked to assign a probability to a certain event. The most suitable elicitation method has to be sought here, to identify the method that best suits the preferences of the interviewee. A large number of elicitation methods are available in literature, based on different principles, as discussed in subsection 6.1.

If possible, probabilities for events as they are estimated by the experts can be compared to the real probabilities. According to Baecher, (1998) (referred to by Renooij (2001)), this feedback, will improve expert calibration. Although Kynn (2008) argues that expert training in terms of calibration is only useful when the questions are directly related to the subject matter, Kynn also states that the training itself is beneficial for improving the expert's understanding of probability. This implies that the training is in itself beneficial; the most important aspect of the training being the fact that it makes the experts comfortable with the concept of probability and with the elicitation method. In the specific case of BBNs, the training of the experts has to be aimed at making them familiar with the concept of conditional probability.

Structuring the questions

Renooij (2001) roughly divides this activity into three separate parts. The first part relates to the definition of the variables and their state space. The second part aims at identifying the problem context. The final part addresses the elicitation format to be used.

As stated in section 3.1, the definition of the variables and their state space is a part of the problem structuring phase of the BBN construction process. In the context of probability elicitation, there are a number of important aspects of the variable definition and the definition of their state-space (Spetzler & Staël Von Holstein, 1975; Kjaerulff & Madsen, 2008), i.e.:

- The state space should be unambiguous and exhaustive.
- The states of the state space should be mutually exclusive.
- The variable definition should pass the "clairvoyance test", i.e.: variables have to be well-defined without needing further clarification (Spetzler & Staël Von Holstein, 1975).
- The variable definition should be meaningful to the subject.

Because of the importance of the definitions of the variables and their state space, documentation containing these definitions should be kept at hand throughout the elicitation process. The context of the problem situation for which the elicitation takes place, also has to be determined. This has to be done, so that the estimates that are made are made under equal circumstances for all respondents. In this stage, it is also important to make sure that the respondents agree with the description of the context, and with the corresponding definitions. The last part of the structuring step addresses the elicitation format, regarding the questions as well as the answers) to be used. Renooij (2001) as well as Cooke (1991) states that both questions and answers should have an attractive format (preferably graphical (Renooij, 2001)). The preparation of questions and answering format are important for the quality of the elicitation outcomes. Therefore, the questions have to be considered carefully, and considerable time has to be spent on formulating the questions (Van der Gaag et al., 2002).

6.1.3. Points of attention: during elicitation

There are two important points of attention that have to be taken into account during the elicitation step, i.e.:

1. The heuristics and biases that play a role during numerical probability elicitation.
2. The way in which the elicitation should take place and be documented.

Heuristics and biases

Many different types of heuristics and biases exist. Heuristics are rules of thumb (Tversky & Kahneman, 1974), which are often introduced by people in the process of estimating a certain value, to simplify the task (Renooij, 2001). Bias reflects systematic error (Chapman & Ward, 2003), a systematic tendency to take factors into account that are irrelevant to the task at hand, or to ignore relevant facts (Renooij, 2001). Bias is a topic that is widely discussed in literature (see e.g. Renooij, 2001; Spetzler & Staël Von Holstein, 1975; Walls & Quigley, 2001), and many different types of bias have been identified. In general, two generic types of bias can be distinguished (Meyer & Booker, 1991; Spetzler & Staël Von Holstein, 1975), i.e.: motivational bias (which has been defined earlier in this subsection as "conscious or subconscious adjustments in the subject's responses motivated by his perceived system of personal rewards for various purposes" (Spetzler & Staël Von Holstein, 1975)), and cognitive bias ("conscious or subconscious adjustments in the subject's responses that are systematically introduced by the way the subject intellectually processes his perceptions"; Spetzler & Staël Von Holstein (1975)). The latter type is a result of the heuristics that play a role (Tversky & Kahneman, 1974). Both types of (views on) bias are discussed in the book of Meyer & Booker (1991). In their book, they discuss not only the two different types of bias, but they also introduce different sources of different forms of bias, indicating points of attention when eliciting probabilities. These points of attention differ depending on the elicitation process that is used and the environmental conditions of the elicitation. The different forms of bias that are presented in this thesis and the way in which they are addressed will be presented during the discussion on the application of the elicitation process in chapter 6.

Eliciting and documenting the expert judgments

A second point of attention during this (second) step of the elicitation process is the way in which numerical probability elicitation takes place and is documented. In this step, the elicitor and the respondent interact. Throughout the elicitation process, the elicitor has a number of tasks (Renooij, 2001), i.e.:

- Clarify problems that the respondent has with definitions and/or interpretations of questions.
- Record additional useful information stated by the respondent (next to the estimates). This can relate to the reasoning process of the experts and also to information regarding the network structure.
- Be aware of possible bias occurring, and intervene if necessary and possible.
- Keep track of time: 30-90 minutes is an adequate time frame for interviews (Spetzler and Staël Von Holstein 1975).

The elicitor has the important task to guide the elicitation process, at the same time making sure that he does not influence the respondent.

Part of the elicitation itself is also direct feedback of the results during the elicitation session. This can be done e.g. through graphical feedback (Kynn, 2008), enabling instant evaluation of the results.

6.1.4. Points of attention: post-elicitation

Finally, after the elicitation, the use and verification of the results need attention.

Combining experts' assessments

In general, it is preferable to include more experts, since this leads to better results (Clemen & Winkler, 1999, Winkler & Clemen, 2004, Jenkinson, 2005). In literature, different ways for combining assessments can be found. The combination of experts' assessments has to be taken into account already before the probability elicitation itself. This is because combining probabilities can be done using either a mathematical (after elicitation) or a behavioural approach (during elicitation) (Clemen & Winkler, 1999). Using a mathematical approach, experts are interviewed individually, and the results are aggregated using processes or analytical models to come to a single 'combined' probability. The behavioural approach aggregates individual probabilities in the elicitation stage, i.e. only one probability is elicited through a group elicitation process.

Mathematical aggregation (averaging forecasts) has been shown to lead to more accurate forecasts (Winkler & Clemen, 2004) than individual forecasts. In mathematical aggregation, there is a choice either to aggregate estimates based on equally based weighting, or to aggregate estimates using performance based weighting. In the latter case, the estimates of the experts that have shown to be able to generate more accurate and precise estimates are more heavily weighted. Accuracy here is related to calibration: “the statistical likelihood that a set of experimental results correspond, in a statistical sense, with the experts’ assessments”, Cooke & Goossens, (2000)); precision is related to information: “the degree to which an experts’ distribution is concentrated relative to some user-selected background measure” Cooke & Goossens, (2000).

In order to calculate a weighted average, seed variables are needed, being variables which value is already known to the elicitor, so that they can be used to determine the accuracy of the expert. These seed variables should have bearing on the field of expertise in which the questions are asked, since ‘good calibration’ means ‘good expertise’ (Cooke & Goossens, 2000). Since no seed variables are available in the case study that is described in this thesis, because the model is based on the views of experts only, it is not possible to use a weighted average based on accuracy of experts (which should be the most important criterion (Cooke & Goossens, 2008). Therefore, equal weights are used.

Evaluating the results

In general, three aspects of the elicited probabilities play a role in the verification process (which is carried out by the elicitor), and have to be taken into account (Renooij 2001), i.e.:

- Verification of the calibration of the expert.
- Coherence of the elicited probabilities.
- Reliability of the results (self-consistency of experts).

The verification of the calibration of the experts can be done by checking the elicited values against real data (if available). As already mentioned, in the context of reliability prediction early in the PDP, there are no data available that can be used to calibrate the experts.

Coherence of the elicited probabilities can be checked by verifying that the summation of the probabilities that represent the state-space of an event equals to 1.

Reliability of the results can be checked by testing whether the respondents are self-consistent (Kynn, 2008), i.e. whether the results of the elicitation process are consistent over time, when the elicitation is repeated.

6.2. Structure of the Probability Elicitation Process

In section 3.2, the instantiation process in BBN construction has been discussed, and a large number of points of attention were identified. In this section, all points will be addressed, discussing choices and their inherent consequences, as well as the extent to which the problems and issues regarding probability elicitation have been addressed. This section will have a descriptive nature, in which the probability elicitation process will be described and accounted for.

One of the most important determinants for the way in which the elicitation process is filled in is the context in which the elicitation takes place. Therefore, it is very important to first identify the conditions and the initial circumstances under which the probability elicitation process takes place. Two important elements have to be taken into account in this situation:

1. Probability elicitation is only one of the two steps of the BBN building phase (see Figure 9).
2. Probability elicitation has to take place in a company.

Since both the problem structuring stage and the instantiation stage are fully dependent on experts in the company (as it takes place in a company), their availability is very important. The fact that their availability is limited (both in number and in time, see section 6.1.2; Renooij, 2001; Fenton et al., 2006) is an important point of attention that has to be taken into account.

The fact that the probability elicitation process follows the intention of the application of the adapted GT process is taken into account as well, i.e.: also in the probability elicitation process, the experts are approached individually, rather than as a group.

6.2.1. Elicitation Method and Methodological principles

The dependency on experts was an important factor in the selection of the probability elicitation method. The choice was made for direct enquiry as elicitation method because of two reasons:

1. Indirect elicitation methods were inappropriate because of the limited availability of the experts: using indirect elicitation methods, the elicitation per probability can take up to 30 minutes (Druzdzel & Van der Gaag, 2000). Therefore, a direct elicitation method had to be chosen.
2. The probability scale is only usable for binary events (Jenkinson, 2005). Since the state space of most variables consisted of three states, the probability scale method was not suitable for use in most situations.

This leaves only the choice for using direct enquiry for the probabilities, in combination with text to describe the conditional probabilities that were asked.

In order to elicit CPTs directly, first, all CPTs that make up the BBN had to be defined. The CPT for a child node is built up as follows (see Table 22, appendix A, page 165):

- The right-most columns contain all the different possible states of the child node
- The left-most columns contain all the (direct) parent nodes of the child node
- The rows represent all possible combinations of states of the parent nodes

As a result of the size of the identified networks, the number of probabilities that had to be elicited from each expert was between 189 probabilities (C&D stage) and 287 probabilities (D&D stage). Because most of the probabilities were part of CPTs, and because the reasoning process for all probabilities in the CPTs was similar, experts were asked to evaluate probability elicitation through direct enquiry of single individual probabilities, and by enquiry of whole CPTs. Although the elicitation of a full CPT was considered cumbersome (especially in the case where the CPT consisted of 81 probabilities), the elicitation of full CPTs was considered by the experts to be favourable over the elicitation of individual probabilities. The most important reason for this was that the experts could adhere to the reasoning process that was used to determine the probabilities throughout the filling in of the CPT. According to the experts, the quality of the estimation of the CPTs was improved by the fact that they were forced to take all factors into account that played a role in the CPTs, resulting in a more consistent reasoning process (Visée, 2009). The gain in both time and (claimed) quality, as well as the clear preference of the experts for this method of probability elicitation, led to the choice for using direct enquiry of probabilities and CPTs for probability elicitation. A negative consequence of using direct enquiry of probabilities for CPTs is that it leads to a number of heuristics and biases. These have been discussed in more detail in subsection 6.1.3; the way in which they are dealt with is discussed in subsection 6.2.3.

The manner, in which the methodological principles are employed, is also influenced by the context in which the probability elicitation has taken place. The way in which the four principles (see Table 12) are adhered to, is discussed underneath:

1. *Scrutability/accountability*
   All data, as well as the experts' details and assessments are available. The tools that were used for the probability elicitation exercise are open to peer review as well.
2. *Empirical control*
   Empirical control regarding expert judgment in the context of reliability management from a systems perspective is not possible, because the constructed model is mainly based on the views of the experts (the experts' mental models), rather than on physical factors. Although empirical experiments may be performed for empirical control of the estimated probabilities, the model represents the beliefs of the experts, which may not be equal to the real probabilities, and may therefore not be verified empirically.
3. *Neutrality*
   Because the probability elicitation is based on the adapted GT approach, it has an individualistic nature. This limits the influence from outside. Also the well-structured elicitation protocol that is used benefits the neutrality of the probability elicitation.
4. *Fairness*
   Finally, the fairness principal, concerned with the judgment of experts prior to assessment results, is adhered to. Both during the process of GT application and during the process of probability elicitation, no difference is made between experts. This can be found in the criterion that is set for including/excluding nodes in/from the BBN: this is based solely on the number of people that mention a certain node, and the times that the node is mentioned. The same counts for the probabilities that are elicited: neither before nor after the probability elicitation process, weights are being attributed to experts' assessments.

As a conclusion, it has to be said that the methodological principles of probability elicitation have been followed to a certain extent, but that it is not possible to fully adhere to them.

6.2.2. Preparation for elicitation

The nature of the GT approach determines to a large extent the individualistic nature of the approach towards instantiation of the BBN. Furthermore, all experts that were involved in probability elicitation also were involved in the GT process, making sure that the experts were aware of the problem area, as well as of the aim of the research, and that they were available and willing to participate. Furthermore, their earlier involvement in the research made a discussion on the topic of reliability, as well as addressing motivational bias (through openly discussing it) easier.

Because of the limited availability of the experts, the training of the experts in the area of probability and specifically probability elicitation was included in the elicitation sessions (Visée, 2009). In this training, first the concepts of probability and conditional probability were dealt with. Then, an example was given of a CPT for a hypothetical scenario. After this, the expert was asked to determine his/her own CPT for this hypothetical scenario in order to make the expert comfortable with the concepts of CPTs and probability elicitation. Because of the lack of data, no seed variables could be used in this training to calibrate the experts.

In the problem structuring phase it was made sure that the variables were well-defined and meaningful to the subject, and that their state-spaces were unambiguous and exhaustive, with mutually exclusive states (see section 5.2). The definitions of the variables as well as their state-spaces were documented and made available (hard copy) for the experts. The format of the questions was a blank table (also hard copy), in which probabilities had to be filled in. Although this elicitation format introduces a number of heuristics and biases (see subsection 6.2.3) it reduces the elicitation time and imposes consistency on the reasoning process.

6.2.3. Elicitation

Throughout the elicitation process, heuristics and biases played an important role. Although they were addressed as much as possible, they still could not be fully discarded. Especially the use of the blank tables for probability elicitation (as mentioned earlier) gave rise to possible biases. In this subsection, a number of causes of cognitive bias will be discussed that are typically addressed by this method of probability elicitation, i.e.:

- Adjustment and anchoring: estimates are based on the most readily available information
- Unstated assumptions
- Availability: specific information is recalled by the respondent and influences the estimate

An important cognitive bias that was introduced through the use of CPTs was anchoring. Although the table presented an aid for maintaining consistency in the reasoning process of the experts, at the same time, it presents anchors. This is illustrated in the following example: if a node has three parents, and all nodes (child and parents) have three states, the CPT belonging to the child node contains 27 distinct scenarios, represented by 81 probabilities. If the state-space of the variables ranges from ‘good’ to ‘bad’, then the best scenario possible (when all parents are ‘good’) will have the most positive outcome for the child node, whereas the worst scenario possible (when all parents are ‘bad’) will have the most negative outcome for the child node. Finally, if all parents are in the state ‘average’, then the child node also will have an average outcome. In this way, these states then provide the ‘anchors’ around which the rest of the CPT can be filled in.

Bias because of unstated assumptions cannot be overcome, since unstated assumptions are represented through the uncertainty that is described by the probabilistic nature of the model. Because not all possible factors that affect reliability are included in the model (all models have a limited number of nodes), not all assumptions are made explicit. To address this bias, the experts were asked to think aloud, and state their assumptions when filling in the CPTs.

Since the problem structuring stage of the BBN was performed just after the development of a real subsystem, the danger of availability bias was large. Even though the probability elicitation took place some time after the problem structuring, still there is a reasonable risk for ‘availability bias’. No explicit measures have been taken to overcome this bias.

During elicitation, the elicitor was present, so that the experts were able to ask questions regarding possible problems with definitions and/or interpretations of questions that might exist. Also, it was possible to keep track of time in this way, and possible interventions could take place if possible and needed.

Additional useful information regarding the reasoning process and network structure was not further investigated. This was because GT already represented the reasoning process of the experts, and possible contradicting statements of experts from different fields of expertise were already identified and resolved. The experts were asked to make the reasoning process explicit, so that the expert could be guided through the elicitation process, and a wrong reasoning process through misunderstandings could be prevented.

Finally, the time was monitored, so that the interviewing time could be limited to 30-90 minutes, which is an adequate time frame for interviews (Spetzler & Staël Von Holstein 1975). If it was estimated beforehand that more time would be needed, two elicitation sessions were planned, so that the focus of the expert could be kept at a high level.

6.2.4. Post-elicitation

Directly after elicitation, graphical feedback was used to evaluate the elicitation results (Visée, 2009). This was done by providing graphical feedback to the expert for a number of scenarios. The graphical feedback consisted of a bar chart, in which three discrete probability distributions for a child node were represented for the three different values of (one of) the parent node(s), keeping all other variables constant. The shapes of the probability distributions were shown to the expert, and the expert was asked whether they matched the expert’s opinion.

As already stated, because of the individualistic nature of GT, the probability elicitation process took place individually per expert, leading to individual assessments as starting point. As a result, the experts' judgments were combined mathematically. Because the BBN for which the probabilities were elicited was based on interviews with experts, it was a representation of the views of the experts and represented their mental model of the situation. For the mental models, there was no data available which could be used for the calibration of the experts, nor was it possible to later verify the model.

Internal coherence of the estimates was guaranteed by the use of CPTs for probability elicitation: the experts were aware that the sum of the probabilities in each row should sum up to one.

The reliability of the results of probability elicitation is by Kynn (2008) considered to be the extent to which probability estimates reflect a persons internal beliefs, relating to the calibration of the expert. Although, as already stated (see subsection 6.1.4), the calibration of the expert cannot be measured directly, an indication of the expert's consistency can be obtained. For this, a second elicitation round was conducted, in order to find to what extent these results confirmed the results of the first elicitation round. Although this second elicitation round could not be used to provide insights in the calibration of the experts, it was able to provide insights in how informative (related to precision in technical literature) the experts were.

In technical literature (Kolarik (1999)) it can be found that the consistency of the measurement (more specifically, the precision of the measurement) is dependent on the quality of the measurement tool and the person that measures the quantity (in terms of Kolarik (1999): instrument, respectively operator). Kolarik (1999) states that deviations in the precision of measurements has two sources: "(1) the failure of our gauge or instrument to exactly repeat itself and (2) the failure of an operator or machine to exactly reproduce the measurement technique or method". In the case of probability elicitation, the instrument is the elicitation protocol, whereas the operator is the elicitor. Analysis of the repeatability and reproducibility of the results are referred to as Gauge R&R studies. In this thesis, the same concept is applied to the extent possible in the context of probability elicitation.

The results of the second round of interviews were obtained at a later point in time in order to avoid experts to be able to recall their previous answers based on short term memory. Furthermore, the interviews were not always conducted by the same persons as a result of the elicitors not being available at the same points in time. Therefore, the results of the study (discussed in section 6.5) have to be treated with care. In order to treat the results as having bearing on repeatability of the elicitation process, the assumptions have to be made that:

- The experts are consistent over time (temporal consistency; they do not change their opinion over time).
- The elicitation process is independent of the elicitor.

In accordance to Zitrou (2006), a formal elicitation protocol was set up and used for the elicitation process. Zitrou also recommended frequent feedback as validation tool. Since the elicitation process in this case study took place in the presence of the elicitor(s), it was possible to give continuous, rather than frequent feedback. The continuous feedback was used to improve the consistency of the results. This is described in the next section.

6.3. Application of Probability Elicitation: Elicitation Protocol

For the elicitation, a protocol was set-up beforehand, and followed throughout the elicitation process. The elicitation protocol consisted of the following components:

- *Training:*
  The experts were trained for reasoning with probabilities. Also, they were acquainted with the subject of conditional probability. Finally, an example was given in order to practice the probability elicitation exercise.
- *Elicitation:*
  During the elicitation process, the expert was given a document, which contained a description of the model, the nodes and their state-space, as well as all CPTs (blank) that needed to be filled in. The expert used this document to estimate the probabilities in the network.
- *Post elicitation:*
  After elicitation, graphical feedback was given to the expert.

In this section, all components of the elicitation protocol are discussed. For performing the whole elicitation protocol (including training, elicitation and post-elicitation), 90 minutes were available per expert (which is, as indicated in Spetzler & Staël Von Holstein (1975), a sufficient time frame) . In most cases, this timeslot proved to be sufficient. In one case, the elicitation took longer, due to the fact that this expert provided much information on his adopted reasoning process.

6.3.1. Training

The training of the experts was done through the use of a short introduction on the topic of conditional probabilities, and a short assignment on the judgement of probabilities (Visée, 2009). The training was in the form of a presentation (see appendix F (Visée, 2009)), which lasted about 15 minutes. During this presentation, a simple BBN was shown, consisting of three nodes: two parent nodes (mode of transport and waking up late), and a child node (arrival at work). The parent nodes all had 2 states (public transport and car, respectively late and on time); the child node had 3 states (early, on time, late). Furthermore, the definitions of the different variable states were given, as well as the corresponding NPT, including fictive probabilities. By explaining the different cells in the NPT, different probabilities were explained, and it was explained what a possible reasoning could be. It was explicitly stated that the example model, as well as the real model, were incomplete, i.e.: not all possible influencing factors were included. Moreover, it was made clear that the absence of certain factors could be expressed through the probabilities in the NPTs.

After this introduction to the topic of conditional, subjective probability, the expert was assigned to construct a BBN of his own, again containing the child node ‘arrival at work’ with the states ‘early’, ‘on time’ and ‘late’. However, instead of using the pre-defined parent nodes, now the expert was asked to think of two other parent nodes and their states (making up an exhaustive set). After this, he was asked to fill in the table, and to give his reasoning behind the filling in of the table, in order to make him comfortable with the procedure.

6.3.2. Elicitation

After the training, the elicitation process was started. Because of the large number of probabilities that had to be elicited for each stage of the PDP, the decision was made to elicit the probabilities of the various stages separately. Furthermore, because of the large number of probabilities that had to be estimated for the M&I stage (287), it was decided to elicit the probabilities for this stage in two sessions, in order not to overburden the experts. The involvement of experts in the different elicitation sessions for the different stages of the PDP, was based on their availability, their willingness to cooperate, and the percentage of factors of the model for the specific stage of the PDP (be it the C&D, D&D or M&I stage) that the expert identified, i.e.: if the expert identified most factors in one stage, he was asked to estimate the probabilities related to that specific stage. In the probability elicitation process in total thirteen people were involved, five for the model of the C&D stage, three for the model of the D&D stage and five for the model of the M&I stage.

Finally, the probabilities had to be estimated for the overall model of the PDP, connecting all separate models to each other. Because the division of the PDP into the three mentioned stages was a result of the data analysis rather than an observation specifically mentioned by an expert, the way in which this overall model was constructed and the estimation of the related probabilities was done by the gate keeper, who was involved in the processes that played a role throughout the PDP, but had not been interviewed in the context of the adapted GT process.

Before the elicitation process started, the expert was given a number of documents (Visée, 2009):

- A page with the figure of the BBN for which they were asked to estimate the probabilities
- A page containing the definitions of the variables including a definition of the different states
- A descriptive, self-explanatory document, which led the expert through the probability elicitation process, and which included an introduction to the research

The document was self-explanatory, i.e.: the document was written so that the reader was guided through the document and that he was able to fill in the tables without any external help. For all of the root nodes, the expert was asked to estimate the probabilities of the node being in a certain state. For all CPTs, an example question was used to ask the expert to estimate the probabilities that represented the discrete probability distribution, conditioned on certain values for the parent nodes. After this, the expert was asked to estimate all probabilities of the CPT, in this way filling in the table as a whole. An example of this is shown in Box 1 below. If people found that they were not able to fill in the CPT, they were given the freedom to leave the CPT blank. If the sum of probabilities for the states of a certain event did not sum up to 1, the distribution among the states was normalized.

2.3 Attention towards reliability

Example question; Consider that reliability is only once a year on the agenda of the MT-meeting & Activities and deliverable are present, without evidence. How likely is it that the attention towards reliability can be considered …

High? …

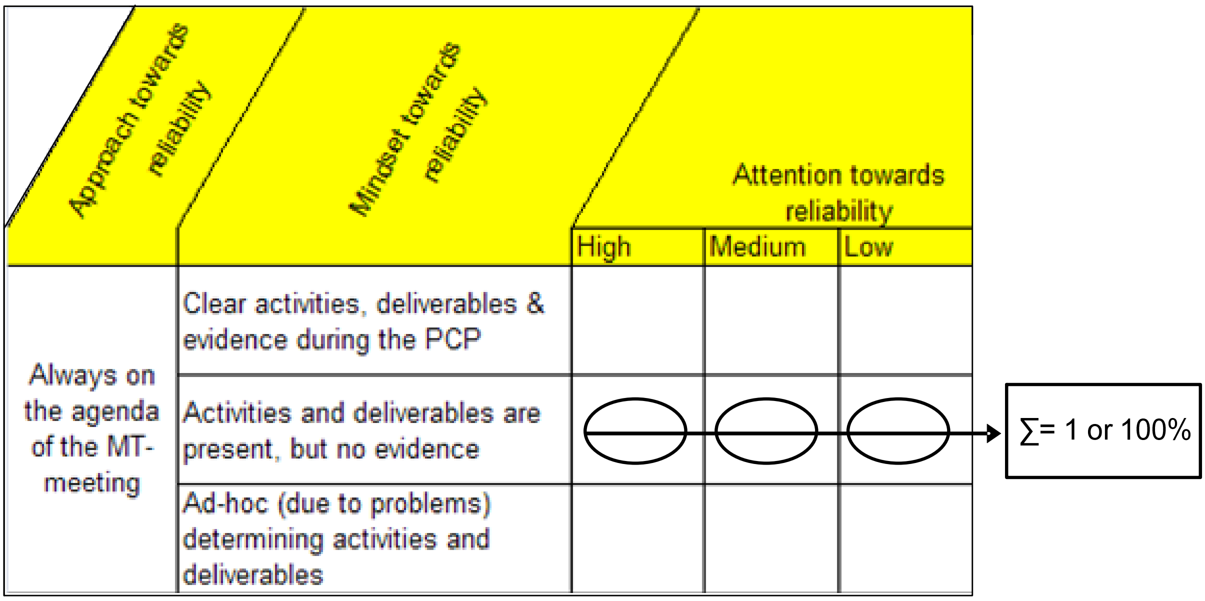
Medium? … ∑ = 1 or 100

Low? …

Use the question above as guide to fill in the table on the next page

How to fill in the answer?

- In case of a table, as shown below, the answers should horizontally sum up to 1 or 100%.



Approach towards reliability	Mindset towards reliability	Attention towards reliability High	 Medium	 Low
Always on the agenda of the MT-meeting	Clear activities, deliverables & evidence during the PCP			
	Activities and deliverables are present, but no evidence	○	○	○
	Ad-hoc (due to problems) determining activities and deliverables			

∑= 1 or 100%

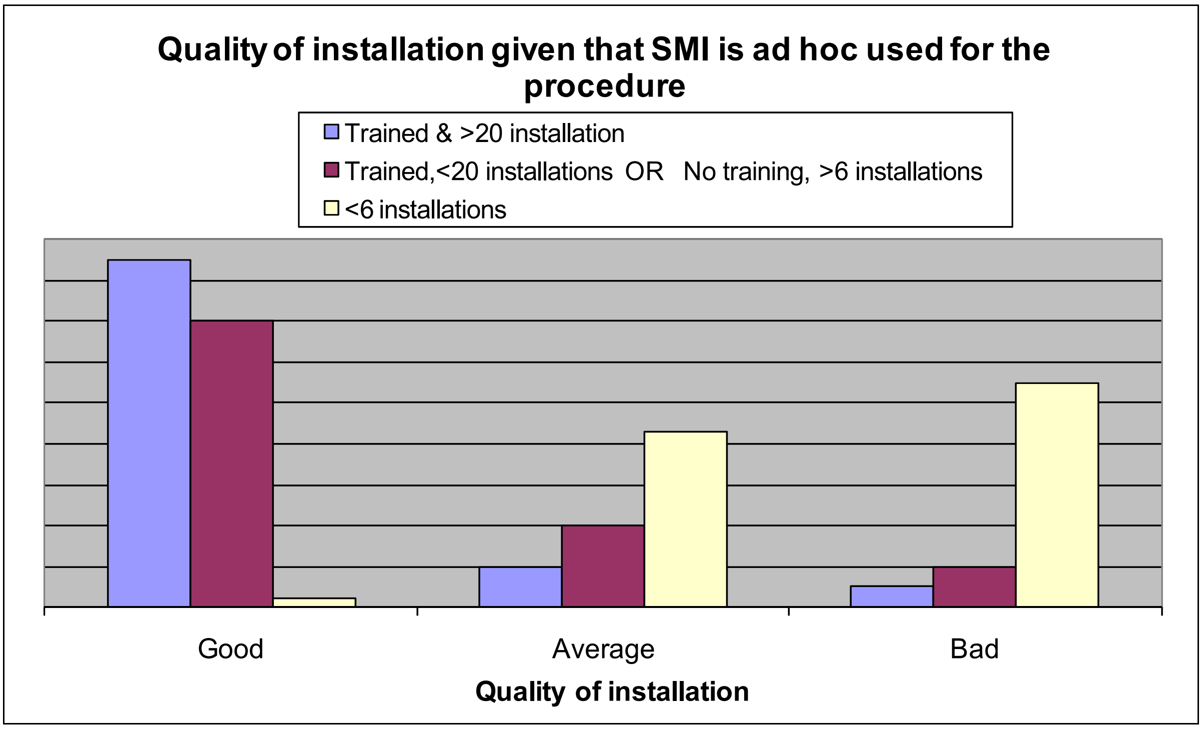
Box 1: Question preceding a table (Visée, 2009)

The elicitors were present at the process to answer possible questions that the experts might have, and to stress the importance of the activity for the research.

It is important to state that, although the influence of the elicitor on the expert (and the associated bias) was minimized by using a self-explanatory document, there were cases where the expert was unable to fill in the document, and asked the elicitor for help. In this case, the elicitor was strongly involved in the elicitation process. Unfortunately, possible bias that resulted from this approach could not be prevented.

6.3.3. Post-elicitation

After eliciting all tables, the elicitor provided graphical feedback for the experts, to verify the results. For this purpose, a bar chart was generated using Microsoft Office Excel©, depicting different probability distributions of the child variable for all different states of one of the parent variables. If the child had more than one parent, then the value (state) of the other parent(s) was defined. The way in which graphical feedback was given is depicted in Box 2.
Because of the large number of probabilities that were elicited, only a limited amount of situations were evaluated graphically. Furthermore, in the cases in which there was no time available for feedback, this component of the elicitation protocol had to be left out.



Box 2: Example of graphical feedback after the elicitation took place (Visée, 2009)

As discussed in subsection 6.1.4, the expert assessments are combined through mathematical aggregation. Since there is no way in which weights can be determined for the different individual experts, the mathematical aggregation is based on equal weights, i.e.: the average result of all obtained results is used as model input. The probabilities for the NPTs and CPTs that could not be estimated by experts were not taken into account in the mathematical combination process, i.e.: the weight of these experts was set to 0 for these NPTs and CPTs.

6.4. Reflection on the Elicitation Process: Continuation of the Case Study

For all models of the three stages of the PDP, as well as for the overall model, the related probabilities were elicited (both the NPTs and the CPTs) by a selected number of experts. All experts were asked for eliciting the probabilities for the model of one stage of the PDP only.

The selection of the experts per model was based on the adapted GT approach. For all nodes in the separate models for the specific stages of the PDP, the interviewees in whose interview reports the node was identified were noted. In this way it could be identified for all models which experts had identified most or all of the nodes in the specific BBN. Experts that identified the most nodes per BBN and that were available for probability elicitation purposes, were then involved in the probability elicitation for the specific model for which they were selected. Although not all experts were able to answer all questions for the BBNs, the experts were mostly able to estimate the probabilities corresponding to the NPTs and CPTs of the BBNs.

Diversity in the areas of expertise of the experts should be leading in the selection process for identifying which and how many experts to involve in the elicitation process of each separate BBN, i.e.: the involved experts should preferably provide a full spectrum of all disciplines that are involved in the stage of the PDP in question. Unfortunately, this was not always possible in the elicitation process. In the case study, identification of the experts that were to be involved in the probability elicitation process was mainly based on availability. This counted both for the identification of the specific experts to include in the elicitation process as well as for the number of experts to include in the process.

The probability elicitation process was done based on a probability elicitation protocol, as described in section 6.3, consisting of a training, an elicitation session, and a post-elicitation session.

In total, thirteen experts contributed to the probability elicitation: five experts involved in the C&D stage, three experts involved in the D&D stage, and five experts involved in the M&I stage. For the experts involved in the C&D and M&I stage, 90 minutes were reserved for an elicitation session. It was possible to elicit the probabilities for all NPTs and CPTs of the networks (10 nodes containing 189 probabilities, respectively 8 nodes containing 207 probabilities) in these sessions. For the D&D stage however, 289 probabilities (distributed over 14 nodes) had to be elicited. Therefore, the network was split up into two pieces, which were elicited separately. The two elicitation sessions that were held to estimate the probabilities both lasted about 60 minutes. The total time spent on elicitation is presented in Table 13.

Table 13: Experts involved and time spent for probability elicitation for the different stages of the PDP

	C&D stage	**D&D stage**	**M&I stage**
Number of people involved	5	3	5
Time spent (minutes, in total)	450	360	450

6.5. Repeatability of the Probability Elicitation Process

Repeatability in the context of probability elicitation is difficult to measure. This is because it relates to the measurement of people's opinions, for which there is no objective mechanism.

Furthermore, there are numerous issues that may influence the elicitation process and its outcomes. Such issues are e.g. heuristics and biases, as discussed earlier in this chapter.

In order to create an indication for the repeatability of the elicitation process, regression analysis is used, comparing two sets of elicited probabilities. These sets of probabilities were elicited by the same experts at two different moments in time. This is further discussed in subsection 6.5.1, after which the results of the regression analysis are discussed (see subsection 6.5.2).

Validation of the results is different from repeatability of the elicitation process. Validation of the results is done using the outputs of the model. As indicated in subsection 3.2.2, model outputs are generated through Bayesian inference, and validation using Bayesian inference will be discussed in the next chapter (section 7.6).

6.5.1. Evaluating the repeatability of the process: regression analysis

As an indicator for the repeatability of the process, the correlation coefficient is used, which can be obtained through regression analysis. The correlation coefficient is a measure for the extent to which two sets of values match. In the case of the probability estimation, the conditional probabilities that were estimated the first time are compared with the conditional probabilities as they were estimated the second time.

A positive correlation coefficient indicates agreement between the second and the first measurement (1 indicating perfect agreement), whereas a negative correlation coefficient indicates disagreement with the first measurement (-1 indicating full disagreement).

The measurement is done at different moments in time, between which the expert is subject to external influences. The expert's probability judgment may be altered over time, and this alteration may be based on a clear reason. In such a case, the correlation coefficient would indicate imperfect agreement, or even disagreement. This would not be the result of a poor elicitation process, but of a change in the expert's opinion.

Furthermore, the elicitation process is done by two different people (because the two elicitors were not available at both points in time). Although the elicitors are aware of heuristics and bias, and aim to avoid these, they cannot be ruled out beforehand. Heuristics and biases may also influence the results of the regression analysis.

As a result, regression analysis cannot be taken as a true measure for the quality of the elicitation process, and the outcomes of regression analysis have to be treated with care. As an example, the interpretation of the correlation coefficient may be that low correlation indicates a bad measurement instrument (low repeatability), but it may also be, that low correlation indicates a change in the opinion of the expert.

The use of the correlation coefficient is most useful to indicate a need for investigating the elicitation process and its results, in the case if many correlation coefficients are strongly negative.

Because the relations are more important for the model than the root nodes (since they reflect the relations between variables, and are assumed to be less subject to change), the CPTs of the child nodes (depicting the relations) are looked at, whereas the NPTs of the root nodes are left aside in the regression analysis.

6.5.2. Results of the regression analysis

In order to evaluate the probability elicitation process, six experts performed the probability elicitation process twice, estimating the same probabilities and CPTs. The second round of probability elicitation took place about two months later than the first round, and (in most cases) involved a different elicitor. The study included four people that were involved in probability elicitation for the BBN representing the C&D stage (189 probabilities) and two people involved in probability elicitation for the BBN representing the M&I stage (207 probabilities).

In the second round of probability elicitation, as was the case in the first round of probability elicitation, the people were given the freedom to leave the CPT blank if they found that they were not able to fill in the CPT. In the cases where either the first or the second round of elicitation did not give any results, no analysis was possible (which was the case for two CPTs). If the sum of probabilities for the states of a certain event did not sum up to 1, the distribution among the states was normalized.

For the first round of probability elicitation, all interviews were limited to 90 minutes, whereas in the second round of probability elicitation, the interviewing time was limited to 60 minutes (since experts already had some experience in probability elicitation). Again, one person used more time than this, but throughout this time provided much information on the expert's adopted reasoning process. Because of the limited availability (with respect to time) of experts during the second session, the post-elicitation component of the elicitation process could not be realized in the second round of interviews.

Analysis of the results of the probability elicitation process by means of the correlation coefficient was done by comparing all individual probabilities that were elicited in the CPTs: for each table, two vectors were created containing all probabilities as they were elicited during the first round, respectively the second round of probability elicitation. The results are presented in Table 14 (C&D stage) and Table 15 (M&I stage).

Table 14:Correlation coefficients for the repeated elicitation of the corresponding CPTs

	Node				
Expert	*Attention towards reliability*	*Configuration variety*	*System diversity*	*Setting specifications*	*Reliability effect C&D phase*
A	**- 0.0133**	0.6124	XXXX	0.8841	0.5500
B	0.7875	**- 0.3592**	0.6586	0.7028	0.7152
C	0.5575	0.2794	0.6006	0.5987	XXXX
D	**- 0.1124**	0.9239	0.8290	0.5628	0.6733

Table 15: Correlation coefficients for the repeated elicitation of the corresponding CPTs

	Node		
Expert	*Quality of installation*	*Quality of tests*	*Reliability effect M&I phase*
E	0.5117	0.6488	0.7689
F	0.7468	0.5893	0.5834

The results of most experts agreed in both rounds of elicitation (resulting in a positive correlation coefficient in most cells of Table 14 and Table 15).

Validation of the results of the elicitation process will take place through analytical and synoptic validation of the resulting BBN, which will be presented in section 7.6 in the next chapter.

6.6. Summary and Conclusions

As identified in section 3.2, there is one main challenge related to the instantiation phase of the BBN construction process, which is the acquisition of the probabilities for the BBN.

The challenge is addressed by using a direct probability elicitation technique (direct enquiry). In this way, the probabilities composing the network could be obtained from experts, making use of the only source of information available. Furthermore, by using direct rather than indirect elicitation, the probability elicitation process is fast (compared to other methods, see e.g. Druzdzel & Van der Gaag, 2000), as can be concluded from Table 13. This lightens the burden for the experts in terms of invested time.

Regarding the reliability/repeatability of the probability elicitation process, it is found that experts have not been able to repeat their estimates perfectly (this is measured using the correlation coefficient). Since this may be the consequence of many factors, no direct conclusions can be drawn from this result. However, it is an indication that the results of the probability elicitation process have to be treated with care, and it stresses the fact that results should not be used offhand, but have to be discussed with experts first, in order to validate the findings. This is discussed in the next chapter (section 7.6). Before going to the next chapter – the application of the BBN – first the full algorithm is presented that was applied in order to construct a BBN.

6.7. Algorithm for Step by Step BBN Construction

The developed approach for BBN construction (in the context of reliability management) is presented in this section. The approach consists of two main stages:

- Problem structuring
- Instantiation

The approach based on GT that is developed for the problem structuring stage (including data gathering) is discussed in the first 5 steps underneath, the approach that is developed for instantiation is presented in steps 7 and 8. For a more in-depth discussion on the preparation and application of the problem structuring stage, see sections 4.4, 5.1 and 5.2. A more in-depth discussion on the preparation and application of the instantiation stage can be found in sections 6.2 and 6.3.

Gather information regarding the way in which the topic under discussion is influenced by conducting interviews (see sections 4.4 and 5.1)

The interviews provide the information needed to identify the problem and its boundaries. Data has to be gathered based on diversity of information sources (e.g. experts from different disciplines in the field).

Identify the factors (i.e. nodes) that influence the topic, by analyzing and coding the interviews (see subsection 5.2.1)

To identify the nodes to include in the network, open coding has to take place. This results in the identification of concepts and categories. Both concepts and categories can represent nodes.

Define the variables by identifying the different possible states (state-space) of the variables through coding and direct conversation with experts (see subsection 5.2.1)

In the open coding process, the different nodes were identified. The identification of the state-space may be done through enhancing the theoretical sensitivity (the second part of the open coding process). If the identification of the state-spaces of the variables is not possible solely based on enhancing theoretical sensitivity (second step of open coding) of the existing information, additional information may be gathered. In this case, state spaces of variables may be defined through direct conversation with experts.

Characterize the relationships between the different nodes using the idioms through analysis and coding of the interviews (see subsection 5.2.2)

During the second part of open coding and during axial/selective coding, relations between nodes in the network are identified. The relations are represented by directed arcs in the BBN, and can be typified by the idioms introduced by Neil et al. (2000). The nodes identified in step 2, together with the arcs, represent the network structure.

Control the number of conditional probabilities that has to be elicited using the definitional/synthesis idiom (see subsection 5.2.2)

The number of conditional probabilities that has to be elicited has to be controlled by limiting the number of input nodes for each node. This can be done using the definitional/synthesis idiom ((Neil et al., 2000). Note that the definitional/synthesis idiom is only usable for the purpose of controlling the number of probabilities to elicit, if the parent nodes that are synthesized are independent of other nodes (that are not synthesized).

Evaluate the BBN, possibly leading to a repetition of (a number of) the first 5 steps (see section 5.4)

After defining the model through the identification and definition of model boundaries, variables, and identifying the model structure, the model can be validated using a survey. Based on the outcomes of the survey, (part of) the problem structuring (steps 1-5), may be repeated to improve the model.

Identify and define the CPTs that define the relationships in the BBN (see subsection 6.2.1)

The relations of a BBN are represented as CPTs. Therefore, in order to define the relationships of the BBN (instantiate the BBN); the CPTs that make up the BBN have to be elicited. The CPT for a child node is built up as follows (see Box 1, page 106):

- The right-most columns contain all the different possible states of the child node.
- The left-most columns contain all the (direct) parent nodes of the child node.
- The rows represent all possible combinations of states of the parent nodes.

Fill in the CPTs, in order to define the relationships in the BBN (see section 6.3)

In order to fill in the CPTs, in the case when no data is available, the CPTs have to be elicited. This means that the experts have to estimate the values for all CPTs. This can be done using a probability elicitation process, consisting of:

- A training presentation (see e.g. appendix F)
- The elicitation process itself, including:
    - Definitions of the variables and their state-space (see e.g. Table 8, page 78 and Table 9, page 79)
    - Self-explanatory documentation for filling in the CPTs (see e.g. Box 1, page 106)
    - Visual feedback in the form of bar charts (see e.g. Box 2, page 107)

Evaluate the BBN, possibly leading to a repetition of (a number of) the first 7 steps (see section 7.6)

After defining the model through the identification and definition of model boundaries, variables, and relations, Through Bayesian inference, scenario analysis and main effects analysis can be performed, which can be used to validate the model. This will be discussed in the next chapter (section 7.6). Based on the validation of the model, (part of) the model building may be repeated to improve the model.

7. Using BBNs: Inference

In this chapter the use and limitations of the BBNs is examined that are constructed using the methodologies developed in earlier chapters. Since the use is limited to the early phases of the PDP, the concern lies not with accurate forecasts of reliability in long term use, but rather with reliability management throughout the three stages of the PDP (C&D, D&D, M&I), using indications of whether the product can be expected to meet the design specifications in terms of reliability. Moreover, the validation of the model is looked at.

In the first section, two different types of analysis that can be performed through Bayesian inference will be introduced, i.e.: scenario analysis and sensitivity analysis. Both types of analyses will be further elaborated on in section 7.2, respectively section 7.3. Different ways in which these types of analyses can be used for reliability prediction and decision support are discussed in section 7.4 and 7.5. In section 7.6, validation of the BBN model is discussed in detail. The chapter ends with a summary and conclusions of this chapter in section 7.7.

7.1. Analysis Using BBN Models

Before going into detail on the purposes for which BBNs can be used, first the BBNs that result from the case study are discussed, since these BBNs will be used as example to illustrate the use of BBNs.

The partial BBNs for the three different phases of the PDP (i.e.: C&D, D&D and M&I), as well as the composite BBN for the entire PDP, which connects all these partial BBNs, have been constructed, using the BBN tool AgenaRisk. This resulted in four models:

1. A partial model for the C&D stage (see Figure 11, chapter 5), using the average of the CPTs that were elicited in both rounds of probability elicitation (see subsection 6.5.2).
2. A partial model for the D&D stage (see Figure 28, appendix C), for which the CPTs were elicited once.
3. A partial model for the M&I stage (see Figure 29, appendix C), using the average of the CPTs that were elicited in both rounds of probability elicitation (see subsection 6.5.2).
4. A model for the overall PDP (see Figure 30, appendix C), for which the CPTs were elicited once.

For all the partial models of the separate PDP stages, as well as for the overall composite model of the entire PDP, the qualitative models that were based on the adapted GT approach, as presented in section 4.3 were used. The probabilistic relationships were used that were based on the CPTs elicited using the direct elicitation method of section 6.2. Because the root nodes in the different models represent input variables for the model, and because scenario analysis as well as sensitivity analysis use their current values, these values were initially defined as non-informative, uniformly distributed (being either ‘0.33 – 0.33 – 0.33’ for input variables with three states or ‘0.50 – 0.50’ for input variables with 2 states). In this way, because all different input variables had the same value, it was possible to make a fair comparison of the different models and input variables using scenario and sensitivity analysis, since it was made sure that the changes in the values of the input variables could be made equally large.

Generally, there are two ways in which Bayesian inference can be used for means of analysis. These are scenario analysis, and sensitivity analysis, which will be addressed in the next two sections.

7.2. Scenario Analysis

Scenario analysis may be used both for prognosis and for diagnosis, i.e.:

- *Prognosis:*
  Given certain conditions (values of the parent/root nodes), the outcome (value of the child/leaf node) may be calculated. This is done using ‘forward reasoning’ (see appendix A) As such; prognosis is the way in which BBNs can provide a prediction of the reliability.
- *Diagnosis:*
  Given a certain outcome (value of the child/leaf node), the most likely conditions (values of the parent/root nodes) can be calculated under which this outcome is obtained. This is done using ‘backward reasoning’ (see appendix A).

The BBN describing the C&D phase of the PDP (see section 5.2, Figure 11) is used here to illustrate scenario analysis. Three different scenarios are defined and illustrated in Figure 14 - Figure 16. These figures represent the resulting values of a part of the BBN for the C&D stage, i.e.: in these figures, the MPDs are presented in the form of a graph including probabilities. The part of the BBN that is presented consists of the three direct parent nodes for the node ‘reliability effect C&D stage’, and the latter node itself. The root nodes are defined, although they are not included in the presented model. The outputs as well as the figures for the different scenarios are generated using the BBN tool AgenaRisk (Agena Ltd., 2009).

Scenario 1

Scenario 1 represents a situation in which there is no information about the values of the input nodes. The probability distributions for the input variables (root nodes) are defined to be non-informative (i.e.: the root nodes are preset with a uniform distribution). Since there is no information in this scenario, this scenario merely represents the prognosis in the case in which inputs are unknown. This is presented in Figure 14, where it can be seen that the value of the node ‘Reliability effect C&D stage’ is ‘0.55 – 0.24 – 0.21’, as a result of the following values for ‘Attention towards reliability’, ‘System diversity’, respectively ‘Setting specifications’: ‘0.27 – 0.21 – 0.51’, ‘0.56 – 0.27 – 0.17’, ‘0.46 – 0.23 – 0.31’.

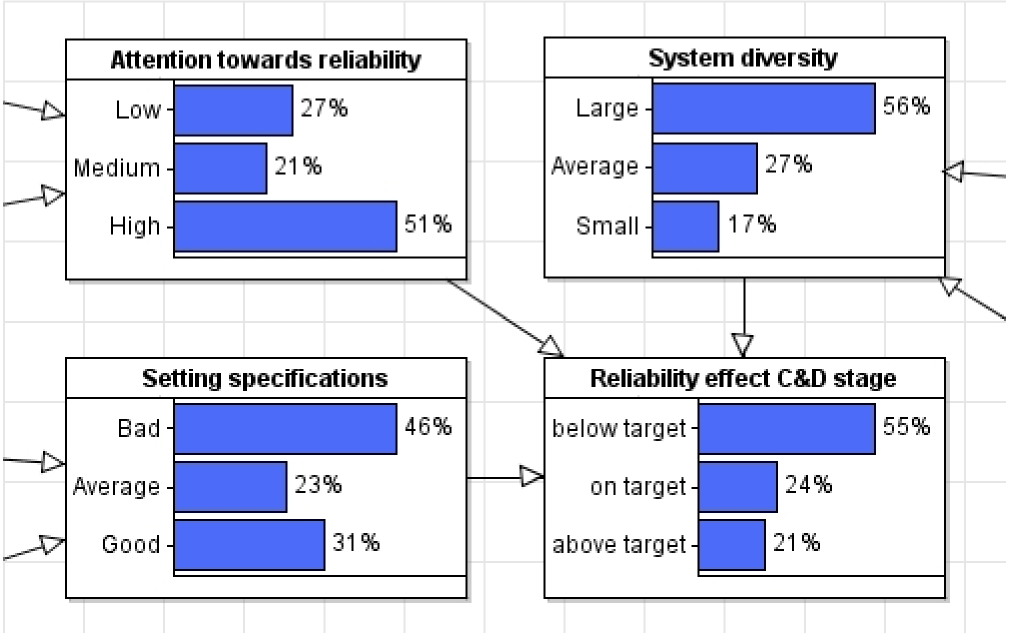


Figure 14: Scenario 1: prognosis (the preset values for the input variables being non-informative)

Scenario 2

Scenario 2 represents a situation in which there is information available regarding values of parent nodes (note: not necessarily root nodes). In Figure 15, it can be seen that the MPD of 'attention towards reliability' is preset, i.e.: evidence that the value (MPD) of the node 'attention towards reliability' is '0.00 – 0.00 – 1.00' is fed into the model, leading to a change in the value (MPD) of 'Reliability effect C&D stage' (i.e.: from '0.55 – 0.24 – 0.20' to '0.49 – 0.26 – 0.25'). This scenario illustrates prognosis as mentioned in the beginning of this section. The values for 'System diversity' and 'Setting specifications' do not change.

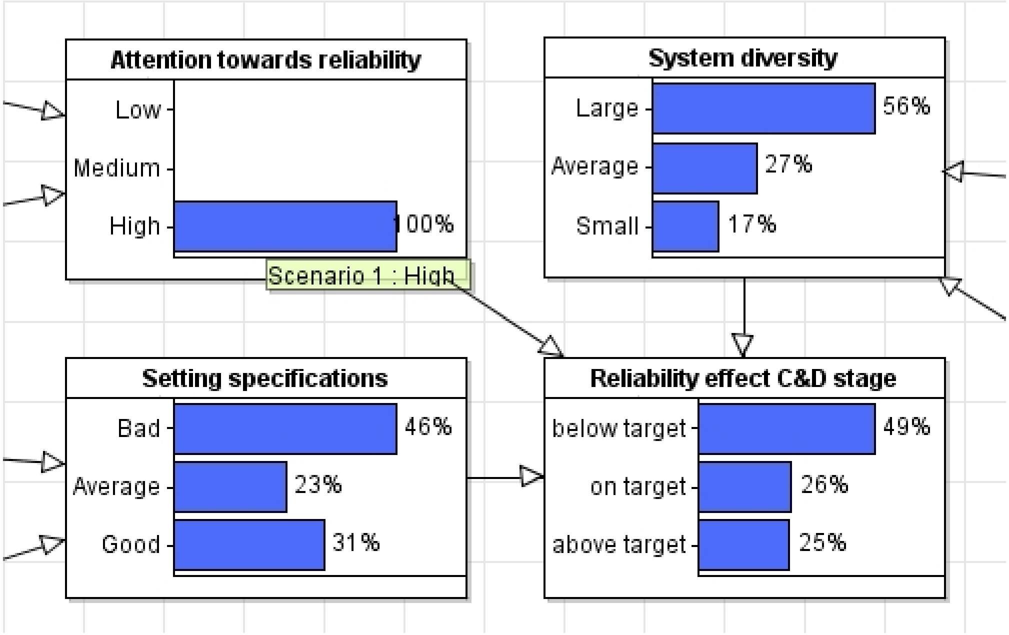


Figure 15: Scenario 2: prognosis ('Attention towards reliability' is 'high')

Scenario 3

Scenario 3 represents a situation, in which there is information available regarding the value of the leaf node, i.e.: a diagnosis scenario. Consider the evidence regarding the variable 'Reliability effect C&D stage' to be 'on target'. The value of 'Reliability effect C&D stage' is then preset as '0.00 – 1.00 – 0.00' and fed back into the model. The BBN can be used for diagnosis of this evidence, which is shown in Figure 16. It can be seen that the most likely state in which e.g. 'Attention towards reliability' is, is 'High' (with a likelihood of 0.54).

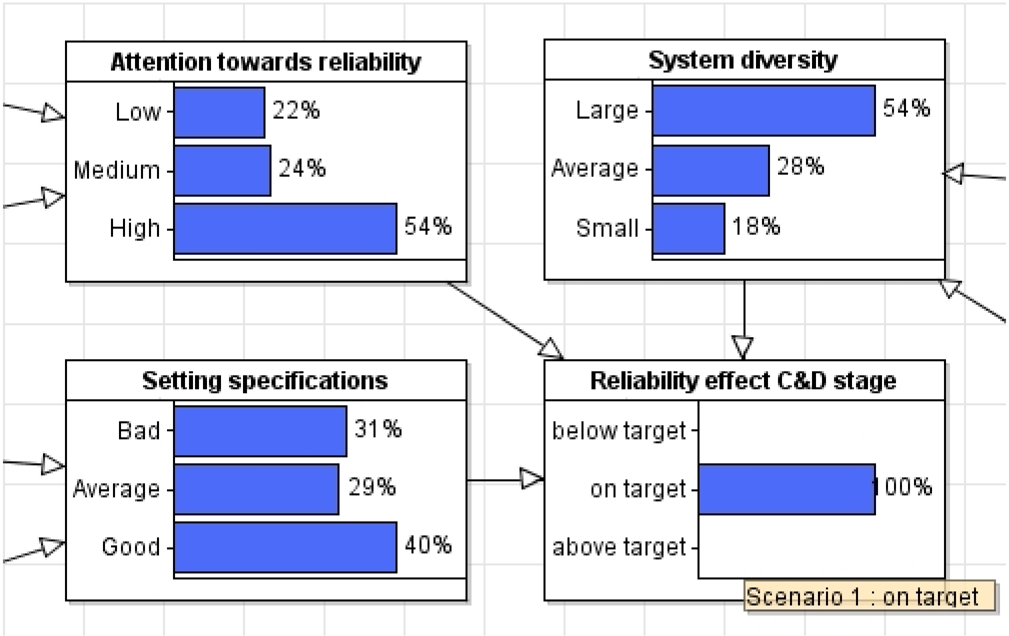


Figure 16: Scenario 3: diagnosis (with desirable outcome: reliability is 'on target')

7.3. Sensitivity Analysis

Next to scenario analysis, Bayesian inference can also provide sensitivity analysis. Basically, there are two main types of sensitivity analysis (Kjaerulff & Madsen, 2008):

- Evidence sensitivity analysis (SE analysis), where the focus lies on identifying how sensitive the belief update (model output; target variable) is to variations in the set of evidence (model input; evidence)
- Parameter sensitivity analysis (SP analysis), where the focus lies on how sensitive the results of a belief update (model output; target variable) are to variation in a parameter value in the model.

Whereas SE analysis clearly focuses on the way in which evidence propagates through the BBN following changes in the evidence, SP analysis focuses on changing the model itself. This means that, whereas SE analysis can be used in the context of Bayesian inference using a systems model, SP analysis is related to the BBN building phase and cannot be used. The adapted GT approach provides a well structured and methodical approach for identifying the variables to include in the model. The pre-defined process for the direct elicitation of the model CPTs provides a means for the systematic elicitation of the CPTs. In this way, the robustness of the model structure is ensured.

SP analysis simulates the effect of changes in the model itself: it shows what happens if the BBN structure is changed. Because the model reflects the perspectives of the experts on reality, the representation of these perspectives (i.e.: the model) cannot be changed by users of the model. Therefore, SP analysis is not suitable for use in the inference stage when the adapted GT approach has been used for variable identification and the CPTs are elicited directly, fixing the model structure.

As stated, in SE analysis, the effect of changing the values of the different input variables on the value of the leaf node is calculated. Since the states of the variables in the BBN are defined qualitatively, and the values of the variables are defined in terms of an MPD, SE analysis can take place both in terms of extreme changes in the values, and in terms of small changes in the values. These two ways of performing sensitivity analysis can be summarized as follows:

1. In the case of extreme changes, the effect of changing the value of one input variable from '1.00 – 0.00 – 0.00' to '0.00 – 0.00 – 1.00' on the model output is calculated. From here onwards, this will be referred to as Extreme Evidence Sensitivity analysis (EESA).
2. In the case of small changes, the effect of a limited change in the value of one input variable is calculated, e.g. from '0.25 – 0.35 – 0.40' to '0.20 – 0.30 – 0.50'. From here onwards, this will be referred to as Subtle Evidence Sensitivity analysis (SESA).

Variables may have individual (main) effects as well as interactional effects. In the following subsections, both types of effects will be discussed.

7.3.1. Main effects analysis: Tornado diagrams

The different input variables (root nodes) have an individual effect (main effect) on the output variable. In order to analyze these main effects, the different variables have to be varied systematically, so that the individual effect of the input variable on the resulting output variable becomes visible. One way of performing a main effects analysis is, is by systematically changing the input variables one by one. The output of the model can be calculated, setting the value of one input variable at the two extremes (i.e. highest and lowest, e.g. '1.00 – 0.00 – 0.00' and '0.00 – 0.00 – 1.00', respectively). Although all variables in the models in this thesis have states that are defined qualitatively, they are ordinal so that setting the values of the input variables at the extremes is possible. After changing the input variable back to its original value (which is in this case a non-informative, uniform distribution), the process can be repeated for all other input variables.

For every state of the output variable, the highest and lowest probability is identified, their difference being the range of influence that the input variable has on the specific state of the output variable. This can graphically be represented by a bar, which reaches from the lowest possible value to the highest possible value of the specific state of the output variable.

By ordering the variables according to their ranges of influence, a ranking of input variables can be determined for all possible states of the output variable. Combining the bars of the ranked input variables, a so-called 'tornado-diagram' is created. Such a tornado-diagram has been created using the software tool AgenaRisk (Agena Ltd., 2009), as can be seen in Figure 17. This tornado diagram shows the results of the EESA – analyzing the effect of the input variables of the C&D stage (shown in Figure 11) – which has been performed for the state the state 'Intrinsic (maximum) reliability is above target' of the node 'Reliability effect C&D stage'. In Figure 17, the vertical line in the middle of the figure represents the current value of the probability 'Intrinsic (maximum) reliability is above target'. The horizontal bars represent the range over which the different input variables may influence this probability (increase or decrease it). This range is defined by the maximal positive effect of the input variable on the probability and the maximal negative effect of the input variable on the probability.

In Figure 17, it can be seen that currently, the probability of the reliability being below target, is 0.55 (indicated by the vertical line in the figure). It can also be seen that, if the variable 'requirements' is pre-set to its most positive value (ceteris paribus), then the probability of reliability being below target would decrease to 0.45. If the variable 'requirements' is pre-set to its most negative value (ceteris paribus), then the probability of reliability being below target would increase to 0.65.

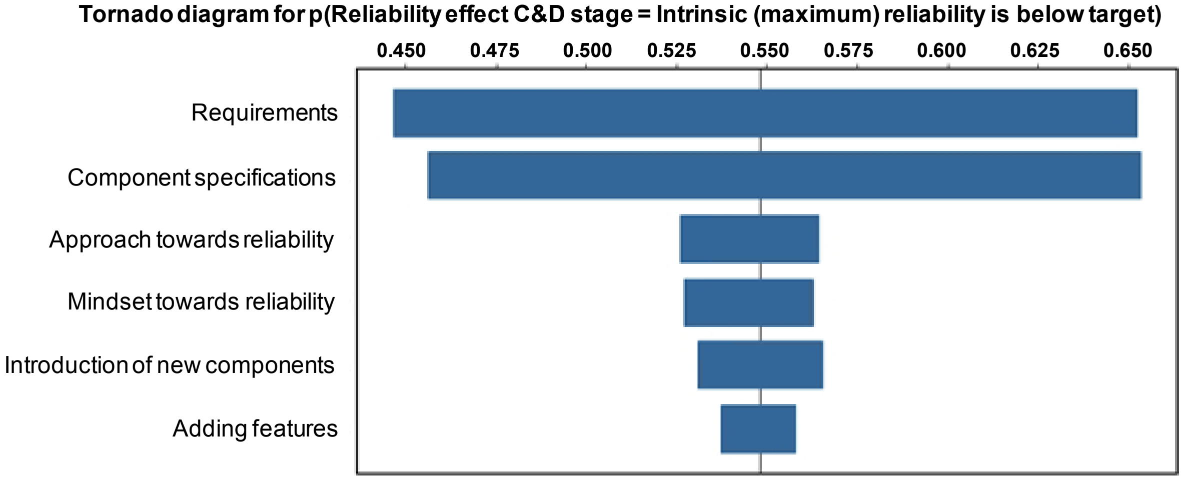


Figure 17: Example of a tornado diagram, constructed using the software tool AgenaRisk (Agena, ltd., 2009). The diagram depicts EESA, performed on the BBN as presented in Figure 11

It is important to note that this type of main effects analysis takes the existing situation as the point of departure. Because evidence for all input variables is propagated through the network, it is important in this case that the input variables are changed back to their original values, before propagating new evidence through the network, since comparison would otherwise be unfair.

7.3.2. Main and interaction effects analysis: DoE

Next to the individual effect of the variables (main effects), there may also be interactions that play a role. For analyzing both the individual and the interaction effect, a DoE set-up can be used. A few points about the application of DoE will be discussed below. A more extensive discussion of DoE can be found in e.g. Clarke & Kempson (1997), Dean & Voss (1999), Montgomery (2009), Montgomery & Runger (2002).

DoE is a systematic way to study the effects and interactions of input variables and aims to quantify outputs corresponding to the whole input space. Because the input space is very large, DoE provides a means of defining a set of inputs which is economical while representing the behaviour of the model with sufficient detail.

In DoE, a number of scenarios are defined in such a way, that comparison of the different scenarios enables the identification of main effects (effects of changes in the value of single input variables on the value of the output variable) as well as interactional effects (effects of changes in the values of combinations of variables on the value of the output variable, where the effect on the output variable of the combination of variables is either more or less than merely a summation of the main effects of the value changes of the individual variables). The number of scenarios that are defined in the context of DoE mainly depends on four factors, i.e.:

1. The number of input variables
2. The number of possible values of the input variables
3. The number of outputs
4. The expectation whether or not high-level interaction effects (interactions between more than two variables) are present

In order to estimate all interaction effects, the output has to be determined for all possible combinations of the different values of the input variables. This is called a full factorial experimental design. The number of scenarios that has to be analyzed can be limited by leaving out higher level interactions (i.e.: interactions between more than two variables). This is especially relevant when there are no indications that higher level effects are present. A DoE which consists of a subset of all possible scenarios is called a fractional factorial design.

In order to illustrate the way in which DoE can be applied in the case of BBNs, the DoE set-up will be discussed for the input variables in the C&D stage (the root nodes in Figure 11), addressing the four factors that affect the DoE set-up.

1. The number of input variables (in the case of BBNs, the root nodes), is six.
2. The number of possible values of the input variables is (theoretically) infinite, since the value is a distribution, consisting of three separate marginal probabilities, and any specific combination represents a specific value. However, for the two different ways of performing sensitivity analysis that have been defined (EESA and SESA), two different values for the input variable are used. Consequently, for each way of analysis, the number of possible values of the input variables is two. These are defined as follows:
    a. For EESA, the values are ‘1.00 – 0.00 – 0.00’ and ‘0.00 – 0.00 – 1.00’. These values are chosen since they represent the largest contrast, i.e.: values of the variables change strongly.

   b. For SESA, the values are ‘0.25 – 0.35 – 0.40’ and ‘0.20 – 0.30 – 0.50’. These values are chosen, because their contrast is limited, i.e.: the values of the variables differ only to a limited extent, the difference being ‘0.05 – 0.05 – 0.10’.

3. Although the number of output variables is 1 (‘Reliability effect C&D phase’, see Figure 11), the value of this variable is represented by a value that consists of three marginal probabilities (for the three states in the state-space of ‘Reliability effect C&D phase’). Therefore, in the analysis and DoE set-up, three outputs have to be taken into account.
4. As there is no reason to assume that higher-level interaction will play a role, higher level interactions are not taken into account in the DoE. Only main effects and first-level interactions (interactions between two variables) are analyzed.

The DoE set-up for sensitivity analysis has been defined using the computer program Statgraphics®. The set-up that has been chosen, is an orthogonal design, which enables the identification of all first-level interactions, and all main effects without these effects being confounded (the main effects can be separated from the interaction effects). In order to analyze main and first-level interaction effects, for EESA as well as for SESA, at least 32 scenarios have to be calculated. For the EESA, the two extreme values are ‘1.00 – 0.00 – 0.00’ and ‘0.00 – 0.00 – 1.00’; for the SESA the values are ‘0.25 – 0.35 – 0.40’ and ‘0.20 – 0.30 – 0.50’.

Using a DoE set-up, in total 64 values (32 for EESA and 32 for SESA) for the output variable ‘Reliability effect C&D phase’ have to be calculated, consisting of 192 probabilities (three marginal probabilities per value). Because through DoE analysis, only one output number per analysis can be taken into account, the resulting value cannot be analyzed as a whole. Rather, the three marginal probabilities for the three different possible states for the output variable have to be analyzed individually. Hence, in total six analyses have to take place, i.e. EESA as well as SESA for the states ‘reliability below target’, ‘reliability on target’, respectively ‘reliability above target’.

DoE analysis: Pareto diagram

The Pareto diagram for the EESA using DoE for the state ‘Reliability is below target’ is shown in Figure 18 underneath. The Pareto diagram shows, from top to bottom, the influences of the different variables (and their interactions). The amount of influence of a variable is depicted by the length of its related bar.

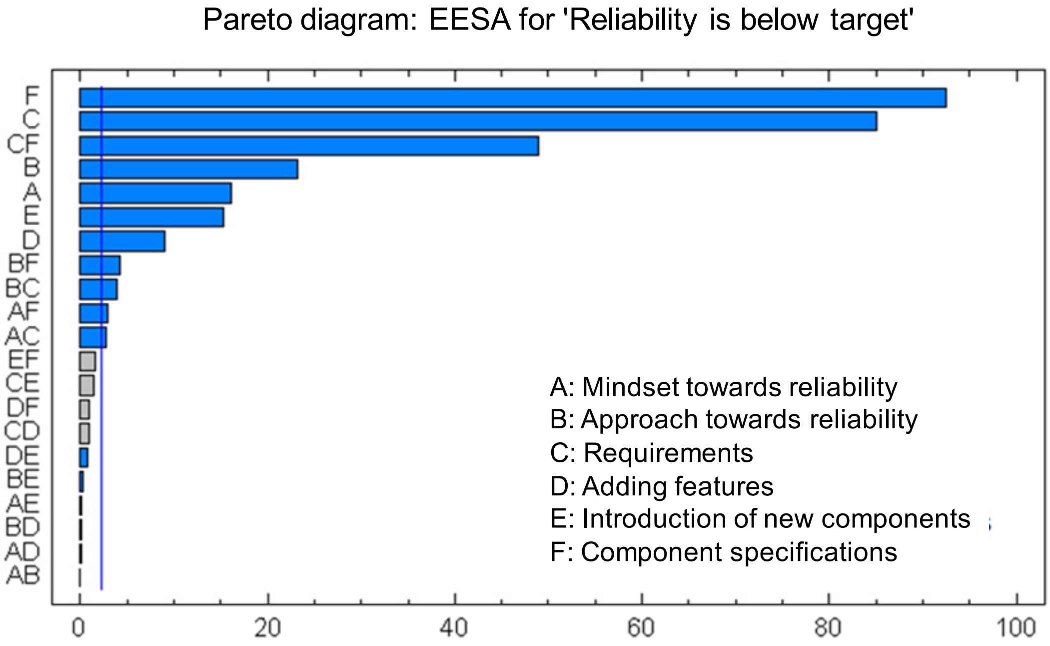


Figure 18: Example of a Pareto diagram, constructed using the software tool Statgraphics®. The diagram depicts EESA using DoE, including first level interaction effects, performed on the BBN as presented in Figure 11, the output variable being 'Reliability is below target'.

Figure 18 shows that, for the state 'Reliability is below target', 'component specifications' is the most influential input variable.

DoE analysis: Main effects plots

Next to the Pareto diagram, for the main effects and first level interactions, also plots can be made of two values of the input variable against one marginal probability of the value of the output variable. This gives a clear indication *how* the values of the input variables can influence the value of the output variable (when the other variables do not change), i.e.: the direction of the change is included in the main effects analysis, where it is not included in the Pareto diagram, where only the relative size of the effect is shown. Such a figure is made for the main effects of the six input variables as shown in Figure 11. In order to better comprehend the figure, it is important to note that the signs '+' and '–' are related to the correct states as defined in Table 9. These relations are represented in Table 16.

The interpretation of this table is as follows: for the variable 'Approach towards reliability' the value is defined to be '+', if the value of this variable is such, that the marginal probability for the state 'Always on the agenda of the MT-meeting' is 1. At the same time, the value of this variable is defined to be '–', if the value of this variable is such that the marginal probability for the state 'Once a year on the agenda of the MT-meeting' is 1 (see Table 16)

Table 16: Variables that are included in the BBN for the C&D stage and their state-spaces

Input variables	**Value as represented in Figure 19**	**State as defined in Table 9**
Approach towards reliability	+	Always on the agenda of the MT-meeting
	–	Once a year on the agenda of the MT-meeting
Mindset towards reliability	+	Clear activities, deliverables & evidence during the PCP
	–	Ad-hoc (due to problems) determining activities and deliverables
Requirements	+	Customer wish translated to system and sub-system level. And insight in the functions influencing this process
	–	Not able to translate requirements to system level
Component specifications	+	Knowledge on component & the conditions under which it has to perform is present
	–	No knowledge on component & no knowledge on the conditions is present
Introduction of new components	+	<10%
	–	>30%
Adding features	–	Many new features & functions
	+	Small number of new features & functions

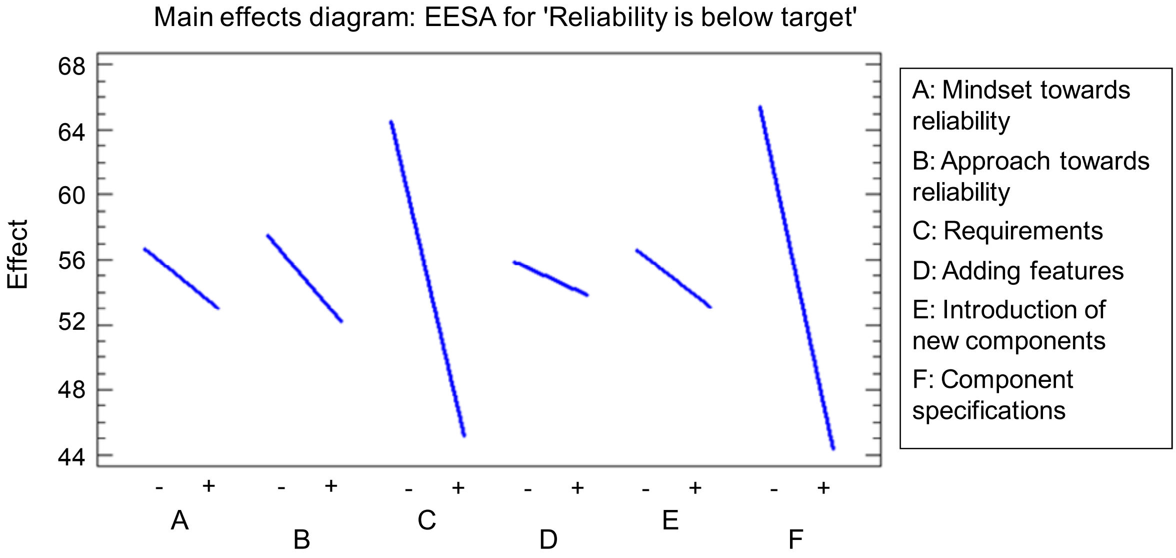


Figure 19: Example of a main effects diagram, constructed using the software tool Statgraphics®. The plots depict the main effects of the input variables on the output variable being 'Reliability is below target' related to the BBN as presented in Figure 11

The relationship between the input variables and the output variable in the main effects plot is depicted by the decreasing lines in Figure 19. It shows e.g. that the variable 'requirements' influences the probability of reliability being below target as follows: if ‘requirements’ has a negative value, then the probability of reliability being below target is around 0.65, whereas if ‘requirements’ has a positive value, then the probability of reliability being below target is around 0.45. The conclusion that is drawn from this picture resembles the conclusion from Figure 17 (sensitivity analysis through a Tornado diagram). However, for the analysis through Tornado diagrams, information of the current situation is needed, whereas this is not the case for main effects analysis through DoE.

DoE analysis: ANOVA

Beside the Pareto diagram and the main effects plots, also analysis of variance (ANOVA) can be performed using DoE. ANOVA is a statistical technique for analyzing whether or not relations between input and output variables are statistically significant. The technique of ANOVA will not be discussed in detail in this thesis. For a more in-depth discussion on ANOVA, see e.g. Montgomery & Runger (2002). In the case that is shown in Figure 18, the ANOVA analysis that is performed using the computer program Statgraphics® identifies 11 significant relations: all individual factors are deemed to have a significant effect, as well as the following combinations:

- ‘Requirements’ and ‘Component specifications’ (CF)
- ‘Approach towards reliability’ and ‘Component specifications’ (BF)
- ‘Approach towards reliability’ and ‘Requirements’ (BC)
- ‘Mindset towards reliability’ and ‘Component specifications’ (AF)
- ‘Mindset towards reliability’ and ‘Requirements’ (AC)

The results of the ANOVA have to be treated with care: although the interactional effect of ‘Requirements’ and ‘Component specifications’ seems logical – good specifications are especially beneficial when the customer wishes are well translated into requirements – the other interactional effects may not be logical or clear-cut. It may e.g. not be clear directly how the number of times that reliability is put on the agenda of the MT meeting (‘Approach towards reliability’) influences e.g. the effect of ‘Requirements’ on reliability.

Furthermore, the results are based on a statistical analysis, for which assumptions regarding the data have to be made. This means that, whereas the effects that are strongly significant can be assumed to be present, the presence of effects that are analyzed to be only marginally significant, should not be accepted blindly. Since the assumptions that have to be made for using the results of the analysis cannot be checked, the results of the ANOVA have to be interpreted with care, rather than used off-hand.

7.4. Reliability Prediction

As already indicated in section 3.3, scenario analysis and sensitivity analysis using Bayesian inference can be used for reliability prediction, decision support and model evaluation. In this section, reliability prediction will be discussed.

By calculating the model output as a result of different input values (prognosis) through inference, reliability may be predicted.

In order to illustrate the way in which reliability can be predicted through scenario analysis, a part of the BBN representing the C&D stage of the PDP is used. The same quantitative model that was used as a basis for the Figure 14 – Figure 16, is used for Figure 20. However, for the generation of Figure 20, the values of the input variables that are used in the model are pre-defined, rather than pre-set to a non-informative MPD. In reality, the values for the input variables may be elicited by experts through probability elicitation as presented in section 6.2 as well. Again, the figure has been generated using the software tool AgenaRisk (Agena Ltd., 2009).

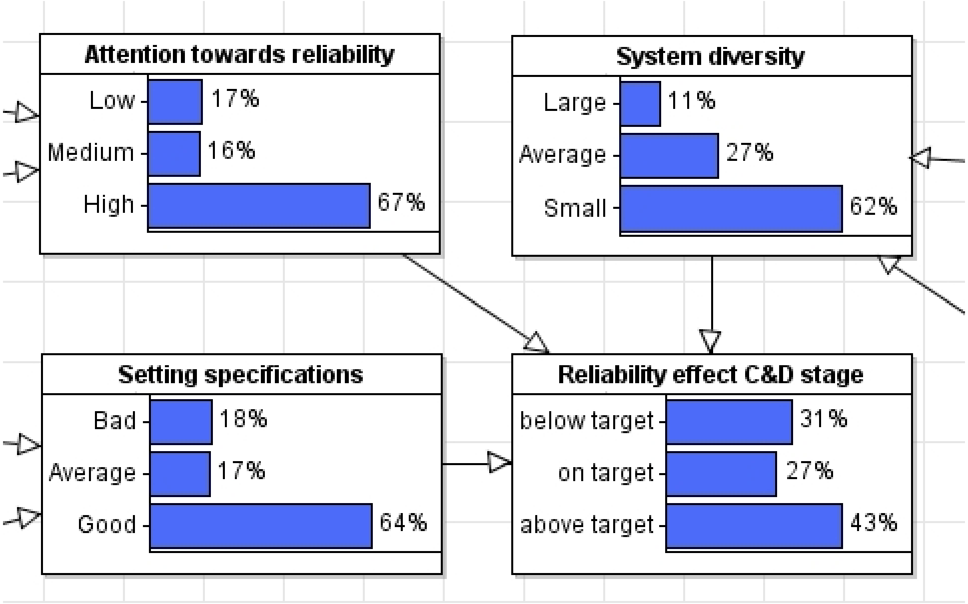


Figure 20: BBN model for reliability prediction based on inputs

As can be seen – based on pre-defined values of the input variables – it is predicted that there is a large probability (0.70) that the reliability of the system after the C&D stage will be on or above target.

7.5. Decision Support

Next to for reliability prediction, Bayesian inference may also be used for decision support. This may be done either through scenario analysis, or through sensitivity analysis. This will be discussed in subsection 7.5.1, respectively subsection 7.5.2.

7.5.1. Decision support through scenario analysis

Generally, there are three ways in which decision support can take place through scenario analysis based on Bayesian inference:

- The outcome may be calculated as a result of different input values (prognosis as identified in section 7.2).
- The likelihood of the values of the input variables can be calculated corresponding to a particular (desired) outcome (diagnosis as identified in section 7.2).
- A combination of both: calculating the most likely values of the input variables based on a (desired) outcome under certain conditions (diagnosis, one or more parent node values being fixed).

In this subsection, these three ways of decision support will be illustrated consecutively, using a part of the BBN representing the C&D stage of the PDP. The analysis is illustrated using Figure 15, Figure 16 and Figure 20. As reference point, Figure 14 (Scenario 1, page 116) is taken.

Decision support: prognosis

A first example of the way in which decision support may take place, is through prognosis: calculating the outcome (value of the child/leaf node), given certain inputs (values of the parent/root nodes). This is shown in Scenario 2 (Figure 15, page 117). In this figure, it can be clearly seen that, even if the attention towards reliability is high, the reliability will likely be below target (a chance of around 0.50). As such, this would not lead to the inclination to spend more attention to reliability. However, comparing Figure 15 with Figure 14, the effects of the high attention towards reliability become clear. Looking at the reliability effect, it can be concluded that the probability of reliability scoring on target has increased a bit (2%), whereas the probability of scoring above target has increased with 4%. As a consequence, the probability of scoring below target has decreased with 6%, marking a clear improvement of reliability performance.

Decision support: diagnosis

An example of the second type of decision support (calculating the most likely value of the input variables, given a desired outcome) is given in Figure 16 (Scenario 3, page 118). The interpretation of the model outcome as it is shown in this figure is not straightforward, i.e. the reliability being ‘on target’ can be reached by several value combinations of its parent nodes. The model represents the likelihood of the variables being in a certain state, when reliability is ‘on target’.

The model also shows that there is a chance that reliability will still be ‘on target’, even if the attention towards reliability is ‘low’, the system diversity is ‘large’ and setting the specifications is ‘bad’, although the likelihood of this combination of values is.

For further interpretation of the model, the results have to be compared again with the results as depicted in Figure 14. It then becomes clear that, in the case of ‘reliability effect C&D stage’ being ‘on target’, the most likely value of ‘setting specifications’, is ‘good’, rather than ‘bad’. This implies that, in order to get reliability ‘on target’, effort should be put into getting ‘setting specifications’ into the state ‘good’. These changes regarding the variable values (MPDs) do not indicate the effort that is needed to achieve this goal. This is one of the weak points of BBNs, i.e. BBNs focus on one attribute only (Fenton & Neil, 2001).

Note that the percentages that are identified through the model represent likelihoods, i.e.: they represent the likelihood of a variable being in a certain state. It is important to keep this interpretation in mind. In reality, a variable assumes a certain state (it does not have an MPD). It has to be stressed that the interpretation of the values of the variables given by the model is not that the value is a certain 'input setting', needed to obtain a result. Rather, the values represent the likelihood with which the variables assume a certain state, given that the output variable takes the value as defined.

Decision support: combining prognosis and diagnosis

Finally, the two types of decision support can be combined (see Figure 21).

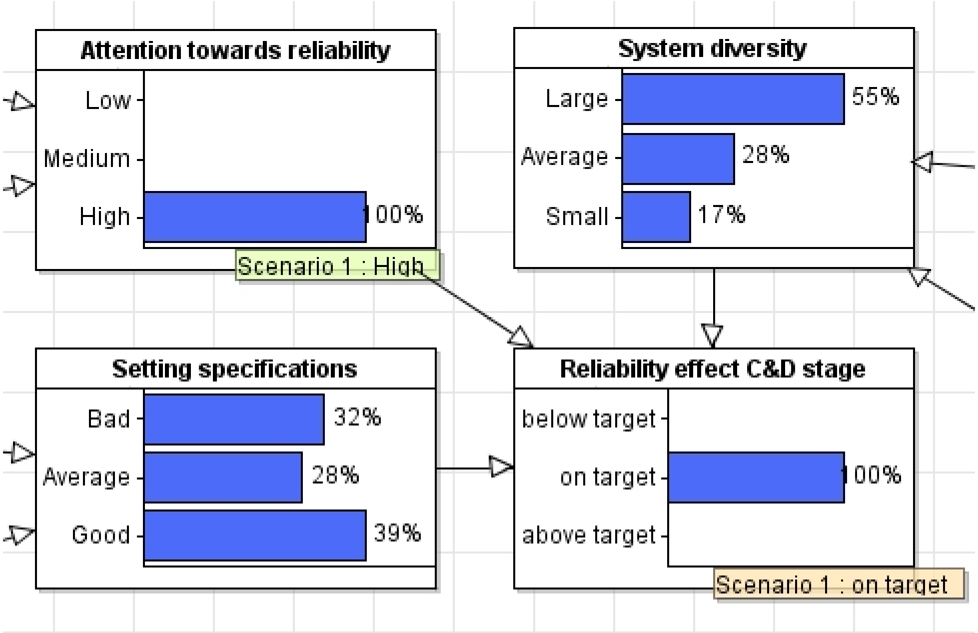


Figure 21: Example BBN model with the condition: 'Attention towards reliability' is 'high', and the desired outcome is 'Intrinsic (maximum) reliability is on target'

Comparing this case to the previous case, where the goal was to obtain reliability 'on target' without any conditions (Scenario 3, page 118, represented by Figure 16), it is clear that observing that 'Attention towards reliability' is 'high', does not affect the implication that is observed. The implication in this case is the same as the implication in the previous case (in order to get reliability 'on target', effort should be put into getting 'setting specifications' into the state 'good').

7.5.2. Decision support through sensitivity analysis

In section 3.3, it was stated that sensitivity analysis through Bayesian inference can be used for the purposes of decision support and model evaluation. In this subsection, the different forms of sensitivity analysis will be discussed, together with the way in which they can be used for decision support. The way in which they can support model evaluation will be discussed in section 7.6. Two types of sensitivity analysis (see section 7.3) are performed, i.e.: EESA and SESA.

Underneath, the use of EESA for decision support is discussed in more detail, followed by an elaboration on the use of SESA for this purpose. The way in which the two types of analysis provide decision support and their limitations are addressed.

Decision support through EESA

As identified in section 7.3, EESA can be useful for identifying the most important influential variables and it indicates the ranges over which the different variables can affect the output variable. As such, the most influential input variable can be identified: this information can be used to prioritize the changes in variables and through that, support decision making.

Furthermore, EESA is able to show the differences in the amount of influence that a certain factor has on the value of the output variable for the different, individual states (marginal probabilities in the MPD). As such, EESA enables decision support for obtaining a target outcome. In order to show this more clearly, an example is introduced. In the case that is represented in Figure 14, the target category 'above target' of the C&D stage is most strongly influenced by the variable 'setting specifications'. However, it might be that in the case of 'Reliability in the C&D stage = below target', a different variable (e.g. 'system diversity') influences the outcome most strongly. With respect to decision support, this would then imply the following:

- In order to obtain reliability results that are *above target*, one should first focus on 'setting specifications'.
- In order to obtain reliability results that are *not below target*, one should first focus on 'system diversity'.

An important characteristic of EESA is that it looks only at the extreme values (MPDs) of input variables, and does not take small changes into account. Because of this, it does not take the shape of the relationship between the value of the input variable and the value of the output of the model into account, i.e.: although the MPD of the output variable may change from '0.70 – 0.20 – 0.10' to '0.20 – 0.25 – 0.55', when the MPD of a model input changes from '1.00 – 0.00 – 0.00' to '0.00 – 0.00 – 1.00', no insight is gained about the relationship between the MPD of the input variable and the MPD of the model output, e.g. whether it is linear, or non-linear. The changes of the MPD of the output variable as a result of incremental changes of the input variables are not known.

As a result, the application of EESA for decision support is especially relevant at early stages of the PDP, where decisions are not very detailed, and are based on crude, preliminary information. This is because at the early stages of the PDP, necessary trade-off decisions can only be made on a high level of abstraction. Resulting changes, based on decisions in the early stages of the PDP are strong rather than subtle. Particularly EESA through DoE is helpful in these stages, since DoE does not need a point of departure for its analysis.

At the same time, the fact that the current state of the variables is not taken into account forms an important drawback of EESA through DoE, i.e.: although the most influential variable may be identified, this is not necessarily the variable with which the most improvement can be obtained at the current point in time. As stated in section 7.3, sensitivity analysis through Tornado diagrams can solve this problem, since they take the current situation into account by taking it as point of departure. At the same time, this implies that tornado diagrams are less useful in the beginning of the PDP, where no point of departure can be defined.

Decision support through SESA

In contrast to EESA, SESA looks at small changes in the value (MPD) of the input variables, and how these small changes affect the value (MPD) of the output variable. Because of this, SESA enables the exploration of the input-output relationship (which is usually a complex one) in the immediate surrounding of the current point of departure. In this way, the most influential input variable can be identified for this surrounding. Note however that only main effects are taken into account.

A clear drawback of this approach is that the scope of this exploration is limited, i.e. it may not be valid outside the immediate surrounding of the current point of departure. In order to create a more complete view on this input-output relationship, the SESA has to be repeated for a large number of different points of departure. Such a repetition of SESA using DoE would enable the identification of the most effective way of improving reliability, defining which variables to change (to a pre-set, limited amount of change in terms of MPD, e.g. '-0.04 – +0.02 – +0.02') and in which sequence, in order to sort the largest effect possible. It should be taken into account that the decision that is to be made may depend on the target of the decision, i.e.: it may be that for obtaining reliability 'above target', a different decision has to be made than for obtaining reliability that is not 'below target'.

The application of SESA for decision support is especially relevant at later stages of the PDP, where decisions are made on a more detailed level, and typically relate to small changes and compromises that have to be made. In this way, SESA can help to identify the most beneficial changes with limited resources, as well as the least harmful cases, if compromises have to be made in the case that resources are insufficient.

By comparing Figure 22 (which shows a Pareto diagram of the SESA of the using DoE for the state 'Reliability is below target') with Figure 18, it can be illustrated that different decisions can be made using EESA than using SESA. Clearly, there are a number of differences between the two graphs. One of these is that the input variables 'Component specifications' (F) and 'Requirements' (C) have switched positions. This implies that, in the case when EESA is used, the focus would lie on 'Component specifications', whereas the use of SESA (from the defined point of departure, i.e.: the values of all input variables are pre-set: '0.25 – 0.35 – 0.40', see subsection 7.3.2) would lead to a focus on the variable 'Requirements'.

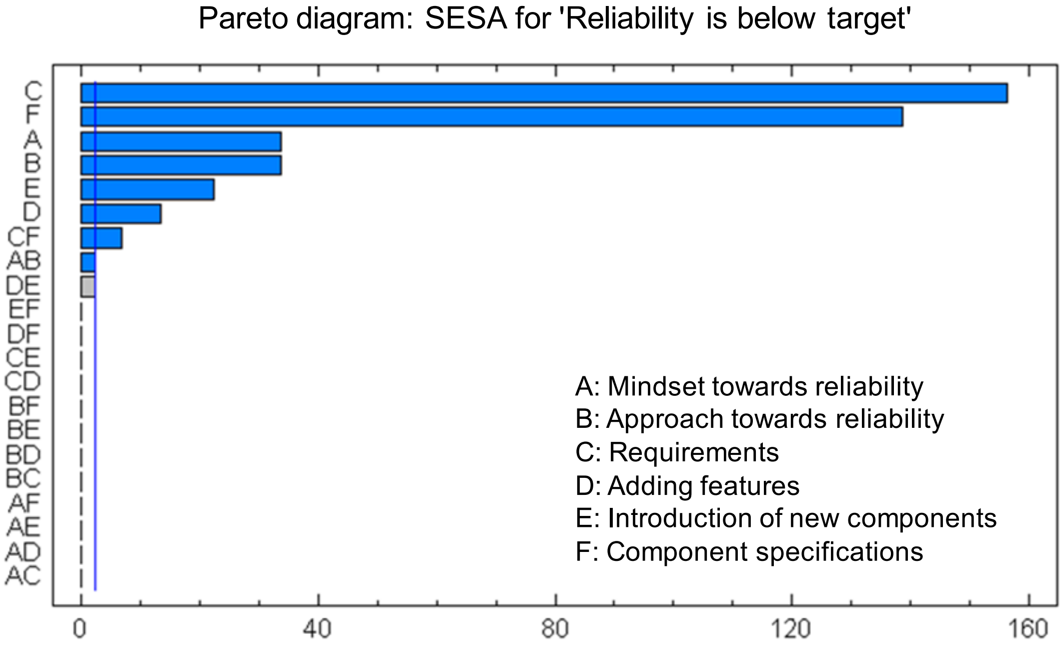


Figure 22: Pareto diagram, constructed using the software tool Statgraphics®. The diagram depicts SESA using DoE, including first level interaction effects, performed on the BBN presented in Figure 11

Next to the Pareto diagram, also an ANOVA can be performed using DoE. As identified in subsection 7.3.2, the results of the ANOVA have to be treated with care. Off-hand use of the results may lead to wrong decisions.

When using sensitivity analysis for decision support, it has to be taken into account that the amount of influence that a certain factor has on the target variable may be different per possible state of the target variable. This can be seen e.g. by comparing Figure 17 with Figure 23: in the situation where the 'reliability effect C&D stage' = 'above target', the range of influence that is related to the variable 'requirements' is 0.20 (see Figure 17), whereas this same range is 0.06 (more than twice as small) for the case where the 'reliability effect C&D stage is 'on target' (see Figure 23). This goes for EESA as well as for SESA.

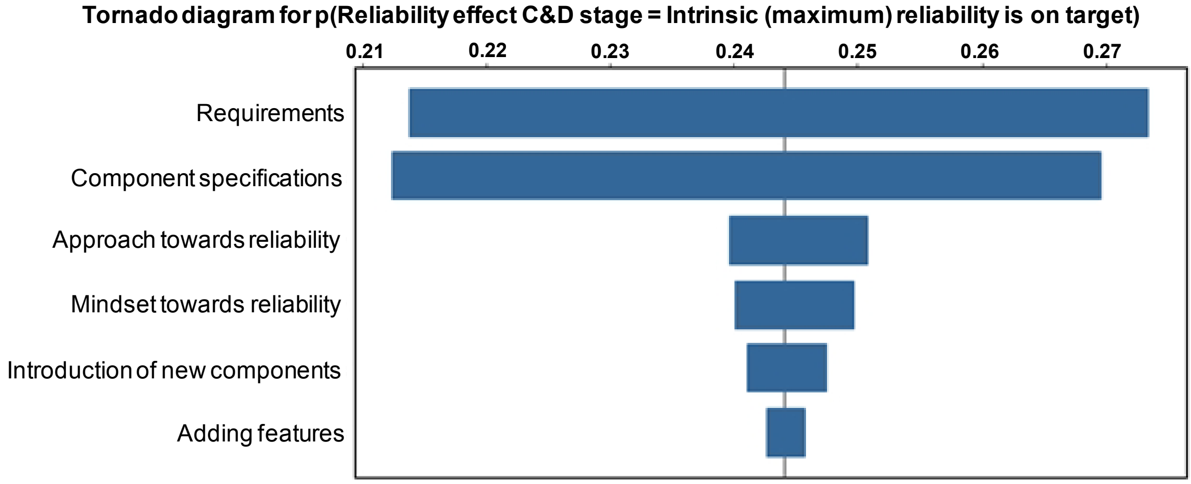


Figure 23: Tornado diagram, constructed using the software tool AgenaRisk (Agena, ltd., 2009). The diagram depicts EES analysis, performed on the BBN presented in Figure 11

7.5.3. Limitations regarding decision support through Bayesian inference

As stated in section 3.3, the use of BBNs for decision support also has its limitations. The most important reason for this (Fenton & Neil, 2001) is that BBNs often focus on one single attribute. This implies that indications that are given by the BBN will only take this single attribute into consideration. Decision making however, is often based on multiple decision criteria. Also in the interviews it has been found that reliability is not the only criterion that is taken into account in the PDP: e.g. also TTM and costs play an important role. The BBNs as developed in this thesis are not capable of making trade-offs between different alternatives based on multiple criteria, e.g. such as cost and effort. For this purpose, Fenton & Neil 2001 propose to complement the BBN with other decision making techniques, such as e.g. multiple criteria decision aids. Since the construction of a BBN in this thesis has not been aimed at incorporating such tools, this will not be discussed further here.

The BBN construction process that is described in this thesis focuses on the construction of a BBN that is able to provide insights in the way in which reliability is affected by different factors that play a role in the PDP. This is done based on the perspectives of the experts (soft systems approach). In the context of decision support, the nature of the model (a systems model) enables the user of the BBN to take both physical and non-physical aspects into account. Furthermore, the nature of the BBN construction process enables addressing the PDP up-front based on subjective information. However, the use of an SSM approach in the BBN construction process enforces the use of crude, mainly qualitative estimations.

With respect to decision support through scenario analysis and sensitivity analysis, these types of analysis only give indications of the ranges over which the input variables influence the output variable. As already discussed in section 3.3, large changes regarding the value (MPD) of an input variable may lead to small changes regarding the value (MPD) of the output variable. Therefore, rather than using sensitivity analysis for decision support in terms of absolute probabilistic changes, sensitivity analysis should be used to provide support in terms of direction and in terms of ranking the input variables from most influential to least influential.

Finally, the use of Bayesian inference for scenario analysis and sensitivity analysis is directed at the current situation, i.e.: the situation as it was modelled in the problem structuring and instantiation phase. In order for scenario analysis and sensitivity analysis to be applied, both the BBN problem structure (influencing variables) and the BBN relations (influences) may not change.

Scenario and sensitivity analysis will only be usable for supporting decision making in the prevailing circumstances at the time at which the model is instantiated. Long-term situation changes regarding variable influences or even which variables to include cannot be incorporated, i.e.: decision support over a longer time window (over time, the included variables may influence reliability differently, or other variables may have to be included) is not possible through Bayesian inference.

Concluding, three important general limitations can be identified with respect to the way in which decision support is provided by BBNs constructed through the process described in chapters 4, 5 and 6:

1. Only one attribute (in this case reliability) is taken into account when constructing the BBN. Other attributes, like e.g. costs and effort, are ignored.
2. The inputs and outputs of the model are crude and qualitative, resulting in a low level of detail regarding decision support.
3. Using Bayesian inference (either through scenario analysis or sensitivity analysis) for decision support reasons from the current situation, and is not valid over a longer time window.

7.6. Model Validation

As discussed in subsection 3.3.3, validation efforts with respect to the instantiation stage of BBN involve the experts that participated in the model construction process. Since the model is built based on the collective knowledge of a number of experts, the validation of the model also has to take place collectively, i.e.: in the form of a group session. The way in which such a session was organized for BBN validation is described in the next subsection (7.6.1). The results of the validation session are described in subsection 7.6.2.

7.6.1. Organization of the validation session

Validation of the model outputs took place in one validation session with a focus group. The choice for validation through a group session was made based on the fact that the model was a result of the collection of knowledge from a number of experts. Since the model is based on collected knowledge, the collective knowledge has to be used for validating the model. A focus group was chosen as the way to obtain information from the experts. This was mainly done because focus groups enable interaction between participants, and are able to make use of the collective knowledge. Krueger & Casey (2009) state that a focus group normally consists of 5 – 10 people and that randomization can be used to help ensure a nonbiased selection of participants. For this focus group meeting, unfortunately availability dictated the selection of participants, rather than that a random sample of people was chosen for the model validation.

Ten people were invited to participate in the group session for model output validation. From the ten people that were invited, five experts participated (see appendix B). Although this number was small, due to availability reasons, it was not possible to involve more people in the focus group meeting.

The experts that were invited for the group session were selected based on information-richness (a criterion also used by Jacobs & Van Moll, 2007, who refer to (Patton, 1990) for this criterion) and on availability. Information richness in the case of this thesis was based on the number of concepts that was identified by the experts: the expert that identified the largest number of concepts (obtained through applying the adapted GT approach) was selected first to be included in the model validation session. Although the experts were already interviewed for eliciting model probabilities, there was a long period of time between model validation and elicitation of model probabilities, so that bias was avoided as much as possible. Unfortunately, because of the limited availability of experts, only a few people that were initially invited were able to attend the focus group meeting. Therefore, two others were invited, who had not participated in the interviews for the GT sessions. Details on the five participants in the group session can be found in appendix B.

Two moderators were present during the session. One of the moderators was the researcher who presented the modelling approach and the final models. The other moderator was the gate-keeper in the company, who, due to his knowledge, was able to guide the discussion on the models. The focus group session was planned to last one and a half hours.

The model as it was defined through adapted GT as well as elicitation of the probabilities were presented and shortly discussed. After that, the working of BBNs was introduced, using a simple, real-world example. Then, the model output was validated by evaluating the behaviour of the models of the separate stages of the PDP, as well as the overall model, combining all stages in the group session.

In the group session, consensus had to be reached on whether or not the model behaviour reflects the collective belief of the experts. If the model behaviour was judged to reflect the view of the experts, then the model was considered valid. If this was not the case, then the shortcomings of the model had to be identified. Possible shortcomings could be identified in terms of the way in which changes in the input variable values affect reliability.

Beforehand, the group session was prepared. Preparation for the group session was done together with the gate keeper, who has insights in the whole PDP. Preparation included (for all models):

- Performing a main-effects analysis for all possible states of reliability (3) in all sub models. This is done in order to evaluate whether changes in the values (MPDs) of the different variables of the model result in expected changes in the value of reliability.
- Identifying sensible scenarios for model validation.
- Identifying sensible changes in these scenarios.

Next to defining scenarios and model changes beforehand, also feedback from the experts during the model validation session should be used to define other model scenarios and other sensible changes. The model had to be validated through a focus group meeting, since it should reflect the ‘opinion of the group’. Within the focus group meeting, the models of the different stages were presented. In the presentation, for every model, the results of the specific main effects analysis were presented. Next to that, the opportunity was provided for the participants to give input for scenario analysis.

Furthermore, for validation purposes, a survey was given to the members of the focus group, which they were asked to fill in at certain set moments throughout the presentation. The questions that were asked in the survey can be found in appendix H. Note: the survey is in Dutch. A translated version is also provided in appendix H; this translation has not been validated. The way in which the questions in the survey helped to validate the model is discussed in the next subsection.

Prior to the meeting, the focus group presentation was defined as follows (Table 17):

Table 17: Focus group presentation as prepared beforehand

Time	Activity
5 minutes	Presentation of the status of the research.
15 minutes	Presentation on the way in which BBNs work, using a simplified, real world example.
15 minutes	The C&D model variables and the results of the main effects analysis are presented. Furthermore, the C&D model is presented and its behaviour is discussed. Finally, the survey question regarding the C&D model is filled in.
15 minutes	The D&D model variables and the results of the main effects analysis are presented. Furthermore, the D&D model is presented and its behaviour is discussed. Finally, the survey question regarding the D&D model is filled in.
15 minutes	The M&I model variables and the results of the main effects analysis are presented. Furthermore, the M&I model is presented and its behaviour is discussed. Finally, the survey question regarding the M&I model is filled in.
15 minutes	The PDP model variables and the results of the main effects analysis are presented. Furthermore, the PDP model is presented and its behaviour is discussed. Finally, the survey question regarding the PDP model is filled in, as well as the remaining survey questions that are related to model evaluation and the focus group session itself.
10 minutes	Extra time reserved for overrunning the schedule

Next to the focus group meeting, also a presentation was given for a decision maker (see appendix B), to evaluate the model individually with the decision maker.

7.6.2. Results of the validation session

In order to obtain an indication of model output validation, for all separate models, the level of reliability of the results according to the participants was measured. This was done using the first four questions presented in appendix H, and using the measure of reliability presented by Miles & Huberman (1994) (see section 5.3). Regarding the questions, the scoring was applied that was proposed by Jacobs & Van Moll (2007), i.e.:

- 1, 2 or 3 is equal to disagreement
- 4 or 5 is equal to agreement

This resulted in the scores as presented in Table 18:

Table 18: Level of reliability with respect to model output

PDP stage	Average score	Reliability
C&D	3.2	0.6
D&D	4.8	1
M&I	4.6	1
Overall	4.2	0.8

Jacobs & Van Moll (2007), referring to Miles & Huberman (1994) state that 0.8 or above is the score for which the reliability of the results are considered to be satisfactory. As a result, the conclusion can be drawn that the models of the D&D stage, the M&I stage and the PDP can be considered satisfactory in the eyes of the participants. The model for the C&D stage was not considered satisfactory. A possible reason for this was found during the discussion in the focus group. In the discussion of the results of the main effects analysis and scenario analysis of the model of the C&D stage, it became clear that there were two possible reasons for this lack of reliability, i.e.:

- Differences in interpretation of some of the included factors.
- Different opinions of the participants within the group

An example was provided in the interpretation of the variable ‘approach towards reliability’, which reflects the priority level of reliability for the company (description). The corresponding state space relates to the place of reliability on the agenda of the MT meeting. One of the participants stated that he assumed that a higher place of reliability on the agenda automatically means that more resources and money are spent on reliability management. This led him to the statement that ‘approach towards reliability’ should have a stronger influence on reliability, than the influence that was given by the model. Others did not agree on this statement, and did not automatically relate a higher place on the agenda of the MT meeting to more resources and money available and spent on reliability management efforts. In this case, implicit assumptions regarding a variable resulted in the model being judged differently (people agreeing or not agreeing on the model).

Evaluation of the model through a focus group meeting provides a good opportunity for improving the model, especially since the participants indicate what problems they have with the model and model variables.

Next to the models themselves, the method with which the models were validated (a focus group meeting, combined with a survey) was evaluated, using the questions as presented in the survey given in appendix H. The results of the survey showed that the participants in the focus group agreed that the models increased their insights in the way in which the factors influence reliability and that the models were useful for managing reliability.

Not all participants judged the ways that were used to validate the models (main effects analysis and scenario analysis) to be adequate (for both, the level of reliability was 0.6). A number of remarks were made regarding the notion that calibration of the model was needed in order to take full advantage of the model.

In the individual discussion with the decision maker, the decision maker assessed the model to be useful, although the applicability of the model was at a high level of abstraction. Relating the model to concrete actions would strengthen the practicability of the model.

Validation in the form of benchmarking the model output against real time performance was not possible in the time period that was given. Model validation using real data obtained in the field may become possible later; BBNs are able to incorporate information regarding the variables in the model, as it becomes available over time.

7.7. Summary and Conclusions

The ways in which Bayesian inference can be used, and the purposes for which they can be used are as follows:

1. Scenario analysis (see section 7.2), either calculating the effect of changing the input values for the model on its output, or calculating the conditions (input values) that are needed for obtaining a certain output.
   Scenario analysis can be used for reliability prediction, decision support and model validation.
2. Sensitivity analysis (see section 7.3), either calculating the effect of extreme changes in the value (MPD) of an input variable on the value (MPD) of the output variable (EESA), or calculating the effect of a small change in the value (MPD) of an input variable on the value (MPD) of the output variable (SESA).
   Sensitivity analysis can be used for decision support and model validation.

Three important limitations with respect to the way in which decision support is provided by BBNs constructed through the process described in chapters 4 – 6 are:

1. Only one attribute (in this case reliability) is taken into account when constructing the BBN, ignoring other attributes like costs, effort, etc.
2. The inputs and outputs of the model are crude and qualitative, resulting in a low level of detail regarding decision support.
3. Using Bayesian inference (either through scenario analysis or sensitivity analysis) for decision support reasons from the current situation, and is not usable in a longer time window.

For model validation purposes, a focus group meeting was organized. Using main effects plots, scenario analysis, and a survey, the different BBNs for the different stages of the PDP, as well as the model for the PDP were evaluated. The results of the meeting show that experts generally agreed upon the model. It has to be noted that unfortunately, only a limited number of experts were able to participate in the meeting.

8. Conclusions and Recommendations for Future Research

In this chapter, the conclusions to the research questions, the scientific and industrial contribution, as well as suggestions for future research are discussed. The first section of this chapter gives a research overview, discussing the research objectives and related research questions, as well as the implications of the research and its validity and reliability, reflecting on the model building process. Then, in section 8.2, the research contribution is discussed. The scientific contribution is discussed in subsection 8.2.1 and the industrial/practical contribution is elaborated in subsection 8.2.2. In section 8.3 a reflection on the research is then given, based on the developed step by step BBN application process. The chapter concludes with recommendations for future research in section 8.4.

8.1. Research Overview and Implications

This section consists of four subsections. In subsections 8.1.1 and 8.1.2, the research questions and their answers are presented. In subsections 8.1.3 and 8.1.4, validity, reliability, and implications of the research will be further discussed. To put the research questions in the right perspective, their related research objectives are presented underneath, i.e.:

Research objectives:

1. To find or develop a reliability prediction and management method which is able to address the criteria that result from the increasing attention that is paid to LCC, as well as the increase in complexity, customer demands, TTM pressure, and globalization

2. To provide a systematic approach for applying the reliability prediction method developed in industry

8.1.1. Research Objective 1

Throughout chapter 2, the first research objective was addressed, including the three research questions related to this first objective:

1.1 Which criteria can be identified for reliability prediction and management methods that relate to the increasing attention for LCC, increasing complexity, customer demands, TTM pressure and globalization?

1.2 Is there a modelling method that can meet these criteria?

1.3 Can this method be applied for reliability prediction and management?

The first three research questions were answered based on literature. Analysis of the current business trends in the capital goods industry, led to the identification of four criteria that apply to a method for reliability prediction and management, i.e.:

1. The method has to give insights in order to support decision making
2. The method has to take non-technical factors into account
3. The method should be usable throughout the PDP
4. The method should be able to incorporate uncertainty

In chapter 2, different available methods for reliability prediction and management in industry are described based on literature research. These include both traditional (FMEA, FTA, database methods) and recently developed methods (REMM, PREDICT, TRACS). The methods are tested against the criteria identified through answering the first research questions, which leads to the identification of BBNs as suitable method for reliability prediction and management. Moreover, literature research has shown that it is possible to apply BBNs in practice for the purpose of reliability prediction and management (Neil et al., 2001); although a systematic approach towards BBN construction is not yet available.

8.1.2. Research Objective 2

In order to address the second research objective, the construction of such a BBN is presented throughout chapters 4, 5 and 6, describing a case study in the capital goods industry. In this way, the three research questions connected to research objective 2 are addressed, i.e.

2.1 How can the system boundaries of a structural model for reliability prediction be identified and defined systematically, when no information regarding the boundaries is available beforehand?
2.2 How can the variables that are to be included in the structural model for reliability prediction be identified and defined in a systematic way?
2.3 How can the structure of the structural model for reliability prediction be identified and defined in a systematic way?

In chapters 4, 5 and 6, the second research objective is addressed by applying an algorithm for BBN construction, based on GT. This algorithm is shown in Table 19.

Research questions 2.2 and 2.3 are specifically addressed through the algorithm itself (in steps 2, 3, 4, 7 and 8)

By involving a large number of people from diverse backgrounds in the adapted GT process, it was possible to include many different viewpoints, and avoid overstressing a single person's opinion. Also, because many people from different backgrounds were involved, the area that was taken into account was deliberately kept large, and no information was needed beforehand. Furthermore, it took care that that no key factors were excluded from the model. The adapted GT algorithm provides a systematic approach which leads to the identification of factors and their relations (addressing research question 2.2 and 2.3), implying the emergence of problem boundaries, when no prior information is available (addressing research question 2.1).

Table 19: BBN construction algorithm (see section 6.7)

1. Gather information regarding the way in which the topic under discussion is influenced by conducting interviews (see section 5.1)
2. Identify the factors (i.e. nodes) that influence the topic, by analyzing and coding the interviews (see subsection 5.2.1)
3. Define the variables by identifying the different possible states (state-space) of the variables through coding and direct conversation with experts (see subsection 5.2.1)
4. Characterize the relationships between the different nodes using the idioms through analysis and coding of the interviews (see subsection 5.2.2)
5. Control the number of conditional probabilities that has to be elicited using the definitional/synthesis idiom (see subsection 5.2.2)
6. Evaluate the BBN, possibly leading to a repetition of (a number of) the first 5 steps (see section 5.4)
7. Identify and define the CPTs that define the relationships in the BBN (see subsection 6.2.1)
8. Fill in the CPTs, in order to define the relationships in the BBN (see section 6.3)
9. Evaluate the BBN, possibly leading to a repetition of (a number of) earlier steps (see section 7.6)

8.1.3. Validity and reliability of the research

In this subsection, both reliability and validity of the case study and the resulting problem structure (problem structuring stage), and validation with respect to the functioning of the model (instantiation and inference stage) have to be discussed.

Problem Structuring

Validity and reliability in the problem structuring phase were addressed in subsection 4.4.6. In that subsection, different strategies were discussed that increased internal and external validity, as well as reliability. A number of strategies were identified that were used to increase internal validity, such as e.g. taking findings back to the field (using a survey, see appendix G), giving a rich, thick description of the performed research (in chapters 4 – 6), and spreading the data collection activities over a longer period of time (from February 2007 until September 2008 (19 months)). Also reliability was addressed, by providing a detailed description of the research focus, the researcher’s role, the position of the interviewees, the basis for their selection and the context from which the data will be gathered.

Because the research was limited to one case study, the external validity (transferability of the research results) was not addressed. Because of the nature of the case study, the expectation is that the constructed BBN is not directly applicable in other situations. Possibly, other BU’s within the same company share the same BBN structure, although the (probabilistic) definitions of the relations may be different. Also, companies in the same field (producing medical scanning equipment), may be able to use the model to some extent, since the activities that they have to display may be assumed to be largely equal to the activities of the company in the case study. It is assumed that when one comes further away from the specified context and environment of the case study, it becomes less plausible that the BBN is valid, and that the extent to which it is valid will decrease.

A more elaborate discussion on validity and reliability of the research and the resulting problem structure can be found in sections 3.1 and 5.3.

Model behaviour

Because the constructed BBN is a representation of the view of the interviewees, the interviewees also had to be involved in the validation process. This was done by using a focus group meeting of 5 people to validate the working of the model. In the focus group meeting, main effects analysis and scenario analysis were used to evaluate the working of the model. A survey instrument was used to let the experts evaluate the model individually, throughout the presentation (see appendix H).

A more elaborate discussion on validation of the model behaviour can be found in section 7.6.

8.1.4. Implications

The adapted GT approach that is introduced in this thesis (see Table 19) provides a systematic and structured way for the identification and definition of model boundaries and variables, as well as for the identification of relations. Although only one case study has been performed, it seems applicable for construction of BBNs in any context, not restricting its use for constructing a BBN for reliability prediction and management. It may even be possible to use parts of the adapted GT approach for the identification of variables, relations and problem boundaries for any qualitative model, not limiting the application of the approach to constructing BBNs.

It is important to take into account that the BBNs constructed through the adapted GT approach are based on the mental models of experts, rather than on physical models. Hence, possible predictions or decisions that are supported through BBNs constructed in this way are based on peoples' perspectives.

The research objectives defined the problem context, i.e.: reliability prediction and management. However, looking at the application of the adapted GT approach for building a BBN for reliability prediction and management, it is important to note that the use of the adapted GT for the construction of BBNs addresses in specific the *criteria* that are *related* to reliability prediction early in the PDP, namely:

- The method has to give insights in order to support decision making
- The method has to take non-technical factors into account
- The method should be usable throughout the PDP
- The method should be able to incorporate uncertainty

As such, the ability of the adapted GT approach to construct BBNs on the basis of expert knowledge appears to be independent of the context of predicting and managing reliability.

The construction of BBNs using the adapted GT approach seems to be applicable in the early stages of the PDP also to other product characteristics, such as e.g. costs or TTM. Looking across the borders of the research objective reliability prediction and management, it looks promising to use the adapted GT approach for the construction of BBNs in the context of many product performance characteristics.

The environment in which the BBN was constructed for the purpose of reliability prediction was a BU in a company that produces medical scanning equipment. This environment seems only partly relevant for the application of adapted GT for BBN construction for reliability prediction. The adapted GT approach seems generally applicable in environments where knowledgeable experts are present.

Predictions and decision support that are provided by the BBNs constructed through adapted GT have a probabilistic, qualitative nature, rather than a quantitative nature. Therefore, BBNs that are constructed through this approach cannot be used for the purpose of providing a quantitative estimate of reliability parameters such as e.g. mean time between failures (MTBF).

Moreover, in order to provide decision support and predict reliability through scenario and sensitivity analysis (see chapter 7), the constructed BBN has to be assumed to be valid. Although this may be true at the moment at which the BBN is constructed, this might not be true at a later point in time, since other variables (as identified and defined in the problem structuring stage of BBN construction) may come to play a role in the problem area, and because relations between variables (as defined in the instantiation stage of BBN construction) may alter.

8.2. Research Contribution

The contribution of this research is twofold, i.e.: both scientific and practical. The scientific contribution relates mainly to the development of a subjective rational approach that provides a structural, systematic way of constructing and using BBNs.

The practical contribution lies in predicting and managing reliability already early in the PDP, when no adequate empirical objective data are yet available that are adequate for reliability prediction. Moreover, the systematic way in which such a reliability prediction can be obtained, provides added value to industry. This approach provides a way to manage reliability (and through reliability, LCC) and focuses on management on a tactical level.

8.2.1. Scientific contribution

As stated, the main contribution lies in the development of a structural, systematic way of applying the BBN modelling technique. The proposed approach is developed by adapting GT.

An important contribution of this research lies in the development of a systematic approach that enables addressing both aspects (i.e.: dependency on experts and lack of objective data) at the same time, taking GT, a method from the social sciences, as point of departure.

GT is a method that enables involvement in the field of study, and including experts in the research. In this way, the lack of objective data early in the PDP is addressed. Furthermore, GT is a methodology that provides a systematic approach for developing (grounding) of theory. Although the purpose in this thesis is not to develop a theory, but rather, to construct a BBN, the GT approach as such was adapted for use in the context of BBN construction, developing a subjective, rational approach to model building.

BBN application consists of three stages: problem structuring, instantiation and inference. In its adapted form, GT provides a systematic approach for problem structuring (see Table 19):

Table 19: BBN construction algorithm (see section 6.7)

1. Gather information regarding the way in which the topic under discussion is influenced by conducting interviews (see section 5.1)
2. Identify the factors (i.e. nodes) that influence the topic, by analyzing and coding the interviews (see subsection 5.2.1)
3. Define the variables by identifying the different possible states (state-space) of the variables through coding and direct conversation with experts (see subsection 5.2.1)
4. Characterize the relationships between the different nodes using the idioms through analysis and coding of the interviews (see subsection 5.2.2)
5. Control the number of conditional probabilities that has to be elicited using the definitional/synthesis idiom (see subsection 5.2.2)
6. Evaluate the BBN, possibly leading to a repetition of (a number of) the first 5 steps (see section 5.4)
7. Identify and define the CPTs that define the relationships in the BBN (see subsection 6.2.1)
8. Fill in the CPTs, in order to define the relationships in the BBN (see section 6.3)
9. Evaluate the BBN, possibly leading to a repetition of (a number of) earlier steps (see section 7.6)

8.2.2. Industrial/practical contribution

The BBN that is constructed using the adapted GT approach addresses the fact that reliability prediction has to take place very early in the PDP, where no objective data are available. As such, the identified approach towards BBN application is able to:

- Predict reliability (taking both soft and hard factors into account, not taking information from the use phase into account)
- Provide a way to manage reliability on a tactical level

Reliability prediction

Reliability prediction is possible through using the BBN. By using scenario analysis, reliability (output of the BBN) can be predicted based on values of the input variables for the

BBN. These values – the MPDs –may be estimated, e.g. through probability elicitation. The output of the model is then a prediction of the reliability.

It is important to note that the reliability prediction and management model incorporates soft, intangible factors (such as e.g. ‘supplier quality’, ‘FMEA’, or ‘quality of the design process’; see appendices D and E). This strengthens the ability of the model to support decisions regarding reliability management in the early stages of the PDP. Note also that no information from the use phase is taken into account in this prediction, indicating clearly the difference between reliability prediction and reliability forecasting.

Tactical reliability management through BBNs

Reliability management can be obtained through BBNs in the form of decision support. Decision support by BBNs can basically be provided through scenario analysis and through sensitivity analysis. Through sensitivity analysis, the effect of extreme changes of the value of the input variable on the output (EESA) can be determined, as well as the effect of small changes in the value of the input variable on the output (SESA). In this way, sensitivity analysis allows for ranking the different input variables according to their order of influence on the output.

In EESA through DoE, no point of departure is identified (it is a general sensitivity analysis, not taking the current situation into account). SESA through DoE (like both types of sensitivity analysis through tornado diagrams) takes the current situation into account. SESA is also able, through repetition of the analysis using DoE, to identify the most effective way of improving reliability, defining which variables to change in which sequence, in order to sort the largest effect (see subsection 7.5.2).

Concluding, EESA through DoE seems to be most effective for supporting decisions early in the PDP, where no decisions have been made yet, and the decision support may be crude. At the same time, SESA through DoE seems to be more effective for supporting decisions that are made on a more detailed level, such as e.g. compromises, i.e.: identifying the most beneficial change with limited resources, or identifying the least harmful case, if resources are insufficient.

8.3. Reflection: Application of BBNs for Reliability Management

In chapter 7, it has been shown how BBNs can be used for reliability prediction and reliability management. Clearly these approaches to decision support and reliability prediction assume that the BBN that is constructed is appropriate, i.e.: that the model is considered to be valid. Although this may be the case at the moment at which the BBN is constructed, the empirical nature of the modelling does not guarantee validity at a later point in time.

8.3.1. Critique

The BBN application process consists of three stages Sigurdsson et al., (2001), i.e.:

1. Problem structuring
2. Instantiation
3. Inference

As identified in this thesis, in the instantiation stage, relationships are defined, whereas in the problem structuring stage, model boundaries and variables are identified and defined, and relations are identified. Over a longer time horizon, the way in which the input variables influence reliability and the input variables themselves may change. Therefore, the BBN construction process may have to be repeated to support reliability management.

The estimation of the way in which the different input variables affect reliability can change over time, as new insights, information, or objective data become available. Measures may be taken to influence the effect of input variables on reliability. Since the change of the estimated influence of input variables on reliability only becomes visible in the instantiation stage, only repetition of the instantiation stage would give insights into the changes and their effect on the predicted reliability and the resulting decisions.

Furthermore, other input variables may become relevant for the estimation of reliability. The rapidly changing environment of product development (due to e.g. the increasing pressure on TTM) may lead to other input variables becoming relevant to reliability. Since input variables are identified in the problem structuring stage, this stage should also be repeated to fully profit from BBN construction in the context of reliability prediction and management.

In conclusion, when BBNs are built for reliability prediction and reliability management, it must not be forgotten that the situation represented may change over time. The usability of the model thus depends on updating throughout the life of the project.

8.3.2. Benchmarking BBNs against FMEA as reliability management tool

In the reliability context the BBN application process has similarities with the application of FMEA (a useful reliability management tool, see section 2.3) as shown in Table 20.

Table 20: Comparison of BBNs and FMEA as reliability management tool

BBNs	FMEA
Identification of factors influencing reliability	Identifying the potential failure modes and the possible faults that are a result of these
Identification and probabilistic description of the relationships between the identified factors and reliability	Identify the RPN, consisting of the characteristics: - Probability that the fault will occur - Probability that the fault will be detected - Severity of the consequences of the fault
Calculations can be made in order to identify the effects of possible alterations of the values (MPDs) of the identified factors	Defining possible actions to reduce RPNs of high-ranked faults

The most important advantage of BBNs over FMEA in reliability management is the fact that BBNs quantify probabilistically the effect of altering the value of a certain influencing factor on reliability, whereas the effect of changing an RPN of a failure mode on reliability is not determined.

More importantly the input for a FMEA is a list of already identified potential failure modes. It has been shown in this thesis that the input for BBNs is not limited to failure modes, BBNs can also incorporate ‘soft’, non-physical, environmental factors, in a systems modelling approach. Because BBNs can handle a much wider range of factors than FMEA, the application of BBNs in the area of reliability management appears to be much broader than the application of FMEA.

BBNs make cause and effect relations explicit and are able to represent chains of causes and effects and can capture and describe interactions between different causes. In contrast, FMEA provides only information regarding single cause and effect relations (each line in a table). Consequently, FMEA does not give insight into possible chains of cause and effect, and is not able to include interactions between multiple causes.

While BBNs calculate the effect of changes in the values of influencing factors, they do not automatically provide a way of defining possible actions to obtain this goal. In the context of using BBNs for reliability management, this would be a valuable addition to the process of BBN application, when comparing it to FMEA.

8.4. Recommendations for Future Research

In the light of the generalization, the conclusions, and the research contribution presented in the previous chapter, and the reflections in the first two sections of this chapter, several directions for future research can be identified. These directions are discussed in this section.

Validation

In this study validation of the problem structure (see section 5.4), and functioning of the model (see section 7.6) was examined.

Validation of the problem structure took place through a survey. Because of the heterogeneity of the group of participants in the research, near 100% response rate was needed to obtain usable results. The low response rate, just 35%, prevented validation of the problem structure in this way.

A focus group provided validation of the functioning of the model, see section 7.6. The session revealed discrepancies between the model output and the participant’s expectations and views about the outputs.
The model may be adjusted in the light of the focus group results by applying another round of adapted GT to increase its validity. The validity may then be tested again using a further focus group meeting. The model can also be validated using real data, as they become available from the field. Data based validation cannot be used for validating the model for decision support early in development, and is only possible when the designed system is already in the field.

This thesis does not address external validation (transferability). The model built in the case study is, as shown in subsection 4.4.6, situation specific and not general. In the thesis only one study at a single site was carried out. Clearly, if more studies at other sites can be done, general results can be sought.

Implementation of BBNs for reliability prediction and management using a systems modelling approach

This thesis demonstrates a systems modelling approach to building BBNs for reliability prediction and for decision support. The validity of the model has been evaluated by a group of experts in a focus group. The extent to which the outcomes of Bayesian inference through these models may be used in practice is not yet researched. Use of Bayesian inference in practice would provide the opportunity to study the usability of the BBN as constructed in this thesis for providing reliability predictions and decision support.

Maintaining a record of reality while using the constructed BBNs would allow the values of these models in terms of reliability prediction and decision support to be determined. In this way it can be seen whether decisions may be supported in terms of ranking possible actions, or whether the implications of the BBN may go further than that. Such an analysis would give an indication of the extent to which the response of the model to changes in inputs reflects the real response. In particular the comparison would show sensitivity of the model to inputs similar to the response of the real system to the inputs.

The use of BBNs in practice might include the use of multiple criteria decision analysis (MCDA), as proposed by Fenton & Neil (2001) allowing more attributes than reliability alone to be taken into account.

Application of adapted GT at a higher level of detail

The adapted GT approach as applied in this research was used to identify and define the variables and was applied at a certain level of abstraction (based on interview reports, rather than full interview transcripts). This made variable identification easier but required additional interviews to define the variables and identify their state-space. The original GT approach might result in the availability of more detailed information for coding and analysis. This approach requires more effort for the identification of the variables, but permits the identification and definition of the state-space of these variables without additional interviews.

Probability elicitation

In the case study described in this thesis, the instantiation stage of the BBN construction process has been approached using direct probability elicitation. Although it was substantiated in chapter 6 that the strongest objections (experts having difficulties with providing probability estimates) against direct probability elicitation were not applicable in this case study, it remains that the use of direct elicitation for the purposes of probability elicitation is, in general, the worst way of probability elicitation (Cooke, 1991). Future research may aim at the identification of a probability elicitation method that is usable also for experts that are less familiar with the concept of probability.

Moreover, the limited availability of the experts for the purpose of probability elicitation should be taken into account. It may be interesting to see whether the work of Fenton et al., (2006), who use ranked nodes to model qualitative judgments (i.e.: judgments in terms of 'good – average – bad'), can be applied in the case in this thesis. The potential time saving of more than 80% in elicitation of the CPTs (Fenton et al., 2006), seems promising. However, an important obstacle for applying the method introduced by Fenton et al. (2006) is that all nodes have to be ranked nodes on an interval scale. Since the adapted GT approach leaves the definition of the state-space of the variables open, it does not require the state-space to be ranked. Therefore, the elicitation approach, presented in the paper by Fenton et al. (2006), is not automatically applicable in combination with the adapted GT approach.

Further use of the developed approach for BBN construction

The adapted GT approach for constructing BBNs based on expert knowledge seems to be applicable in different contexts and in different environments. The investigation of the extent of the applicability and its robustness against differences offers a potentially fruitful area of work.

The combination of the adapted GT approach with objective data remains to be studied. The analysis in this thesis has been essentially static leaving the development of the model through time as data becomes available to be investigated. In particular the incorporation of objective data within the soft systems methodology remains open. This counts both for model building and for evidence propagation.

References

Agena Ltd., 2009. www.AgenaRisk.co.uk.

Anderton, R.H., Checkland, P.B. 1977. *On learning our lessons*, Internal discussion paper, 2/77, Department of Systems, University of Lancaster.

Anjard, R.P., 1996. Process Mapping: One of Three, New, Special Quality Tools for Management, Quality and all other Professionals, *Microelectronics Reliability*, Vol. 36, No. 2, pp. 223-225.

Ayyub, B.M., Klir, G.J., 2006. *Uncertainty modelling and Analysis in Engineering and the Sciences*, Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton.

Babchuk, W.A., 1996. Glaser or Strauss? Grounded Theory and Adult Education, *Midwest Research-to-Practice Conference in Adult, Continuing and Community Education*.

Badoux, R.A.J., 1991, Introduction to Bayesian Methods in Reliability, In: Badoux, R, Sander, P. (Eds.), *Bayesian Methods in Reliability* (Kluwer Academic Publishers, Dordrecht.

Baecher, G.B., 1998. *Judgemental probability in risk assessment*, Technical report, prepared for The Office of the Chief, US Army Corps of Engineers.

Ballou, R.H., 1999. *Business Logistics Management*, 4th edition, Prentice-Hall, Inc., New Jersey.

Bedford, T., Cooke, R.M. 2001, *Probabilistic Risk Analysis Foundations and Methods*, Cambridge University Press, Cambridge.

Bedford, T., Quigley, J., Walls, L. 2006. Expert Elicitation for Reliable System Design, *Statistical Science,* Vol. 21, No. 4, pp. 428-450.

Biazzo, S., 2002. Process mapping techniques and organisational analysis Lessons from sociotechnical system theory, *Business Process Management Journal*, Vol. 8, No. 1, pp. 42-52.

Birolini, A. 2007. *Reliability Engineering: Theory and practice*, 5th edition, Springer, Berlin.

Blair, R.N., Whitston, C.W., 1971. *Elements of Industrial Systems Engineering*, Prentice Hall, Englewood Cliffs.

Blanchard, W.S., Fabrycky, W.J. 2006. *Systems Engineering and Analysis*, 4th edition, Prentice Hall, Upper Saddle River.

Blischke, W.R., Murthy, D.N.P. 2000. *Reliability: modelling, prediction, and optimization*, Wiley-Interscience, New York.

Booker, J.M., Bement, T.R., Meyer, M.A., Kerscher III, W.J., 2003. PREDICT: A New Approach To Product Development and Lifetime Assessment Using Information Integration Technology, In: C.R. Khattree and R. Rao (Eds.), *Handbook of statistics: statistics in industry*, Elsevier Science Direct, Amsterdam.

Borenstein, D. 1998. Towards a practical method to validate decision support systems, *Decision Support Systems*, Vol. 23, pp. 227-239.

Brombacher, A.C., Sander, P.C., Sonnemans, P.J.M., Rouvroye, J.L. 2005. Managing product reliability in business processes 'under pressure', *Reliability Engineering and System Safety,* Vol. 88, pp. 137-146.

Bryant, A., 2002. Re-grounding Grounded Theory, *The Journal of Information Technology Theory and Application (JITTA), Vol. 4, No. 1, pp. 25-42.*

Carrico, M.A., Girard, J.E., Jones, J.P. 1989. *Building Knowledge Systems Developing and Managing Rule-Based Applications*, McGraw-Hill Book Company, New York.

Chapman, C., Ward, S., 2000. Estimation and evaluation of uncertainty: a minimalist first pass approach. *International journal of project management*, Vol. 18, pp. 369-383.

Chapman, C., Ward, S. 2003. Constructively simple estimating: a project management example. *Journal of the Operational Research Society*, Vol. 54, pp. 1050-1058.

Charniak, E., 1991. Bayesian Networks without tears. *AI Magazine*, Vol. 12, No. 4, pp. 50-63.

Checkland, P.B. 1986. *Systems Thinking, Systems Practice*. Chichester, John Wiley & Sons Ltd.

Checkland, P.B., 1989. Soft Systems Methodology, in: Rosenhead, J. (ed.), *Rational Analysis for a Problematic World*, John Wiley & Sons, Ltd., Chichester.

Checkland, P.B. & Scholes, J. 1990. *Soft Systems Methodology in Action*. Chichester: John Wiley & Sons Ltd.

Checkland, P.B. & Holwell, S. 1998. *Information, Systems and Information Systems making sense of the field*. Chichester, John Wiley & Sons Ltd.

Checkland, P. 2000. Soft Systems Methodology: A Thirty Year Retrospective, *Systems Research and Behavioral Science*, Vol. 12, pp. 12-58.

Cheng, J., Greiner, R., Kelly, J., Bell, D., Liu, W., 2002. Learning Bayesian networks from data: An information-theory based approach, *Artificial Intelligence*, Vol. 137, pp. 43-90.

Clandinin, D.J., Connelly, F.M., 2000. *Narrative inquiry: Experience and story in qualitative research*, Jossey-Bass, San Fransisco.

Clarke, G.M., Kempson, R.E., 1997. *Introduction to the design & analysis of experiment*s. Arnold, London.

Clemen, R.T., Winkler, R.L., 1999. Combining Probability Distributions from Experts in Risk Analysis, *Risk Analysis,* Vol. 19, No. 2, pp. 187-203.

Cohen, M.A., Eliashberg, J., Ho, T-H., 1996. New Product Development: The Performance and Time-to-Market Tradeoff, *Management Science*, Vol. 42, No. 2, pp. 173-186.

Condra, L.W., 2001. *Reliability improvement with design of experiments*, 2nd edition. Dekker, Basel.

Cooke, R.M. 1991. *Experts in uncertainty: opinions and subjective probability in science*, Oxford University Press, Oxford.

Cooke, N.J., 1994. Varieties of knowledge elicitation techniques, International Journal of Human-Computer Studies 41, pp. 801-849.

Cooke, R.M., Goossens, L.H.J., 2000. Procedures Guide For Structured Expert Judgement in Accident Consequence Modelling, *Radiation Protection Dosimetry,* Vol. 90, No.3, pp. 202-209.

Cooke, R.M., Goossens, L.H.J., 2008. TU Delft expert judgment data base, *Reliability Engineering and System Safety*, Vol. 93, No. 5, pp. 657-674.

Cooper, R.G., 2001. *Winning at new products: accelerating the process from idea to launch*, 3rd edition. Perseus, Cambridge.

Cooper, G.E., Herskovits, E., 1992. A Bayesian Method for the Induction of Probabilistic Networks from Data, *Machine Learning* 9, pp. 309-347.

Corbin, J., Strauss, A.L., 1990. Grounded Theory Research: Procedures Canons and Evaluative Criteria, *Qualitative Sociology* 13 (1), pp. 3-21.

Cowell, R.G., Dawid, P.A., Lauritzen, S.L. 1999. *Probabilistic Networks and expert systems*, Springer, Berlin.

Creswell, J.W., 1994. *Research Design Qualitative and Quantitative Approaches*, SAGE Publications, Thousand Oaks.

Creswell, J.W., 1998. *Qualitative inquiry and research design*, SAGE Publications, Thousand Oaks.

Creswell, J.W., 2003. *Research Design Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd edition, SAGE Publications, Thousand Oaks.

D'Ambrosio 1999. Inference in Bayesian Networks, *AI Magazine*, Vol. 20, No. 2, pp. 21-36.

Davis, T.P., 1998. The fallacy of reliability prediction in automotive engineering. *Automotive Excellence,* Spring 1998, pp. 19-21. Published by the American Society for Quality (also available from http://www.timdavis.co.uk/predict.PDF)

Dean, A., Voss, D. 1999. *Design and analysis of experiments*, Springer, Berlin.

Den Ouden, P.H. 2006. *Development of a design analysis model for consumer complaints: revealing a new class of quality failures*, Technische Universiteit Eindhoven, Eindhoven.

Druzdzel, M.J., Van der Gaag, L.C. 2000. Building Probabilistic Networks: "Where do the numbers come from?" Guest Editors' Introduction, IEEE Transactions on Knowledge and Data Engineering, Vol. 12, No. 4, pp. 481-486.

Duane, J.T., 1964. Learning curve approach to reliability monitoring, *IEEE Transactions on Aerospace*, Vol. 2, no. 2, pp. 563-566.

Easterby-Smith, M., Thorpe, R., Lowe, A., 1991. *Management Research – An Introduction*, SAGE Publications Ltd., London.

Eden, C., 1988. Cognitive mapping – Invited review, *European Journal of Operational Research*, No. 36, pp. 1-13.

Eden, C., 2004. Analyzing cognitive maps to help structure issues or problems, *European Journal of Operational Research*, No. 159, pp. 673-686.

Eden, C., Simpson, P., 1989. SODA and cognitive mapping in practice, in: Rosenhead, J. (ed.), *Rational Analysis for a Problematic World*, John Wiley & Sons, Ltd., Chichester.

Eom, S.B., Lee, S.M., Kim, E.B., Somarajan, C., 1998. A Survey of Decision Support System Applications (1988-1994), *The Journal of the Operational Research Society*, Vol. 49, No. 2, pp. 109-120.

Fenton, N., Neil, M., 2001. Making decisions: using Bayesian nets and MCDA, *Knowledge-based systems*, Vol. 14, pp. 307-325.

Fenton, N., Neil, M., Caballero, J.G., 2007. Using Ranked Nodes to Model Qualitative Judgments in Bayesian Networks, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 10, pp. 1420-1432.

Fenton, N., Marsh, W., Neil, M., Cates, P., Forey, S., Tailor, T., 2004. Making Resource Decisions for Software Projects, *Proceedings 26th International Conference on Software Engineering*, IEEE Computer Society, Edinburgh, United Kingdom. ISBN 0-7695-2163-0, pp. 397-406.

Fenton, N., Neil, M., Marsh, W., Hearty, P., Marquez, D., Krause, P., Mishra, R. 2007. Predicting software defects in varying development lifecycles using Bayesian nets, *Information and Software Technology,* Vol. 49, pp. 32-43.

Finlay, P., 1995. *Introducing Decision Support Systems*, reprint, NCC Blackwell Ltd., Manchester.

Finlay, P.N., Forsey, G.J., Wilson, J.M. 1988. The Validation of Expert Systems – Contrasts with Traditional Methods, *The Journal of the Operational Research Society*, Vol. 39, No. 10, pp. 933-938.

Flood, R.L., Carson, E.R., 1990. *Dealing with complexity – An introduction to the Theory and Application of Systems Science*. Plenum Press, New York.

Flood, R.L., Jackson, M.C. 1991. *Creative Problem Solving – Total Systems Intervention*. John Wiley & Sons Ltd, Chichester.

Gandy, A., Jäger, P., Bertsche, B., Jensen, U., 2007. Decision support in early development phases – A case study from machine engineering. *Reliability Engineering and System Safety*, Vol. 92, pp. 921-929.

Gass, S.I. 1983. Decision-Aiding Models: Validation, Assessment, and Related Issues for Policy Analysis, *Operations Research*, Vol. 31, No.4, pp. 603-631.

Gilchrist, W., 1993. Modelling Failure Modes and Effects Analysis, *International Journal of Quality and Reliability Management*, Vol. 10, No. 5, pp. 16-23.

Goel, A., Graves, R.J. 2006. Electronic System Reliability: Collating Prediction Models, *IEEE Transactions on Device and Materials Reliability*, Vol. 6, No.2, pp. 258-265.

Gorry, G.A., Morton, M.S.S., 1971. A Framework for Management Information Systems, *Sloan Management Review*, Vol. 13, No. 1, pp. 55-70.

Hall, W.A., Callery, P., 2001. Enhancing the Rigor of Grounded Theory: Incorporating Reflexivity and Relationality, *Qualitative Health Research*, Vol. 11, No. 2, pp. 257-272.

Healy, J.D., Jain, A.K., Bennett, J.M., 1997. "Reliability Prediction" *tutorial notes, 1997 Annual Reliability and Maintainability Symposium.*

Heckerman, D. 1996. *A Tutorial on Learning with Bayesian Networks*. Technical Report # MSR-TR-95-06, Microsoft Research, Washington.

Hodge, R., Evans, M., Marshall, J., Quigley, J., Walls, L., 2001. Eliciting Engineering Knowledge about Reliability during Design – Lessons learnt from Implementation. *Quality and Reliability Engineering International* 17, pp.169-179.

Hora, S.C., 1996. Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management, *Reliability Engineering and System Safety*, Vol. 54 (Nov.), pp. 217-223.

Hora, S.C., Von Winterfeldt, D., 1997. Nuclear Waste and Future Societies: a Look into the Deep Future, *Technological Forecasting and Social Change,* Vol. 56, pp. 155-170.

Houben, M.J.H.A., Sonnemans, P.J.M., Newby, M.J., 2009. Uncertainty in reliability estimation: when do we know everything we know? In: Bris, R., Guedes Soares, C., Martorell, S. (Eds.), *Reliability, Risk and Safety: Theory and Applications,* Taylor & Francis group, London, pp. 61-64. ISBN: 978-0-415-55509-8.

Houben, M.J.H.A., Sonnemans, P.J.M., Stollman, G.M., Newby, M.J., 2009, Reliability Prediction Through Bayesian Inference Based On Product Change. *Proceedings Fifty-fifth Annual Reliability and Maintainability Symposium*, paper 9D3, Fort Worth, TX, USA. ISBN: 978-1-4244-2509-9.

Howard, R.A., Matheson, J.E., 2005. Influence Diagrams, *Decision Analysis*, Vol. 2, No. 3, pp. 127-143.

IEC 60300 series, International Electrotechnical Commission, Geneva.

Jackson, M.C. 1991, *Systems methodology for the management sciences*. Plenum Press, New York.

Jacobs, J.C., Van Moll, J.H., 2007. *Effects of virtual product development on product quality and their influencing factors*, Technische Universiteit Eindhoven, Eindhoven.

Jenkinson, D., 2005. *The Elicitation of Probabilities – A Review of the Statistical Literature.* Available online: http://www.shef.ac.uk/content/1/c6/03/09/33/review3.pdf. Visited: 30 June, 2009.

Jensen, F.V., 2001. *Bayesian Networks and Decision Graphs*, Springer-Verlag, New York, Inc.

Jones, J., Marshall, J., Aulak, G., Newman, B., 2003. Development of an Expert System for Reliability Task Planning as Part of the REMM Methodology, *Proceedings Annual Reliability and Maintainability Symposium.*

Keeney, R.L., Von Winterfeldt, D., 1991. Eliciting Probabilities from Experts in Complex Technical Problems, *IEEE Transactions on Engineering Management*, Vol. 38, No. 3, pp. 191-201.

Kerscher III, W.J., Booker, J.M., Meyer, M.A., 2001. PREDICT: A Case Study, *Proceedings of the European Conference on Safety and Reliability (ESREL).*

Kerscher III, W.J., Booker, J.M., Meyer, M.A., Smith, R.E., 2003. PREDICT: A Case Study, Using Fuzzy Logic, *Proceedings Annual Reliability and Maintainability Symposium.*

Kjaerulff, U.B., Madsen, A.L., 2008. *Bayesian networks and influence diagrams: a guide to construction and analysis*, Springer, New York.

Kolarik, W.J. 1999. *Creating Quality: Process Design for Results*, McGraw-Hill, New York.

Krueger, R.A., Casey, M.A., 2009. *Focus groups: A practical guide for applied research*, 4th edition, SAGE Publications, London.

Kynn, M., 2008. The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society: Series A,* Vol. 171, Part 1, pp. 239-264.

Langseth, H., Portinale, L. 2007. Bayesian networks in reliability, *Reliability Engineering and System Safety,* Vol. 92, pp. 92-108.

Larrañaga, P., Kuijpers, C.M.H., Murga, M.H., Yurramendi, Y., 1996. Learning Bayesian Network Structures by Searching for the Best Ordering with Genetic Algorithms. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* 26 (4), pp. 487-493.

Lecompte, M.D., Schensul, J.J. 1999. *Designing & Conducting Ethnographic Research*, Altamira press, Walnut Creek.

Leslie, L.L., 1972. Are High Response Rates Essential to Valid Surveys?, *Social Science Research,* No. 1, pp. 323-334.

Lewis, E.E. 1996. *Introduction to reliability engineering,* 2nd edition. Wiley, New York.

Lindley, D., 1975. *Making decisions*, Wiley-Interscience, London.

Locke, K., 2001. *Grounded Theory in Management Research*, SAGE Publications Ltd, London.

Lu, Y. 2002. *Analysing Reliability Problems in Concurrent Fast Product Development Processes*, Technische Universiteit Eindhoven, Eindhoven.

Mackenzie, A., Pidd, M., Rooksby, J., Sommerville, I., Warren, I., Westcombe, M., 2006. *Wisdom,* decision support and paradigms for decision making, *European Journal of Operational Research*, No. 170, pp. 156-171.

Magniez, C.A.A., 2007. *Combining information flow and physics-of-failure in mechatronic products*, Technische Universiteit Eindhoven, Eindhoven.

Mansourian, Y., 2006. Adoption of grounded theory in LIS research, *New Library World*, Vol. 107, No. 1228/1229, pp. 386-402.

Marinai, L., Probert, D., Singh, R. 2004. Prospects for Aero gas-turbine diagnostics: a review, *Applied Energy*, Vol.79, Iss. 1, pp. 109-126.

Marquez, D., Neil, M., Fenton, N. 2007. A new Bayesian Network approach to Reliability modelling, *Proceedings Mathematical Methods in Reliability.*

Marshall, J., Walls, L., Jones, J., 2002. Enhancing Product Reliability using REMM, *Proceedings Annual Reliability and Maintainability Symposium.*

Marshall, J., Balderstone, M., Davies, J., Lumbard, D., 2003. REMM Evaluation Process, *Proceedings Annual Reliability and Maintainability Symposium.*

Marshall, J., Jones, J., 2006. Practical Evaluation of the REMM Process, *Proceedings Annual Reliability and Maintainability Symposium.*

McCall, Jr. M. W., Lombardo, M. M. 1982. Using Simulation for Leadership and Management Research: Through the Looking Glass, *Management Science,* Vol. 28, No. 5, pp. 533-549.

McDermott, R.E., Mikulak, R.J., Beauregard, M.R., 2009. *The basics of FMEA*, 2nd edition, Productivity Press, Taylor & Francis Group, New York.

McGraw, K.L., Harbison-Briggs, K. 1989. *Knowledge Acquisition Principles and Guidelines*, Prentice-Hall International, Inc., Englewood Cliffs.

Meeker, W.Q., Escobar, L.A. 2004. Reliability: The Other Dimension of Quality, *Quality Technology & Qualitative Management,* Vol. 1, No. 1, pp. 1-25.

Meeker, W.Q., Hamada, M. 1995. Statistical Tools for the Rapid Development & Evaluation of High-Reliability Products, *IEEE Transactions on Reliability*, Vol. 44, No. 2, pp. 187-198.

Merriam, S.B., 1998. Qualitative research and case study applications in education, Jossey-Bass, San Francisco.

Meyer, M.A., Booker, J.M., 1991. *Eliciting and Analyzing Expert Judgment – A Practical Guide*. Academic Press Limited, London.

Miles, M.B., Huberman, A.M., 1994. *Qualitative Data Analysis – An Expanded Sourcebook*, 2nd edition, SAGE Publications Inc., Thousand Oaks.

Mingers, J. 2006. *Realising Systems Thinking Knowledge and Action in Management Science*, Springer Science + Business Media, Inc., New York.

Mitchell, G., 1993. *The practice of operational research*, Wiley, Chichester.

Montgomery, D.C. 2009. Design and analysis of experiments, 7th edition. Wiley, Hoboken.

Montgomery, D.C., Runger, G.C., 2002. *Applied Statistics and Probability for Engineers*, 3rd edition, John Wiley & Sons, Inc., New York.

Morgan, M.G., Henrion, M., 1990. *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*, Cambridge University Press, Cambridge.

Moustakas, C., 1994. *Phenomenological research methods*, SAGE Publications, Thousand Oaks.

Munro, I., Mingers, J. 2002. The use of multimethodology in practice – results of a survey of practitioners, *Journal of the Operations research society,* Vol. 59, No. 4, pp. 369-378.

Murthy, D.N.P., Rausand, M., Østeras, T. 2008. *Product reliability: specification and performance,* Springer, Berlin.

Nadkarni, S., Shenoy, P.P., 2004. A Causal Mapping Approach to Constructing Bayesian Networks, *Decision Support Systems* 38, pp. 259-281.

Neapolitan, R.E. 2004. *Learning Bayesian Networks*. Pearson/Prentice Hall, London.

Neil, M., Fenton, N., Forey, S., Harris, R. 2001. Using Bayesian belief networks to predict the reliability of military vehicles, *Computing and Control Engineering Journal,* (Feb.) pp. 11-20.

Neil, M., Fenton, N., Forey, S., Harris, R., 2003.Assessing Global Vehicle Reliability using Bayesian Networks, In: Strutt, J.E., Hall, P.L. (Eds.), Global vehicle reliability: prediction and optimization techniques, The Cromwell Press, Trowbridge.

Neil, M., Fenton, N., Nielson, L. 2000. Building large-scale Bayesian networks, *The Knowledge Engineering Review*, Vol. 15, No. 3, pp. 257-284.

Neil, M. Fenton, N., Tailor, M. 2005. Using Bayesian Networks to Model Expected and Unexpected Losses, *Risk Analysis*, Vol. 25, No. 4, pp. 963-972.

Nieswiadomy, R.M., 1993. *Foundations of Nursing Research*, 2nd edition, Appleton & Lange, Norwalk.

Norrington, L., Quigley, J., Russell, A., Meer, v.d., R., 2008. Modelling the reliability of search and rescue operations with Bayesian Belief Networks, *Reliability Engineering and System Safety* 93, pp. 940-949.

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F., Ferson, S., 2004. Challenge problems: uncertainty in systems response given uncertain parameters, *Reliability Engineering and System Safety*, Vol. 85, pp. 11-19.

O'Hagan, A., Forster. J., 2004. *Bayesian inference*, 2nd edition, Arnold, London.

O'Leary, T.J., Goul, M., Moffitt, K.E., Radwan, A.E. 1990. Validating Expert Systems, *IEEE Intelligent Systems*, Vol. 5, No. 3, pp. 51-58.

Oniśko, A. 2008. Medical Diagnosis. In: Pourret, O., Naïm, P., Marcot, B. (Eds.) *Bayesian Networks A Practical Guide to Applications*, John Wiley& Sons Ltd, Chichester, pp. 15-32.

Orlikowski, W.J., 1993. CASE Tools as Organisational Change: Investigating Incremental and Radical Changes in Systems Development, *MIS Quarterly*, Vol. 17, No. 3, pp. 309-340.

Ormon, S.W., Cassady, C.R., Greenwood, A.G. 2002. Reliability Prediction Models to Support Conceptual Design. *IEEE Transactions on Reliability*, Vol. 51, No. 2, pp. 151-157.

Østeras, T., Rausand, M., Murthy, D.N.P., 2008. Reliability specification in new product development, *International Journal of Product Development,* Vol. 5, Nos. 1/2, pp. 17-38.

Partington, D., 2000. Building Grounded Theories of Management Action, *British Journal of Management*, Vol. 11, pp. 91-102.

Patton, M.Q., 1990. *Qualitative evaluation and research methods*, 2nd edition, SAGE Publications, London.

Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo.

Pearl, J., 2000. *Causality: models, reasoning and inference*, Cambridge University Press, Cambridge.

Petkova, V.T., 2003. *An analysis of field feedback in consumer electronics industry*, Technische Universiteit Eindhoven, Eindhoven.

Ragatz, G.L., Handfield, R.B., Scannell, T.V., 1997. Success Factors for Integrating Suppliers into New Product Development, *Journal of Product Innovation Management*, Vol. 14, pp. 190-202.

Rausand, M., Høyland, A. 2004. *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd edition, John Wiley & Sons, Inc., Hoboken.

Renooij, S. 2001. Probability elicitation for belief networks: issues to consider, *The Knowledge Engineering Review*, Vol. 16, No. 3, pp. 255-269.

Renooij, S., Witteman, C., 1999. Talking probabilities: communicating probabilistic information with words and numbers, *International Journal of Approximate Reasoning* 22, pp. 169-194.

RIAC-HDBK-217Plus, 2006. *Handbook of 217Plus™ reliability prediction models*.

Robrecht, L.C., 1995. Grounded Theory: Evolving methods, *Qualitative Health Research*, Vol. 5, No. 2, pp. 169-178.

Rupp, O. 2002. Hannover, GmbH, Langenhagen, Germany, MTU Maintenance.

Sarker, S., Lau, F., Sahay, S., 2001. Using an Adapted Grounded Theory Approach for Inductive Theory Building About Virtual Team Development, *The DATA Base for Advances in Information Systems*, Vol. 32, No. 1, pp. 38-56.

Schmitt, S.A., 1969. *Measuring uncertainty: an elementary introduction to Bayesian statistics*, Addison-Wesley, London.

Shachter, R.D., 1988. Probabilistic Inference and Influence Diagrams, *Operations Research*, Vol. 36, No. 4, pp. 589-604.

Sigurdsson, J.H., Walls, L.A., Quigley, J.L. 2001. Bayesian Belief Nets for Managing Expert Judgement and Modelling Reliability, *Quality and Reliability Engineering International*, Vol. 17, pp. 181-190.

Singh, M., Valtorta, M., 1995. Construction of Bayesian Network Structures From Data: A Brief Survey and Efficient Algorithm. *International Journal of Approximate Reasoning* 12, pp. 111-131.

Smith, A.F.M., 1991, An Overview of the Bayesian Approach, in: *Bayesian Methods in Reliability* (eds.: Badoux, R, Sander, P.), Kluwer Academic Publishers, Dordrecht.

Soliman, F., 1998. Optimum level of process mapping and least cost business process re-engineering, *International Journal of Operations & Production Management*, Vol. 18, No. 9/10, pp. 810-816.

Spetzler, C.S., Staël Von Holstein, C-A, S., 1975. Probability Encoding in Decision Analysis, *Management Science,* Vol. 22, No. 3, pp. 340-358.

Stake, R.E., 1995. *The Art of Case Study*, SAGE Publications, Thousand Oaks.

Stamatis, D.H., 2003. *Failure mode and effect analysis: FMEA from theory to execution*, 2nd edition, ASQ Quality Press, Milwaukee.

Strauss, A.L., Corbin, J., 1990. *Basics of qualitative research: grounded theory procedures and techniques*, SAGE Publications, Newbury Park.

Strauss, A.L., Corbin, J., 1998. *Basics of qualitative research: grounded theory procedures and techniques*, 2nd edition, SAGE Publications, Newbury Park.

Studer, R., Benjamins, V.R., Fensel, D., 1998. Knowledge Acquisition: Principles and Guidelines, *Data & Knowledge Engineering* 25, pp. 161-197.

Teng, S. (G.), Ho, S. (M.), 1996. Failure mode and effects analysis, *International Journal of Quality & Reliability Management*, Vol. 13, No. 5, pp. 8-26.

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Bias, *Science,* Vol. 185, No. 4157, pp. 1124-1131.

US MIL-HDBK-217F 1992. *US-MIL-HDBK-217F: Military Handbook Reliability Prediction of Electronic Equipment*, National Technical Information Service, Springfield.

Van der Gaag, L.C., 1996. Bayesian Belief Networks: Odds and Ends, *The Computer Journal*, Vol. 39, No. 2, pp. 97-113.

Van der Gaag, L.C., Renooij, S., Witteman, C.L.M., Aleman, B.M.P., Taal, B.G. 2002. Probabilities for a probabilistic network: a case study in oesophageal cancer, *Artificial Intelligence in Medicine,* Vol. 25, pp. 123-148.

Verschuren, P., Doorewaard, H. 1999. *Designing a Research Project*, Uitgeverij Lemma, Utrecht. Dutch original: Het ontwerpen van een onderzoek (2nd edition, 1998).

Visée, J., 2009. *An adequate method to quantify probabilities : case studies in the context of early reliability prediction using Bayesian networks restricted to subjective data*, Technische Universiteit Eindhoven, Eindhoven.

Von Mises, R., 1939. *Probability, Statistics and Truth*. Hodge, London, United Kingdom. German original 1928.

Walls, L., Quigley, J. 2001. Building prior distributions to support Bayesian reliability growth modelling using expert judgement, *Reliability Engineering and System Safety*, Vol. 74, pp. 117-128.

Walls, L., Quigley, J., Marshall, J. 2006. Modeling to Support Reliability Enhancement During Product Development With Applications in the U.K. Aerospace Industry, *IEEE Transactions on Engineering Management*, Vol. 53, No. 2, pp. 263-274.

Wheelwright, S.C., Clark, K.B. 1992. *Revolutionizing product development: Quantum leaps in speed, efficiency and quality*, Free press, New York.

Winkler, R.L. 1996. Uncertainty in probabilistic risk assessment, *Reliability Engineering and Systems Safety,* Vol. 54, pp. 127-132.

Winkler, R.L., Clemen, R.T., 1999. Combining Probability Distributions From Experts In Risk Analysis, *Risk Analysis* 19 (2), pp. 187-203.

Winkler, R.L., Clemen, R.T., 2004. Multiple Experts vs. Multiple Methods: Combining Correlation Assessments, *Decision Analysis,* Vol. 1, No. 3, pp. 167-176.

Wise, R., Baumgartner, P. 1999. Go Downstream: The New Profit Imperative in Manufacturing, *Harvard Business Review*, Vol. 77, pp. 133-141.

Witteman, C., Renooij, S. 2003. Evaluation of a verbal-numerical probability scale, *Journal of Approximate Reasoning 33*, pp. 117-131.

Yadav, O.P., Singh, N., Goel, P.S., Itabashi-Campbell, R., 2003. A Framework for Reliability Prediction During Product Development Process Incorporating Engineering Judgments, *Quality Engineering*, Vol. 15, No. 4, pp. 649-662.

Yin, R.K., 2003. *Case study research design and methods*. Third edition, Sage Publications, Inc. Thousand Oaks.

Zhang, X., Pham, H., 2000. An analysis of factors affecting software reliability, *The Journal of Systems and Software*, No. 50, pp. 43-56.

Zitrou, A. 2006. *Exploring a Bayesian Approach for Structural Modelling of Common Cause Failures*, University of Strathclyde, Strathclyde.

Appendices

A. BBNs: Formalism

BBNs are a graphical model representing probabilistic relationships between variables (Heckerman, 1996). The graphical model consists of nodes that represent variables with uncertain values (i.e. stochastic variables), and directed arcs, that represent the relationships (causal or influential (Van der Gaag, 1996)) between the variables. Absence of a connection between two nodes implies (conditional) independence. In this section, the concept of BBNs is elaborated using an illustrative example model. For a more mathematical discussion of the subject of BBNs, see e.g. Pearl (1988); Cowell et al. (1999); Pearl (2000).

The example in this appendix (inspired on the 'Bayesian nets and probability tutorial', see http://www.dcs.qmw.ac.uk/~norman/BBNs/BBNs.htm, but adapted for use in this thesis) is used to describe the different elements of the model:

Jack and Thomas work for the same company. However, Jack and Thomas live in different places. Since Jack lives nearby a railway station, he goes to work either by train or by car. Thomas however does not live in the neighbourhood of a railway station and always goes to work by car. Whereas Jack always comes out of bed early, Thomas sometimes oversleeps. Hence, the time of arrival at work for Jack– if he travels by train – depends on whether the train is on time or is delayed (e.g. because of a signal disturbance) or – if he travels by car – whether there are problems in traffic (e.g. traffic jam) On the other hand, the time of arrival at work for Thomas, depends on possible problems in traffic and whether or not he oversleeps.

From this example, the arrival at work of Thomas and Jack can be modelled using a BBN. Hence, the model contains elements that affect their arrival at work. For Thomas, this is his oversleeping, or there being a traffic jam. For Jack, this is there being a train delay, or there being a traffic jam. Furthermore, since a train delay can be caused by a signal disturbance or a derailed train, these two factors can be included in the model as well. The resulting model is shown in Figure 24.

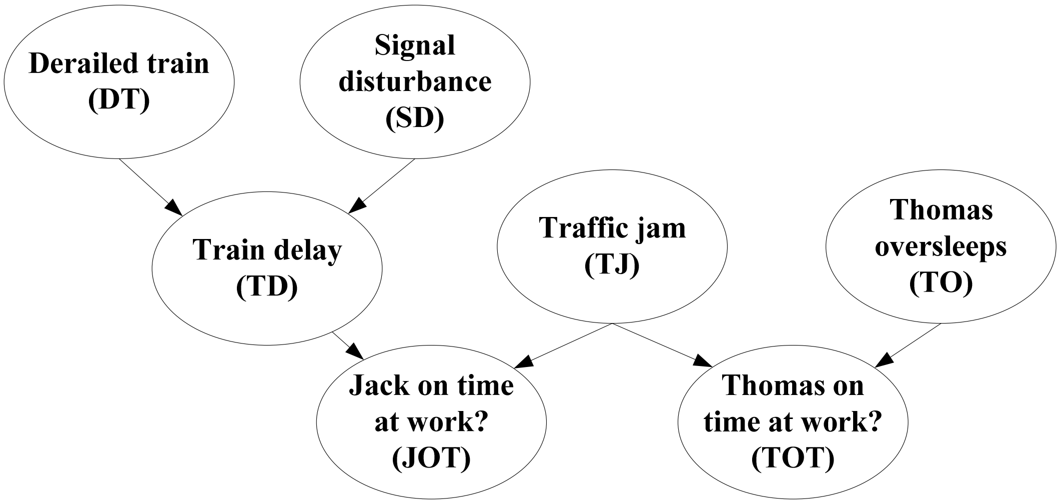


Figure 24: Example BBN

Variables

Random variables that are represented in the model are represented by nodes. Nodes that stand at the beginning of an arc are called ‘parent nodes’, whereas nodes that stand at the end of an arc are called ‘child nodes’. Nodes that have no parents themselves are called ‘root nodes’ (e.g. ‘signal disturbance’); nodes that have no children themselves are called ‘leaf nodes’ (e.g. ‘Jack on time at work?’) (Sigurdsson et al., 2001). The ‘root nodes’ often represent input nodes for the network, since they provide input, rather than being related to the way in which the BBN functions. Variables are described by their name, and can assume certain potential values (states). These values can be different in nature, being e.g. continuous numerical, discrete numerical, categorical, dichotomous, etc. (Charniak, 1991) All states together, make up the state space of a variable. The uncertainty regarding the actual state of this variable is depicted using probability theory.
As example, the root node variable ‘Signal disturbance’ is taken. For this variable, 2 dichotomous states are defined: ‘Yes’” (there is a signal disturbance) and ‘No’ (there is no signal disturbance)” (Note that in the example model, all variables are defined to have a discrete nature, either being labelled variables, or dichotomous). The uncertainty regarding the state of this variable is expressed using probabilities. As an example, there may be a probability of 0.9 that there is no signal disturbance (at the same time implying that there is a probability of 0.1 of the signal being disturbed). The probabilities for the different states of ‘signal disturbance’ (p(SD)) probability table that is related to the root node signal disturbance as shown in Table 21. Together, the probabilities form a discrete MPD: ‘Yes – No’: ‘0.10 – 0.90’.

Table 21: NPT for the root node SD

Signal disturbance (SD)	
Yes	*No*
0.10	0.90

Kjaerulff and Madsen (2008) identify two important points of attention that have to be taken into account regarding the variables that are part of the BBN.

- Regarding the definition: the variables have to be well-defined.
- Regarding the range of potential values of the variables (state-space): the states of the variables have to be unambiguous and mutually exclusive, and the state-space has to be exhaustive.

In the case of the variable ‘signal disturbance’, these points are addressed as follows:

- The definition of the variable is: the presence of a signal disturbance on the section that Jack has to travel by train.
- The range of potential values is: yes, no. Their meaning is clear (unambiguous), and there is no overlap in their meaning (mutually exclusive). Furthermore, the full range of possible values for this variable is covered by these two potential values, ensuring exhaustiveness of the state-space.

Because of the fact that each variable has to have a state space, and because of the exhaustiveness property, the total sum of the probabilities, for all states represented in the probability tables, is 1. Hence, the value of every variable is determined by a probability distribution.

Relationships

Conditional Probability Tables (CPTs) represent the uncertain relationship of a child node with its parent node(s). It is also possible to represent an uncertain relationship between variables using a pdf that is connected to the child node (Neil et al., 2000).

An example of such a CPT is related to the child node 'Train delay' (p('Train delay' |'Signal disturbance', 'Derailed train')), and is shown in Table 22 (in terms of mathematical expressions) and Table 23 (in terms of probabilities) In these tables, the columns represent the different possible states of the child node, whereas the rows represent the different possible states of the parent nodes (the different possible conditions).

Table 22: CPT for the child node TD

Signal Disturbance (SD)	**Derailed Train (DT)**	**Train delay (TD)**	
		Yes	*No*
Yes	*Yes*	p('Train delay'='Yes '| 'Signal disturbance'='Yes', 'Derailed train'='Yes')	p('Train delay'='No '| 'Signal disturbance'='Yes', 'Derailed train'='Yes')
	No	p('Train delay'='Yes '| 'Signal disturbance'='Yes', 'Derailed train'='No')	p('Train delay'='No '| 'Signal disturbance'='Yes', 'Derailed train'='No')
No	*Yes*	p('Train delay'='Yes '| 'Signal disturbance'='No', 'Derailed train'='Yes')	p('Train delay'='No '| 'Signal disturbance'='No', 'Derailed train'='No')
	No	p('Train delay'='Yes '| 'Signal disturbance'='No', 'Derailed train'='No')	p('Train delay'='No '| 'Signal disturbance'='No', 'Derailed train'='No')

Table 23: CPT for the child node TD filled in

Signal disturbance (SD)	**Derailed train (DT)**	**Train delay (TD)**	
		Yes	*No*
Yes	*Yes*	0.95	0.05
	No	**0.60**	0.40
No	*Yes*	0.80	0.20
	No	0.05	0.95

From Table 23, the conditional probabilities related to p('Train delay' |'Signal disturbance', 'Derailed train') can be determined. As an example, in Table 23, the conditional probability of the train being delayed p('Train delay'='Yes ') given that there is a signal disturbance p('Train delay'='Yes '|'Signal disturbance'='Yes'), but there is no derailed train, p('Train delay'='Yes '|'Signal disturbance'='Yes', 'Derailed train'='No') is 0.60. This probability is represented bold and underlined in Table 23. Because of the fact that the state-space of the nodes has to be exhaustive, and the child node has to have a value (has to be in some state), the sum of the probabilities in every row of the CPT has to be 1. As an example: in the same case, where there is a signal disturbance, but no derailed train, either there is a train delay, or there is no train delay. However, one of the statements has to be true. Hence, the sum of both probabilities has to be 1.

It may be strange to notice that, even when there is a signal disturbance, and there is a derailed train, there still is a chance that there is no train delay. However, this might be the case, .e.g. if the derailed train can be put back on the tracks in a short period of time, and the signal disturbance is only temporary. p('Train delay'='Yes'|'Signal disturbance'='Yes', 'Derailed train'='Yes') is therefore not 1, but close to 1.

Using Bayes' theorem, probabilistic information (like e.g. in Table 21 and Table 23) can be combined to infer new information. The general form of Bayes' theorem is as follows (Jensen, 2001):

$$P(A|B)=\frac{P(B|A)\cdot P(A)}{P(B)}$$

Extended with one extra conditioned variable, it becomes:

$$P(A|B,C)=\frac{P(B|A,C)\cdot P(A|C)}{P(B|C)}$$

Using Bayes' formula, it is possible to enter "evidence" regarding a node into the network, and use Bayesian inference to propagate the effect of this on the MPDs of other nodes in the network. This can be done using either forward reasoning (reasoning from cause to consequence) or backward reasoning (reasoning from consequence to cause).

Dependency types and propagation

From the CPT for the node 'Train delay' itself, there is no possibility to determine what the probability is that the train is delayed. This is because the real values of the variables 'Derailed train' and 'Signal disturbance' are not (yet) known: there is no evidence regarding either of the variables. Only if the values of 'Derailed train' and 'Signal disturbance' are known, one can also determine the values of the variable 'Train delay'. The dependency of 'train delay' on the variables 'derailed train' and 'signal disturbance' is a so-called converging d-connection, which is one of three possible types of dependence connections (d-connections) (Neil et al., 2000). All three types of d-connections will be discussed in this subsection. Note: for all types of d-connections, examples are taken from Figure 24.

Converging d-connection

Typically, in a converging d-connection, one consequence (child) node is directly affected by two (or more) causal (parent) nodes. In the case of Figure 25, 'Train delay' (being the consequence node) is conditionally dependent on 'Derailed train' and 'Signal disturbance'. Evidence at either of these two nodes will therefore "update" the node 'train delay'. Also, *if* there is already evidence at node 'train delay', then evidence for one of the two remaining nodes will lead to updating of the other: 'signal disturbance' and 'derailed train' are conditionally dependent on each other, *given* evidence for 'train delay'.

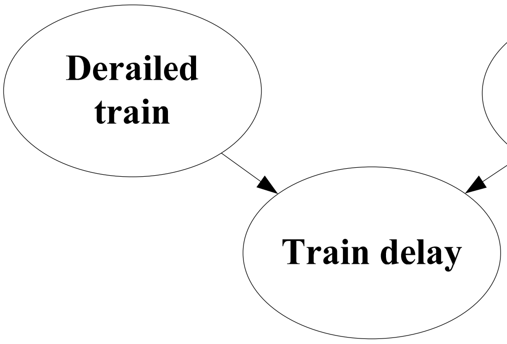


Figure 25: converging d-connection

As an example, the CPT represented in Table 23 can be used. If evidence regarding the variables 'derailed train' (p(DT)) and 'signal disturbance' (p(SD)) would be entered, then the marginal distribution for 'train delay' (p(TD)) can be determined.

Using Table 23 we can infer that *if* there is no derailed train on the tracks (evidence), and *if* there is no signal disturbance (evidence), there is a 95% likelihood of the train not being delayed. However, *if* there is no derailed train on the track (evidence) *and* it becomes known that there is a signal disturbance (new evidence), the likelihood of the train being delayed reduces to 40%, and the likelihood that there will be a delay rises from 5% to 60%. The evidence at the nodes 'derailed train' and 'signal disturbance' propagates through the network, causing the MPD for the node 'train delay' to alter.

Serial d-connection

A second type of d-connection, i.e. serial d-connection, is shown in Figure 26. In this connection, 'Jack on time at work?' is dependent on 'train delay', whereas 'train delay' is dependent on 'derailed train' and 'signal disturbance'. Therefore, evidence for 'derailed train' and/or 'signal disturbance', influences the node 'Jack on time at work?' However, *if* there is evidence for 'train delay', then new evidence for 'derailed train' will not influence the node 'Jack on time at work?': the latter two nodes ('train delay' and 'Jack on time at work?') are therefore conditionally independent, *given* evidence at the node 'train delay'

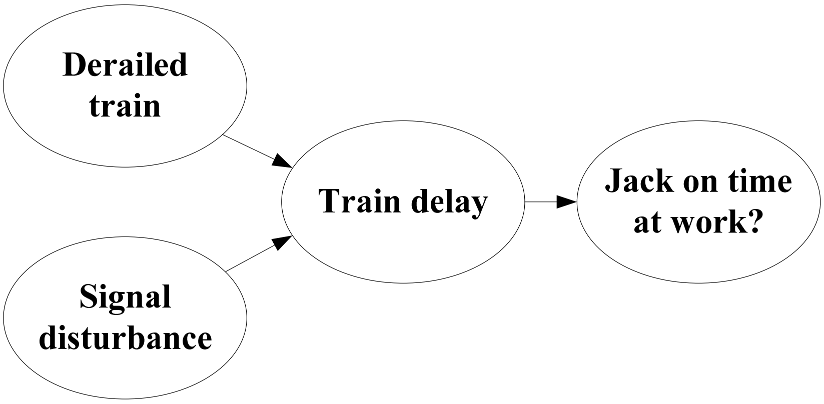


Figure 26: serial d-connection

In order to show the way in which the serial connection works, the CPTs for p(TD|SD,DT) (Table 23) and the CPT for p(JOT|TD)) (Table 24) are given.

Table 24: CPT for the node JOT

Train delay (TD)	**Jack on time at work? (JOT)**	
	Yes	*No*
Yes	0.10	0.90
No	0.80	0.20

Using the evidence that there is a derailed train on the tracks, *and* there is a signal disturbance can be used to calculate the marginal distribution of Jack being on time at work. Taking the evidence into account, the MPD p(TD|DT=Yes, SD=Yes) can be inferred from Table 23:

Table 25: MPD of TD, given DT=Yes and SD=Yes

Train delay (TD)	
Yes	*No*
0.95	0.05

Since the marginal distribution of 'Train delay' is known, the marginal distribution of 'Jack on time at work?' (p(JOT|TD) can be inferred as follows:
p((JOT|TD) = (p(JOT=Yes|TD), p(JOT=No|TD)).
p(JOT=Yes|DT=Yes) = p(TD=Yes)·p(JOT=Yes|TD=Yes) + p(TD=No)·p(JOT=Yes|TD=No)
= 0.95·0.10 + 0.05·0.80 = 0.135.
p(JOT=No|DT=Yes) = p(TD=Yes)·p(JOT=No|TD=Yes) + p(TD=No)·p(JOT=No|TD=No)
= 0.95·0.90 + 0.05·0.20 = 0.865.
The results are presented in Table 26:

Table 26: MPD of JOT, given TD=Yes

Jack on time at work (JOT)	
Yes	*No*
0.135	0.865

Hence, using Table 26, there is a likelihood of 0.865 of Jack not being on time at work, *if* there is a derailed train and a signal disturbance. Therefore, if Jack would be aware of a train being derailed and a signal disturbance before he leaves his house (e.g. through a news report on the radio), he might decide to go by car instead, because of the high likelihood of being late at work if he would take the train. In this way, a BBN supports decision making.

Diverging d-connection

A third type of d-connection is called diverging d-connection and is depicted in Figure 27: the node 'traffic jam' being connected as a parent node to both the child node 'Jack on time at work?' and the child node 'Thomas on time at work?' In this connection, evidence at the node 'traffic jam' will influence both 'Jack on time at work?' and 'Thomas on time at work?' At the same time, neither of the latter nodes can influence each other, *given* evidence at the node 'traffic jam' Hence, 'Jack on time at work?' and 'Thomas on time at work?' are conditionally independent given evidence at the node 'traffic jam'.

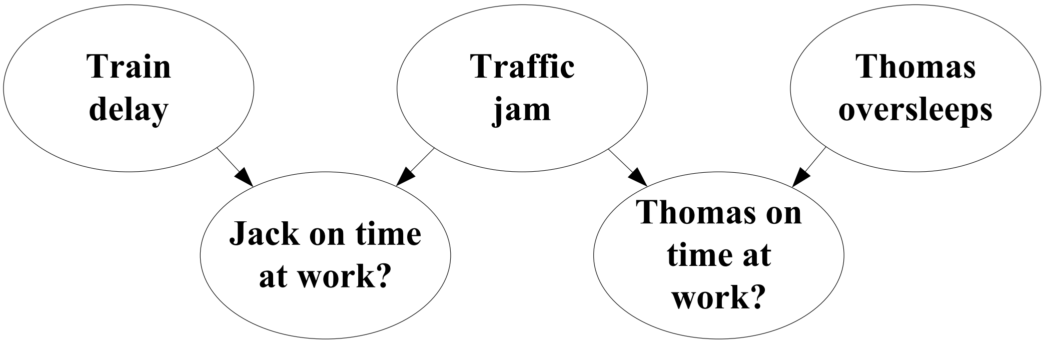


Figure 27: diverging d-connection

In this situation, it is interesting to know the probabilities whether Thomas and Jack are on time at their work. In order to make a calculation regarding this situation, the CPTs of p(TOT|TO,TJ) and p(JOT|TD,TJ) are needed. These are given in Table 27 and Table 28 as an example.

Table 27: CPT for the node TOT given TO and TJ

Traffic jam (TJ)	**Thomas oversleeps (TO)**	**Thomas on time at work? (TOT)**	
		Yes	*No*
Yes	*Yes*	0.05	0.95
	No	0.40	0.60
No	*Yes*	0.25	0.75
	No	0.90	0.10

Table 28: CPT for the node JOT given TD and TJ

Traffic jam (TJ)	**Train delay (TD)**	**Jack on time at work? (JOT)**	
		Yes	*No*
Yes	*Yes*	0.10	0.90
	No	0.60	0.40
No	*Yes*	0.30	0.70
	No	0.90	0.10

Assuming that there is no train delay, Thomas has not overslept, but there is a traffic jam, then the PTs for Jack being on time and Thomas being on time, become:

Table 29: MPD of TOT, given TO=No, and TJ=Yes

Thomas on time at work? (TOT)	
Yes	*No*
0.40	0.60

Table 30: MPD of JOT, given TD=No, and TJ=Yes

Jack on time at work? (JOT)	
Yes	*No*
0.60	0.40

In this case, evidence that Jack is on time does not change the PT of ‘Thomas on time?’ given that he does not oversleep, and there is a traffic jam (because of the conditional independence of ‘Thomas on time?’ and ‘Jack on time?’ given evidence at ‘Traffic jam?’).

However, if there would be no evidence whether there is a traffic jam or not, then the marginal PT of Thomas being on time and Jack being on time would change, assuming a non-informative PT for ‘traffic jam’ (uniformly distributed among the states “Yes” and “No”. This would lead to p(JOT=Yes) becoming 75%, and p(TOT=Yes) becoming 65%. Now, if there was evidence that Jack was on time p(JOT=Yes) becomes 100% (it increases with 25%). Backwards reasoning gives that the likelihood of Thomas being on time increases by 5% to 70%. For this, it is needed to first calculate the marginal probability that Jack is on time (p(JOT=Yes)). This is calculated as follows:
$p((JOT=Yes) = p(TJ=No)\cdot(p(JOT=Yes|TJ=No) + p(TJ=Yes)\cdot p(JOT=Yes|TJ=Yes)) =$
$0.50\cdot 0.60 + 0.50\cdot 0.90 = 0.75.$
p(TJ=Yes) can be derived from the PT for p(TJ). Since that PT is non-informative, p(JOT=Yes) = 0.50.
p(JOT=Yes|TJ=Yes) can be derived from (p(JOT=Yes|TD=No,TJ=Yes)), which is 0.60.
Using Bayes’ theorem, this results in the probability of there being a traffic jam, given that Jack is on time, taking into account that there is no train delay:

$$p(TJ=Yes|JOT=Yes) = \frac{p(JOT=Yes|TJ=Yes)\cdot p(TJ=Yes)}{p(JOT=Yes\ )} = \frac{0.60\cdot 0.50}{0.75} = 0.40$$

Conversely, p(TJ=No |JOT=Yes) is 0.60. Using these numbers, the marginal PT for p(TOT) can be calculated:
$p((TOT=Yes) = p(TJ=No)\cdot(p(TOT=Yes|TJ=No) + p(TJ=Yes)\cdot p(TOT=Yes|TJ=Yes))$
$= 0.60\cdot 0.90 + 0.40\cdot 0.40 = 0.70$, leading to Table 31:

Table 31: Probability table of TOT, given p(TJ=No) = 0.60, and p(TJ=Yes) = 0.40

Thomas on time at work?	
Yes	*No*
0.70	0.30

B. Interviewed Experts and their Disciplines

Expert	**Number**	**Format**	**Date**	**Interviewer(s)**
First round of information collection: 7 experts, semi-structured interviews, 2 interviewers				
System test manager	1	Report	28-02-2007	I & II
Sr. Mechanical designer	2	Report	28-02-2007	I & II
New Products Introduction Manager	3	Report	03-04-2007	I & II
Service Innovation Engineer	4	Report	03-04-2007	I & II
Coordinator helpdesk / Field engineer and help desk	5	Report	16-04-2007	I & II
Quality Data Engineer	6	Report	18-04-2007	I & II
System Architect	7	Report	23-04-2007	I & II
Second round of information collection: 9 experts, unstructured interviews, 2 interviewers				
Production (Manager Engineering (IS OPS_OXB Mngt*)	8	Full transcript	07-06-2007	II & III
Marketing (International Senior Product Manager*)	9	Full transcript	07-06-2007	II & III
System design (System Architect*)	10	Full transcript	13-06-2007	II & III
Integration and testing (Competence Area Manager System Integration &Testing*)	11	Full transcript	14-06-2007	II & III
Customer Support & New Product Introduction (CV CS Innovation*)	12	Full transcript	14-06-2007	II & III
Customer support and head of the “flying quality squad” (Customer Satisfaction Manager*)	13	Full transcript	14-06-2007	II & III
Customer Services (Tier 3 support engineer Customer Service*)	14	Full transcript	14-06-2007	II & III
Application & Clinical Science CV (CV Marketing [Interus expert]*)	15	Full transcript	15-06-2007	II & III
Software architect/reliability engineer (CV Software Architect*)	16	Full transcript	19-09-2007	II & III

**: For these participants, only the area of expertise was known at the moment of interviewing; not the name of the function. The name of their function has been available only from February 23rd, 2009, and has been added to the description of the expert in brackets.*

Third round of information collection: 10 experts, structured interviews, 1 interviewer				
Principal Scientist and Technology Manager	17	Summary of factors	15-05-2008	III
Software architect	18	Summary of factors	15-05-2008	III
CV electronics	19	Summary of factors	15-05-2008	III
Test Architect	20	Summary of factors	15-05-2008	III
System Design	21	Summary of factors	16-05-2008	III
Manufacturing Engineer	22	Summary of factors	16-05-2008	III
Product Quality Improvement Engineer	23	Summary of factors	21-05-2008	III
System Designer	24	Summary of factors	21-05-2008	III
Team leader Service Innovation	25	Summary of factors	03-06-2008	III
PQ&M Manager	26	Summary of factors	03-06-2008	III
Fourth round of information collection: 2 experts, unstructured interviews, 1 interviewer				
System Architect	27 (7)	Report	22-09-2008	III
System Architect	28 (10)	Report	22-09-2008	III
Focus group session: 5 experts that participated				
System Architect		Survey	03-03-2010	III & Gatekeeper
System Architect		Survey	03-03-2010	III & Gatekeeper
PQ&M Manager		Survey	03-03-2010	III & Gatekeeper
System designer		Survey	03-03-2010	III & Gatekeeper
NPI Manager CS		Survey	03-03-2010	III & Gatekeeper
Individual discussion: 1 expert (decision maker)				
Development Director Cardio Vascular		Discussion	15-03-2010	III & Gatekeeper

C. BBN Models for D&D Stage, M&I Stage, and Full PDP

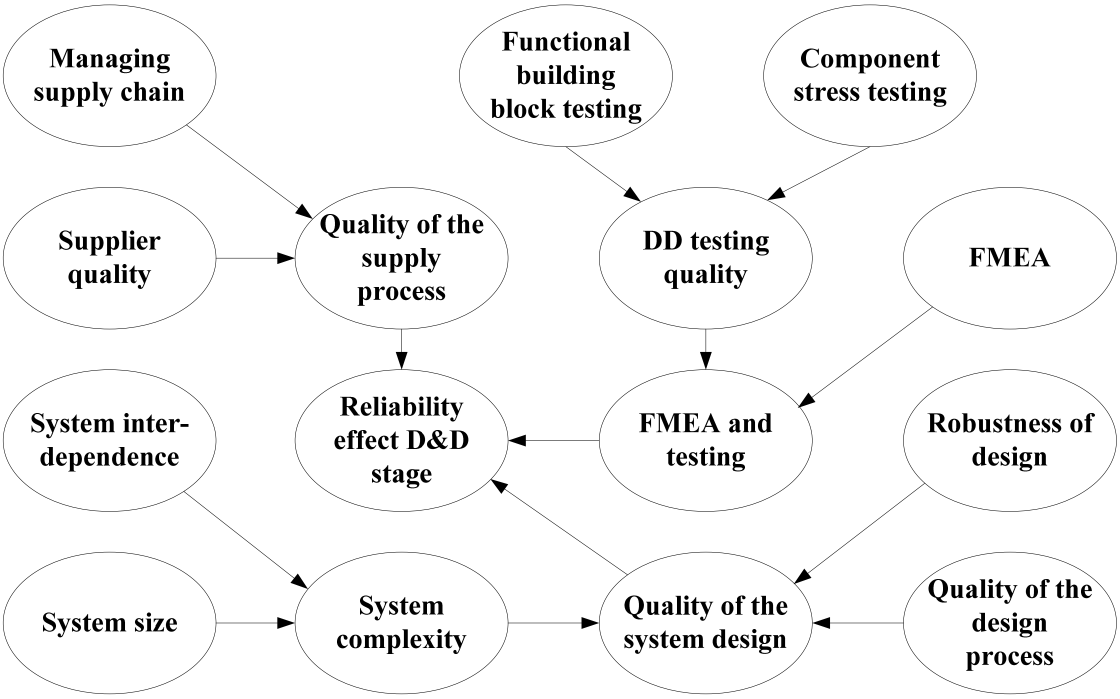


Figure 28: Resulting BBN for the D&D stage

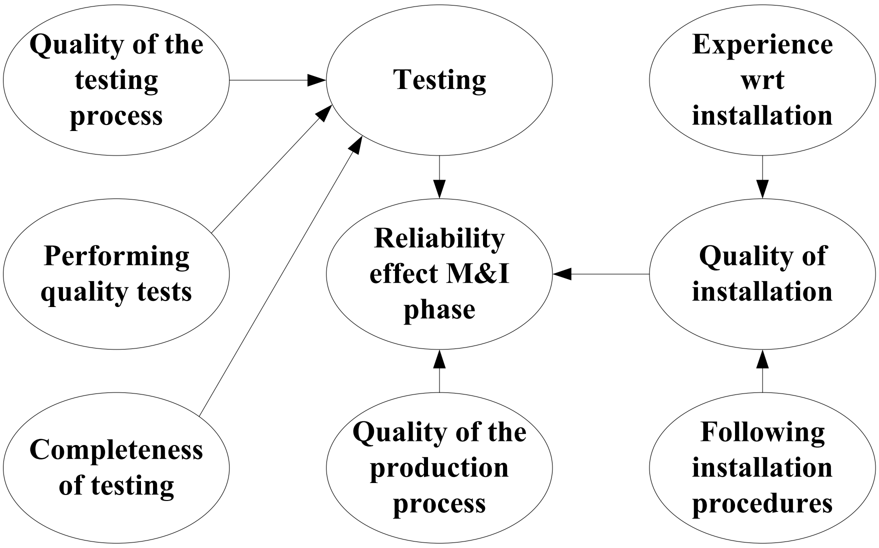


Figure 29: Resulting BBN for the M&I stage

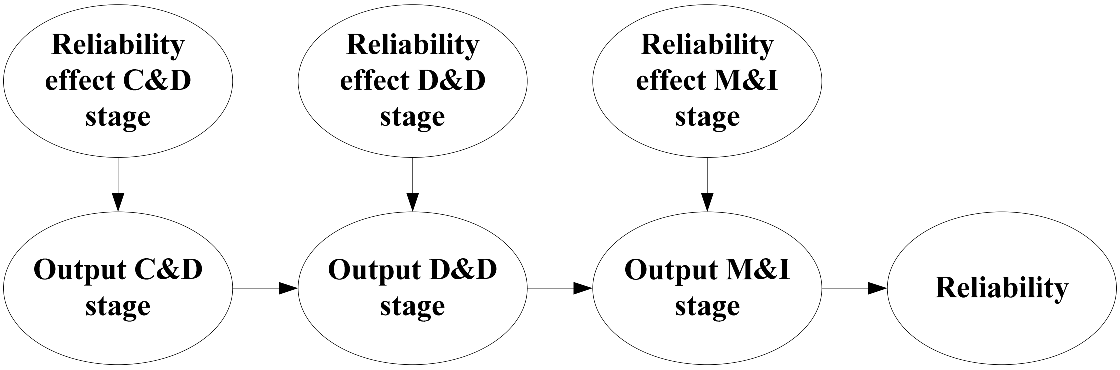


Figure 30: Resulting BBN for the full PDP

D. Variable descriptions for D&D Stage and M&I Stage

Table 32: Descriptions of variables that are included in the BBN for the D&D stage

Variable name	Description
Functional testing of building blocks	This variable gives an indication of the amount attention that is given to subsystems through functional tests. Particularly the critical subsystems are of importance.
Component stress testing	Component stress testing depicts stress tests of components that give insight into the reliability of the separate components. Are stress tests on component level being performed either by the company or by its suppliers, or are data collected in a different way?
Quality of the supply process	The quality of the supply process is determined as the combined effect of supplier quality and supplier management; it determines the quality of the components that are supplied.
FMEA	This represents how well FMEAs are applied. A number of aspects are of importance regarding FMEAs, and determine the quality with which it is applied. These are: - multidisciplinary group - actions are being performed with respect to redesign and testing of the product - The product is being redesigned through defined actions
FMEA & Testing	This is the quality of FMEA and testing, taken together as one “influencing variable”.
System complexity	Here, the following things play a role: the functionality (number of functions) of the system and the number of interactions between components within the system (interdependency)). What is the level of complexity of the system?
System size	How large is the system in terms of components and functions compared to the previous product.
System interdependence	Interdependence is related to both functions and components. How many interactions between functions and components are there within the system, *compared to the previous system.*
Robust design	Here, the robustness of the design is a characteristic of the design rather than a direct measure for the reliability. It reflects how insensitive to changes in the environment the design is. It represents the sensitivity of the system with respect to changes in circumstances, where the change is such that the circumstances are not any longer within specifications?
Quality of system design	The quality of system design is determined by the systems robustness, the quality of the design process, and the changes that have been made in the design.

Quality of the design process	The quality of the design process is represented by: - the skills of the people - the history of the design process in the company - use of tools and techniques regarding reliability - the available resources
Supplier quality	Supplier quality is determined by the following 2 characteristics: 1. Supplier has process control and ensures 0-hr quality 2. Supplier explicitly gives attention to reliability?
Managing the supplier	Regarding management of the supplier, 2 things play a role: 1. The amount of insight that the company has in the processes (both development and production) of the supplier. 2. The amount of influence that the company can exert influence on the supplier.
DD testing quality	The DD testing quality represents the combined effect of the functional building block tests and the component stress tests.

Table 33: Descriptions of variables that are included in the BBN for the M&I stage

Variable name	Description
Quality of installation	This is the quality of installation, being a combination of the experience of people with respect to the installation, and the extent to which they follow procedures.
Experience with respect to installation	Here, the number of systems that the engineers have installed plays a role, but also the knowledge that the engineers have from other sources. How much experience do the engineers have with respect to installing systems at the customer?
Following installation procedures	Do the engineers that install the system at a customer site do the right activities in the right order?
Quality of production testing process	What is the quality of the way in which a specific completed, produced system is tested at the end of production?
Performing quality tests	These tests relate to the functional system tests, where the performance of the system is tested to be within limits. What is the quality of the functional tests that are performed on a general systems level?
Completeness of test coverage	How much of the sum of the total amount of functions (CTQ, Safety, SPC, etc) and their performance parameters are tested.
Quality of the production assembly process	The quality of the production assembly process is influenced by the experience of the people that work in production-assembly and how well the procedures are followed.
Overall quality of tests during production / assembly	This is a variable that measures the amount of problems found during testing in relation to the total intrinsic problems.

E. Variable state-spaces for D&D Stage and M&I Stage

Table 34: Variables that are included in the BBN for the D&D stage and their state-spaces

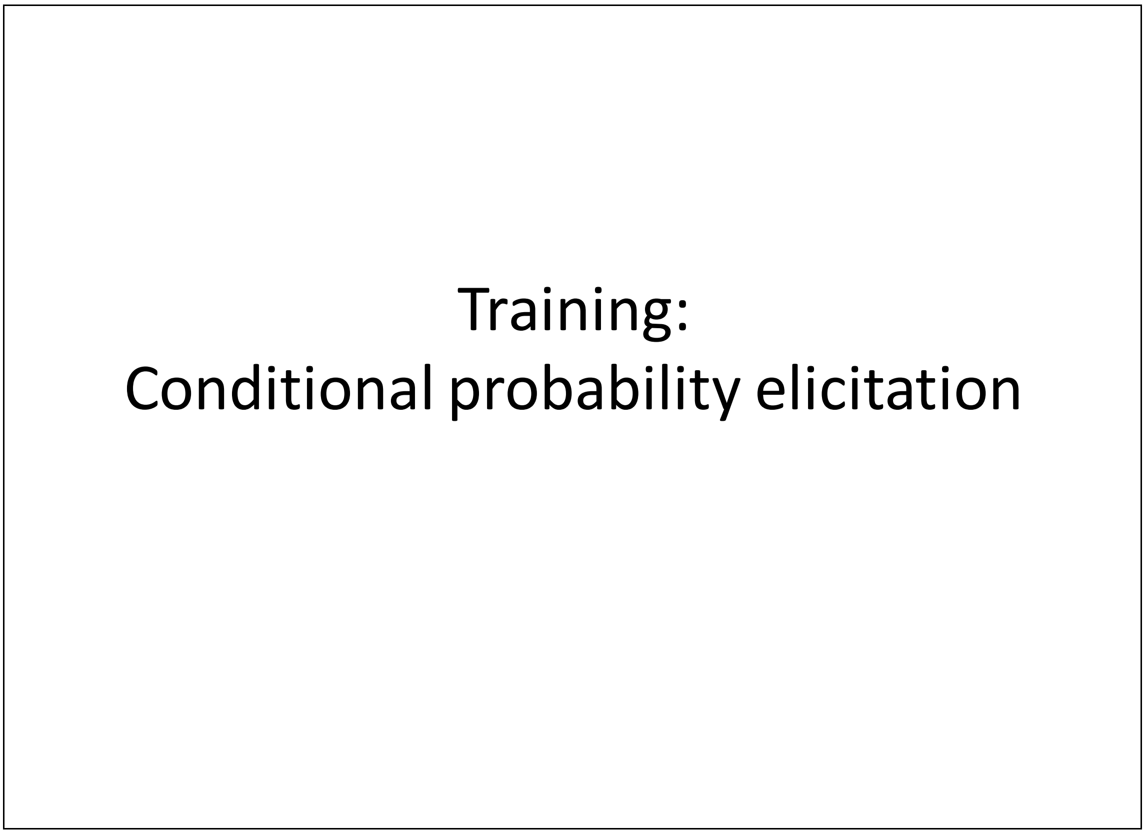
Variable name	State space
Functional testing of building blocks	1. >75% of the critical subsystems is tested 2. Between 50 and 75% of the critical subsystems is tested 3. <50% of the critical subsystems is tested
Component stress testing	1. >75% of the critical components is tested 2. Between 50 and 75% of the critical components is tested 3. <50% of the critical components is tested
Quality of the supply process	1. High 2. Average 3. Low
FMEA	1. High 2. Average 3. Low
FMEA & Testing	1. Good 2. Average 3. Bad
System complexity	1. High 2. Average 3. Low
System size	1. >10% new components and functions 2. <10% new components and functions
System interdependence	1. Many more 2. Few more 3. Less
Robust design	1. Both primary and secondary functionality are insensitive 2. Primary functionality is insensitive, secondary functionality is sensitive 3. Both primary and secondary functionality is sensitive
Quality of system design	1. Good 2. Average 3. bad
Quality of the design process	1. High 2. Average 3. Low
Supplier quality	1. High: supplier has both 1 and 2 2. Average: supplier has 1, but not 2 3. Low: supplier has neither

Managing the supplier	1. High: supplier has both 1 and 2 2. Average: supplier has 1, but not 2 3. Low: supplier has 1 only for a part
DD testing quality	1. High 2. Average 3. Low

Table 35: Variables that are included in the BBN for the M&I stage and their state-spaces

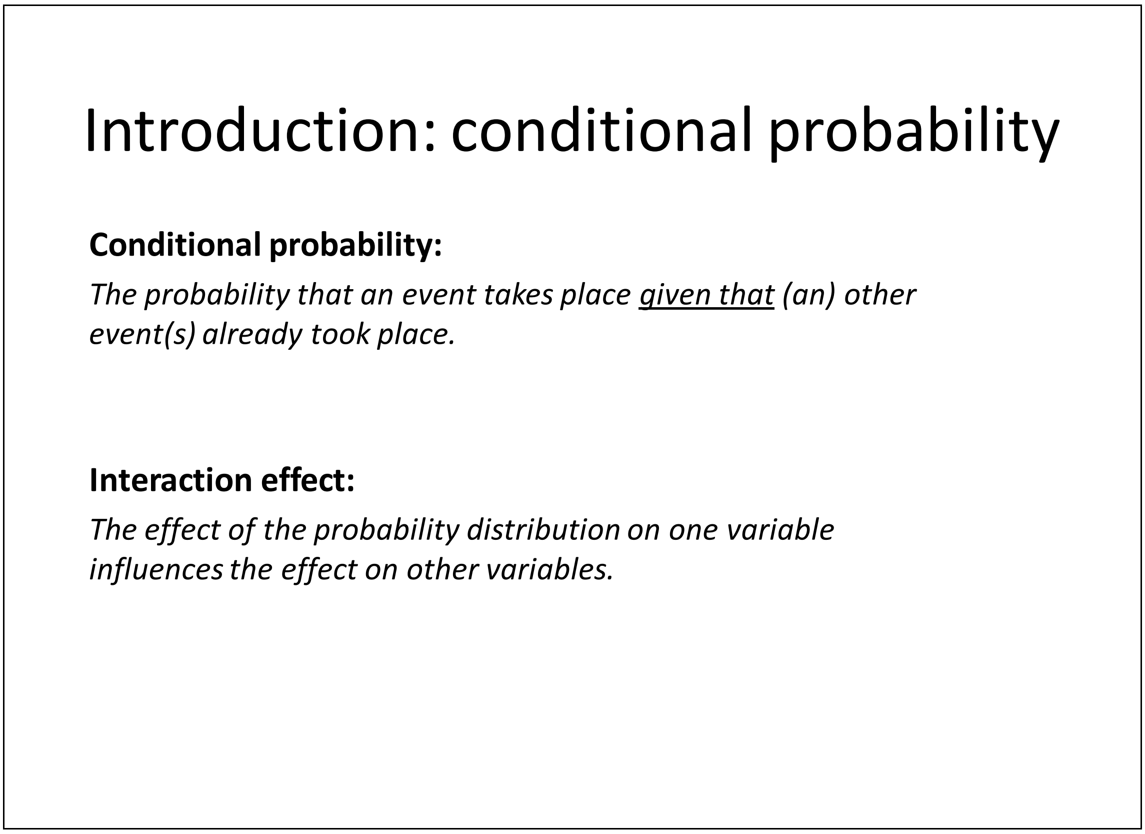
Variable name	State space
Quality of installation	1. Good 2. Average 3. Bad
Quality of installation	1. Good; all aspects are good: o The performance after installation is equal to that at the factory o No customer complaints wrt installation o Room is equal to room lay-out 2. Average; 1 of three aspects is missing 3. Bad; >1 aspect of three is missing
Experience with respect to installation	1. Performed more than 20 installations 2. Performed less than 20 installations, OR more than 6 installations 3. Performed less than 6 installations
Following installation procedures	1. Fully according to SMI (Service Manual of Installation) 2. SMI is used ad-hoc; all the activities have been performed taking a look at the order. 3. All activities have been performed, without taking notice of the SMI
Quality of production testing process	1. CTQs are defined, a test strategy and plan have been defined and tests are evaluated for the future 2. Test plan has been defined, but only little attention to quality 3. No structured testing program
Performing quality tests	1. Up to performance level testing 2. Up to functional testing 3. Safety testing
Completeness of test coverage	1. >95% 2. between 70 - 95% 3. < 70%
Quality of the production assembly process	1. High quality; Trained, assembled > 40 tables and worked according to procedure 2. Average quality; 1 of three aspects is missing 3. Low quality; >1 of three aspects is missing
Overall quality of tests during production/assembly	1. >90% of intrinsic problems found 2. 70%-90% 3. <70%

F. Sheets for Probability Elicitation Training



Training:
Conditional probability elicitation

Box 3a: Presentation for probability elicitation training, sheet 1



Introduction: conditional probability

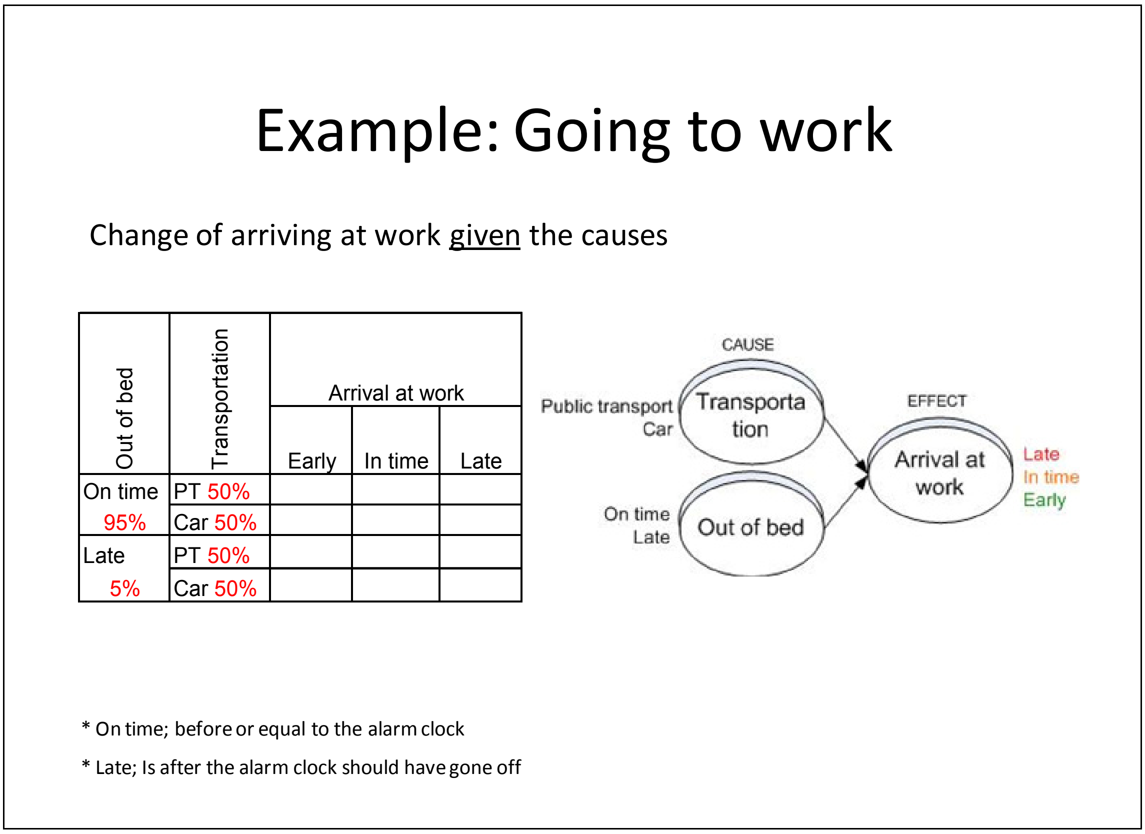
Conditional probability:

The probability that an event takes place given that (an) other event(s) already took place.

Interaction effect:

The effect of the probability distribution on one variable influences the effect on other variables.

Box 3b: Presentation for probability elicitation training, sheet 2



Example: Going to work

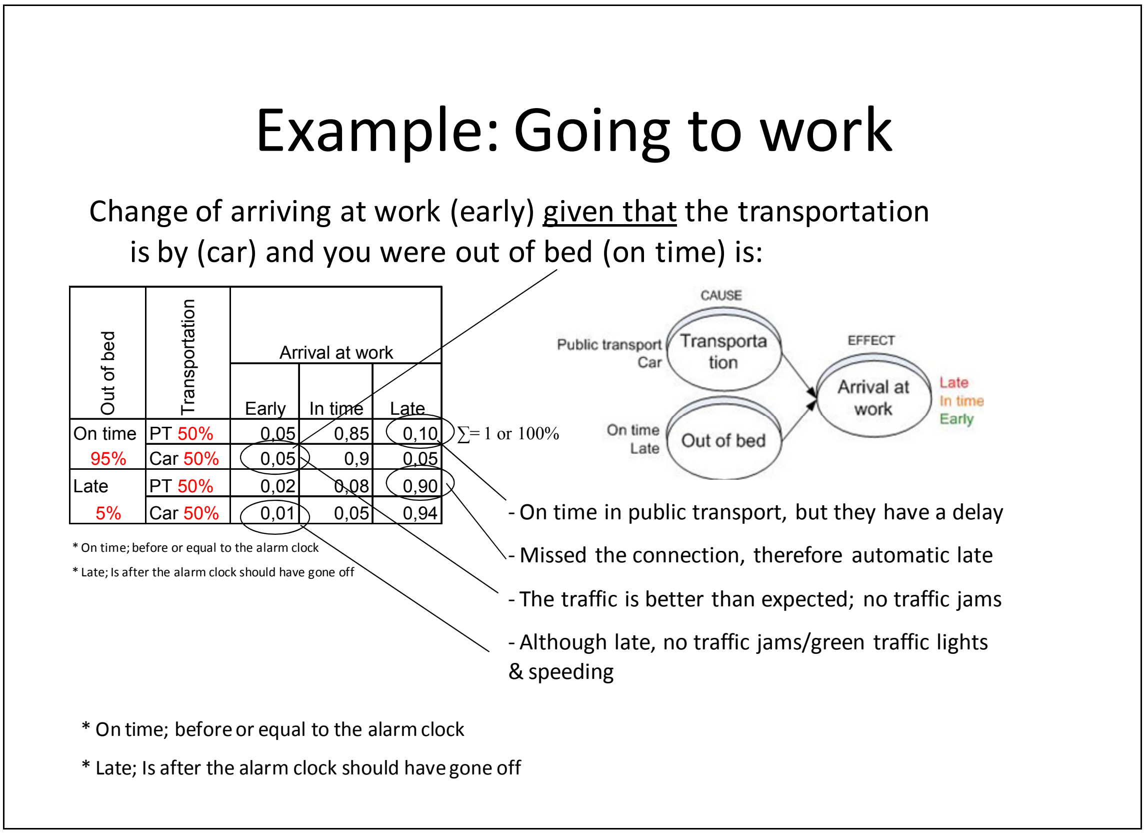
Change of arriving at work given the causes

Out of bed	Transportation	Arrival at work		
		Early	In time	Late
On time 95%	PT 50%			
	Car 50%			
Late 5%	PT 50%			
	Car 50%			

* On time; before or equal to the alarm clock

* Late; Is after the alarm clock should have gone off

Box 3c: Presentation for probability elicitation training, sheet 3



Example: Going to work

Change of arriving at work (early) given that the transportation is by (car) and you were out of bed (on time) is:

Out of bed	Transportation	Arrival at work		
		Early	In time	Late
On time 95%	PT 50%	0,05	0,85	0,10
	Car 50%	0,05	0,9	0,05
Late 5%	PT 50%	0,02	0,08	0,90
	Car 50%	0,01	0,05	0,94

∑= 1 or 100%

* On time; before or equal to the alarm clock

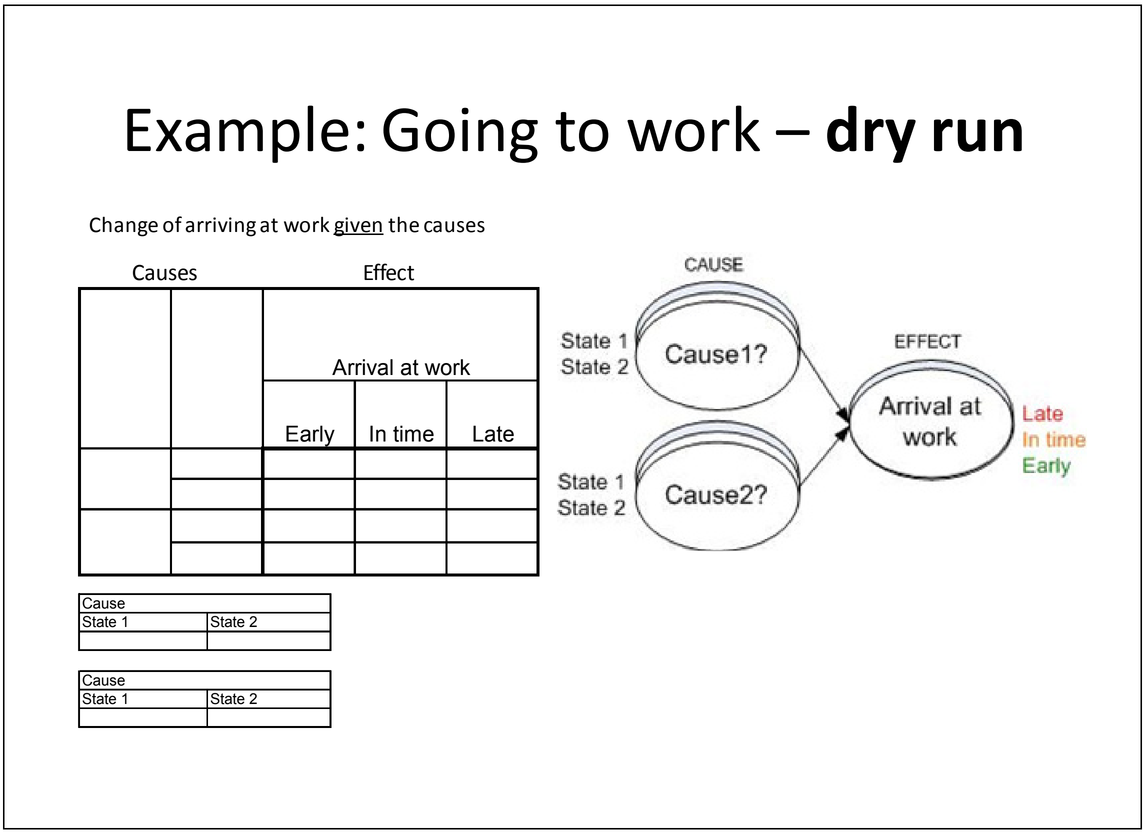
* Late; Is after the alarm clock should have gone off

- On time in public transport, but they have a delay
- Missed the connection, therefore automatic late
- The traffic is better than expected; no traffic jams
- Although late, no traffic jams/green traffic lights & speeding

* On time; before or equal to the alarm clock

* Late; Is after the alarm clock should have gone off

Box 3d: Presentation for probability elicitation training, sheet 4



Example: Going to work – **dry run**

Change of arriving at work given the causes

Causes | Effect

		Arrival at work		
		Early	In time	Late

Cause	
State 1	State 2

Cause	
State 1	State 2

Box 3e: Presentation for probability elicitation training, sheet 5

G. Survey/Questionnaire for model validation

Dear mr./mrs.,

In the past years, I have performed research at [company] regarding reliability. In my research, I have identified a number of factors that influence reliability (to a larger or smaller extent). This is a survey questionnaire, which is aimed at validating the results of this earlier research.

This is a survey questionnaire regarding factors that affect reliability. These factors affect reliability throughout the product creation process (PCP), in the concept and definition (C&D) phase, the design and development (D&D) phase and the manufacturing and installation (M&I) phase. In this study, we include both factors that have a technical nature (e.g. 'system size'), and factors that have a non-technical nature (e.g. 'mindset towards reliability').

The purpose of this survey questionnaire is to obtain your opinion on how strongly the factors influence the reliability of the product. For this purpose, we want to ask you to rank the following factors in terms of their influence on reliability:

- If you think that the factor has **very much influence** on reliability, please rank it at "7".
- If you think that the factor has **no influence** on reliability, please rank it at "1".
- If you are **not able to estimate the influence** of a factor on reliability, then please fill in "0" (at the extreme right of the form).

The questionnaire consists of four sections: one section for every phase in the product development process (C&D; 1 page, D&D; 2 pages, M&I; 1 page).

The ranking can be performed by clicking the box that corresponds to your answer. An answer can be corrected by clicking on the checkbox again (emptying it), and after that clicking on the right box.

After filling in all questions, I would like to ask you to save the document, and return the survey to the e-mail address below (M.J.H.A.Houben@tue.nl), or to: Name@company.com.

The results of the survey will be processed anonymously.

Thank you very much in advance for your cooperation,

Maurits Houben

PhD Student, Eindhoven University of Technology (TU/e)

Phone: (+31 40) 247 5053 / 3601
E-mail: M.J.H.A.Houben@tue.nl

C&D phase: variables and their influence on Reliability

Variable name description	*No influence*						*Very much influence*	*No opinion*
Approach towards reliability The priority level of reliability for the company (putting reliability on the agenda of the management-team meeting	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Mindset towards reliability How the people think about reliability throughout the PLC (product lifecycle) (in terms of activities and deliverables during the PCP)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Cost saving The need to save costs during the concept and definition phase of a product	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Product innovation context The type of innovation that is introduced, (the change being from evolutionary (small) to revolutionary (large))	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Historical reputation regarding reliability The reputation of the company in the area of reliability, related to the way in which reliability was treated in the past	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Requirements The familiarity of the designers with customer wishes and their ability to translate these to system and subsystem level, taking the function into account, as well as the level on which, and the circumstances under which it has to perform	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Component specifications The quality of the specifications of the components (knowledge on components and on the conditions under which it has to perform)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
New technology The ability of the company to deal with the technological change of products and the (increasing) influx of new technology	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Introduction of new components The percentage of the total functions of the product that is new	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Pressure on development time The amount of time that is made available for product conceptualization and definition	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Continuously changing specifications The extent to which the company is able to cope with the continuously changing in product specifications throughout the PDP	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Adding features The number of features and functions that is added to the specific product looked at in the light of previous developments.	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0

D&D phase (I): variables and their influence on Reliability

Variable name description	*No influence*						*Very much influence*	*No opinion*
Supplier quality Supplier quality is determined by the following 2 characteristics: 1. Supplier has process control and ensures 0-hr quality 2. Supplier explicitly gives attention to reliability?	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Functional testing of building blocks The percentage of critical subsystems that is tested	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Interconnectivity The way in which functions and building blocks are related to each other (from ‘one-to-one’ to ‘one-to-many’)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
System tolerances The tightness of the tolerances of the different components and subsystems with respect to the way in which they physically connect to other components and subsystems	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
FMEA The quality with which FMEAs are applied. The quality of appliance of FMEAs is determined by: - Using a multidisciplinary group - Actions are being performed with respect to redesign and testing of the product - The product is being redesigned through defined actions	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Standard components Using off-the shelf components and subsystems instead of self-developed components and subsystems	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Knowledge regarding the use of tools in development How much knowledge the people that are involved in development have *regarding the use of reliability tools*	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Learning from previous problems into account The extent to which problems in earlier designs are taken into account when designing a new product, and how much is learnt from previously designed systems	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
System size The change in system size in terms of number of components and functions, compared to the previous product	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Operational tests not according to use The extent to which the way in which the system is tested in the D&D phase reflects the way in which the system is used	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0

D&D phase (II) : variables and their influence on Reliability

Variable name description	*No influence*						*Very much influence*	*No opinion*
System interdependence The number of interactions between functions and components within the system, compared to the previous system	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Component stress testing The percentage of the critical components that is tested	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Robust design The sensitivity of the system with respect to changes in circumstances, where the circumstances are not any longer within specifications	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Speed of development The speed with which new components are introduced	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Software design changes The way in which changes that are made in the software affect reliability	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Quality of the design process The quality of the design process is represented by: - the skills of the people - the history of the design process in the company - use of tools and techniques regarding reliability - the available resources	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Managing the supply chain Regarding management of the supplier, 2 things play a role, i.e.: the amount of *insight* that [company] has in the processes (both development and production) of the supplier, and the amount of *influence* that [company] can exert on the supplier	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Experience of design engineer The amount of experience the design engineers have with the design of products	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0

M&I phase: variables and their influence on Reliability

Variable name description	*No influence*						*Very much influence*	*No opinion*
Manufacturing reorganization Reorganization of the production/assembly process within the company	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Following installation procedures The extent to which the engineers install the system at a customer site according to SMI (Service Manual of Installation)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Human errors The effect of human errors during the production/assembly process	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Quality of production testing process The quality of the way in which a specific completed, produced system is tested at the end of production (in terms of CTQs (critical to quality's), test strategy, plans and evaluation of tests	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
SPC (Statistical Process Control) The extent to which the production/assembly process is under control	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Completeness of test coverage The percentage of the total amount of functions (critical to quality, Safety, etc) and their performance parameters that is tested.	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Quality of material Whether the appropriate materials are used during production/assembly	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Transportation The influence that the transportation of the system to the field has on reliability	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Quality of the production assembly process The quality of the production/assembly process, which is influenced by the experience of the people that work in production/assembly, and the quality with which the procedures are followed	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Performing quality tests The level of testing up to where the systems are tested (ranging from only safety testing to safety, functional *and* performance level testing)	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Experience with respect to installation Both the number of systems that the engineers already have installed in the past, and their knowledge from other sources	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0
Quality of tools The quality of the tooling in the production/assembly process on reliability	☐1	☐2	☐3	☐4	☐5	☐6	☐7	☐0

H. Questions in survey for model validation presentation

Questions in Dutch:

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
Het model van de C&D fase geeft goed weer wat mijn visie is op de manier waarop verschillende factoren in het model reliability beïnvloeden	☐	☐	☐	☐	☐

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
Het model van de D&D fase geeft goed weer wat mijn visie is op de manier waarop verschillende factoren in het model reliability beïnvloeden	☐	☐	☐	☐	☐

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
Het model van de M&I fase geeft goed weer wat mijn visie is op de manier waarop verschillende factoren in het model reliability beïnvloeden	☐	☐	☐	☐	☐

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
Het model van het PDP geeft goed weer wat mijn visie is op de manier waarop de verschillende fases van het PDP zijn gekoppeld, en hoe zij reliability beïnvloeden	☐	☐	☐	☐	☐

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
Main Effects Analysis is een goede manier om de kwaliteit van een model te toetsen	☐	☐	☐	☐	☐
Scenario analyse is een goede manier om de kwaliteit van een model te toetsen	☐	☐	☐	☐	☐

	mee oneens	enigszins mee oneens	niet mee eens/niet mee oneens	enigszins mee eens	mee eens
De modellen die zijn gepresenteerd kunnen gebruikt worden voor reliability management	☐	☐	☐	☐	☐

	mee oneens	**enigszins mee oneens**	**niet mee eens/niet mee oneens**	**enigszins mee eens**	**mee eens**
Deze presentatie heeft mijn inzichten in de manier waarop verschillende factoren, gedurende het PDP, reliability beïnvloeden vergroot	☐	☐	☐	☐	☐

English translation:

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
The model of the C&D stage is a good representation of my view on the way in which the different factors in the model affect reliability	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
The model of the D&D stage is a good representation of my view on the way in which the different factors in the model affect reliability	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
The model of the M&I stage is a good representation of my view on the way in which the different factors in the model affect reliability	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
The model of the PDP is a good representation of my view on the way in which the different stages in the PDP are linked, and affect reliability	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
Main Effects Analysis is a good way to evaluate the quality of the model	☐	☐	☐	☐	☐
Scenario analysis is a good way to evaluate the quality of the model	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
The models that were presented can be used for reliability management	☐	☐	☐	☐	☐

	strongly disagree	**disagree**	**neutral**	**agree**	**strongly agree**
This presentation helped increase my insights in the way in which different factors throughout the PDP influence reliability	☐	☐	☐	☐	☐

Curriculum Vitae

Maurits Houben was born in Deurne, the Netherlands, on April 28th, 1982. In 2006 he received his Masters degree (with distinction) in Industrial Engineering and Management Science from the Eindhoven University of Technology. The topic of his graduation project was the presence of soft reliability problems in other industries than the consumer electronics industry.

In March 2006, he started his Ph.D. research project at the sub department of Quality and Reliability Engineering at the faculty of Technology Management at the Eindhoven University of Technology. From January 2008 onwards, this project was continued at the sub department of Operations, Planning, Accounting and Control at the same faculty, which was renamed to the faculty of Industrial Engineering and Innovation Sciences. The research project was funded by the Innovation-Oriented Research Programme ‘Integrated Product Creation and Realization (IOP IPCR)’ of the Netherlands Ministry of Economic Affairs. Diverse industrial and academic partners were involved in the program.

Since April 9th 2010, he is involved in valorisation of his research at the Eindhoven University of Technology. From July 1st 2010, he is working as RAMS advisor/specialist at the Directorate-General of Public Works and Water Management.