



A Prescriptive Approach to Eliciting Decision Information

Mona Riabacke

©Mona Riabacke, Stockholm University 2012

ISSN 1101-8526
ISBN 978-91-7447-517-3
Report Series No. 12-008

Printed in Sweden by Universitetservice US AB, Stockholm 2012
Distributor: Department of Computer and Systems Sciences

To my family

ABSTRACT

The importance of making good decisions is well-known in the business world and the perception of decision making as rational choice is the most common view. The amount of information involved in many decision making situations has increased dramatically in recent years and support of some kind is often needed. Consequently, fields like Business Intelligence (BI) and Decision Support Systems (DSS) have advanced.

Decision analysis applications belong to the latter category and aim to support decision making activities in businesses and organizations, and provide more clearly structured decision material to use as a basis for decisions. In spite of a belief in their potential, their employment is still limited in practice. This can partly be attributed to the fact that existing tools are incomplete to support decision processes sufficiently in real settings. Moreover, in traditional decision process models, the *elicitation* of input data to the decision model is one of the key components during modelling. However, in spite of the fact that it involves quite a few problematic elements and its importance to the quality of the process as a whole, the specification and execution of the elicitation process is often left to the discretion of the user.

This thesis focuses on the elicitation of information in decision analysis applications and the need for more prescriptive methods (realistic and practically useful) than what is offered today. A process model emphasizing the importance of structured elicitation of adequate input data throughout decision processes is also suggested. In order to further define the problematic aspects of elicitation, three empirical studies were conducted, where problems from existing literature were further studied. The problems with eliciting precise decision data suggests that using imprecise values within elicitation is a more realistic and useful approach to strive for, and perhaps even more important within preference elicitation. Based on theory and the findings of the studies presented in papers I-III, an elicitation method for imprecise statements and noisy input was formalized and the method was further refined into the Cardinal Rank Ordering of Criteria (CROC) weight elicitation method, presented in paper VII. The CROC method is both compatible with an adapted prescriptive decision making model, focused on a more structured elicitation component as well as algorithms for dealing with such data. It was employed and validated in two real-life cases (described in papers V and VI), which is not so common within decision analysis research, i.e. there are relatively few examples of real-life cases employing such techniques.

SAMMANFATTNING

Vikten av att fatta bra beslut är välkänt inom företagsvärlden och uppfattningen om att beslutsfattande handlar om rationella val är den mest vanliga synen. Under de senaste åren har mängden information i många beslutssituationer ökat markant och stöd av något slag behövs ofta. Följaktligen har områden som Business Intelligence (BI) och Beslutsstödssystem (BSS) framskridit.

Beslutsanalysverktyg tillhör den senare kategorin och syftar till att fungera som stöd vid beslutsfattande inom företag och organisationer och tillhandahålla mer strukturerat underlag för beslut. Trots en tro på deras potential, så är deras användande begränsat i praktiken. Detta kan delvis tillskrivas det faktum att existerande verktyg är inkompleta för att stödja beslutsprocesser i tillräcklig utsträckning i verkligheten. Dessutom är utvinningen (*eliciteringen*) av indata till beslutsmodellen en viktig komponent i traditionella beslutsprocessmodeller. Trots det faktum att detta involverar ett antal problematiska element och dess vikt för kvalitén av hela processen så förutsätts ofta att användaren själv klarar av att specificera och utföra utvinningen (*eliciteringen*) av indata.

Denna avhandling fokuserar på elicitering av information i beslutsanalysapplikationer och behovet av mer preskriptiva eliciteringsmetoder (realistiska och praktiskt användbara) än vad som finns att tillgå idag. En processmodell som betonar vikten av strukturerad elicitering av adekvata indata genom hela beslutsprocessen föreslås också. För att ytterligare definiera de problematiska aspekterna av elicitering utfördes tre empiriska studier där problem från existerande litteratur studerades vidare. Problemen med att utvinna precisa beslutsdata antyder att användandet av oprecisa värden inom elicitering är en mer realistisk och användbar ansats att sträva efter och kanske än mer viktigt när det gäller utvinning av preferenser. Baserat på teori och resultaten av studierna beskrivna i artiklarna I-III har en eliciteringsmetod för oprecisa utlåtanden och osäkra indata formaliserats och metoden förfinades ytterligare till Cardinal Rank Ordering of Criteria (CROC) vikteliciteringsmetoden som presenteras i artikel VII. CROC-metoden är både kompatibel med en anpassad preskriptiv beslutsmodell fokuserad på en mer strukturerad eliciteringskomponent samt algoritmer för att hantera denna typ av data. CROC-metoden användes och validerades i två riktiga fall (beskrivna i artiklarna V och VI), vilket inte är så vanligt inom beslutsanalys forskning, dvs. det finns relativt få exempel på riktiga fall där sådana tekniker används.

ACKNOWLEDGEMENTS

There are many persons to whom I would like to express my gratitude. First of all, Prof. Mats Danielson for indispensable supervision and support during this work. Prof. Love Ekenberg for taking me on as a Ph.D. student and being an excellent co-supervisor. Further, Anna Ahlström and Ellen Terserus foundation for funding my research up until my licentiate degree. All my friends and colleagues at DSV, not least Prof. Harald Kjellin for constructive input on this work. PUC, Pontificia Universidade Católica do Rio de Janeiro, for the exchange and stay there during the spring of 2006, and especially Prof. Tara Baidya for cooperation and friendliness. Mittuniversitetet for hospitality during my exchange there during the fall of 2006.

My former fellow graduate students within the DECIDE research group, Dr. Karin Hansson, Dr. Jim Idefeldt, Dr. Aron Larsson, Dr. Ari Riabacke, and Dr. David Sundgren for contributing to a stimulating research environment. I would like to express a unique and very special thank you to Dr. Riabacke for outstanding teamwork and making the seemingly impossible possible, Dr. Larsson for valuable cooperation at later stages of this work and Dr. Lisa Brouwers for cooperation early on. I would also like to thank Dr. Brouwers, Dr. Hansson and Dr. Anette Hulth for guidance into the Ph.D. world and our KLMN tradition.

My dearest and oldest friends - Rozana, Sofia and Kristin - for hanging in there, although I have been rather absent during the last few years.

My grandparents, in heaven, I know you would have been extremely proud if you could have been here today.

My parents, for all your support and encouragement during my research, not least all your baby sitting assistance when I needed it the most! My sister Suzanne, for all your encouragement. I am so happy to have you in my life!

The four most wonderful children in the world - Philippa, Mi, Nelly and Maximilian - for all the beautiful words, hugs and kisses that helped me remember what is really important in life! And last but not least, my incredibly supportive and absolutely fabulous husband Ari - you are a great part of this work in so many ways. I could not have done it without you!

I dedicate this work to my whole family - with all my love.

TABLE OF CONTENTS

1. Introduction.....	1
1.1 Problem Background	4
1.2 Problem Focus	7
1.3 Research Focus of Thesis.....	8
1.4 Methodological Issues	9
1.5 Outline.....	10
2. Decision Theories.....	11
2.1 Normative Models	12
2.2 Multi-Criteria Decision Analysis	13
2.3 Descriptive Models	15
2.4 Prescriptive Decision Analysis.....	17
3. Decision Process Outline.....	23
4. Elicitation.....	31
4.1 Probability and Utility Elicitation.....	32
4.2 Weight Elicitation.....	34
4.2.1 Ratio Weight Procedures	35
4.2.2 Imprecise Weight Elicitation	37
4.3 Issues on Elicitation	41
4.4 Approaching Elicitation Prescriptively	43
4.5 Handling Elicitation in Software.....	47
5. Methodology	51
5.1 Conducting and Validating Research within Decision Analysis.....	52
5.2 Validating Elicitation Methods	55
5.3 Methodological Approach of Thesis	57
6. Contributions	63
7. Concluding Remarks	71
7.1 Summary	72
7.2 Further Research	74
References.....	77

LIST OF INCLUDED PAPERS:

PAPER I

Påhlman, M. and Riabacke, A. A Study on Framing Effects in Risk Elicitation. In *Proceedings of the International Conference on Computational Intelligence for Modelling Control and Automation (CIMCA)*, Vienna, 2005.

PAPER II

Riabacke, A., Pålman, M., and Larsson, A. How Different Choice Strategies Can Affect the Risk Elicitation Process. *IAENG International Journal of Computer Science* 32(4): 460-465, 2006.

PAPER III

Riabacke, A., Pålman, M., and Baidya, T. Risk Elicitation in Precise and Imprecise Domains - A Comparative Study, Sweden and Brazil. In *Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA)*, Sydney, 2006.

PAPER IV

Riabacke, M., Danielson, M., Larsson, A., and Ekenberg, L. Employing Cardinal Rank Ordering of Criteria in Multi-Criteria Decision Analysis. In *Proceedings of the 10th International FLINS Conference on Uncertainty Modeling in Knowledge Engineering and Decision Making (FLINS)*, Istanbul, 2012.

PAPER V

Danielson, M., Ekenberg, L., Larsson, A., and Riabacke M. Transparent Public Decision Making - Discussion and Case Study in Sweden. *E-Democracy: A Group Decision and Negotiation Perspective*, Rios Insua, D. and French, S. (eds.), Springer, 2010.

PAPER VI

Riabacke, M., Åström, J., and Grönlund, Å. E-participation galore? Extending Multi-Criteria Decision Analysis to the Public. *International Journal of Public Information Systems* 7(2):79-99, 2011.

PAPER VII

Riabacke, M., Danielson, M., Ekenberg, L., and Larsson, A. Cardinal Rank Ordering of Criteria – Addressing Prescription within Weight Elicitation. In last revision stage, *International Journal of IT and Decision Making*, 2012.

A note should be made that the author has changed surname, and therefore, Pålman, M. is now Riabacke, M.

RELATED PAPERS (not included in thesis):

- Riabacke, M., Danielson, M., Ekenberg, L., and Larsson, A. A Prescriptive Approach for Eliciting Imprecise Weight Statements in an MCDA Process. In *Proceedings of 1st International Conference on Algorithmic Decision Theory (ADT)*, Venice, 2009.
- Brouwers, L. and Riabacke, M. Consensus by Simulation: A Flood Model for Participatory Policy Making. To appear in: *Integrated Catastrophe Risk Modelling, supporting policy processes*, Amendola, A., Ermolieva, T., and Mechler, R. (eds.), Springer, forthcoming spring 2012.
- Danielson, M., Ekenberg, L., Larsson, A., and Riabacke, M. Structured eDemocracy beyond eDiscussion. In *Proceedings of E-challenges*, Istanbul, 2009.
- Riabacke, M., Bohman, S., Danielson, M., Ekenberg, L., Faye, W., and Larsson, A. Public Decision Making Support: A Developing Country Perspective. In *Proceedings of IST-Africa*, Durban, 2010.

CHAPTER 1

1. INTRODUCTION

We are frequently exposed to the concept of decision making in everyday life. Most of the decisions we make do not require much thought, whereas other, more complex matters require extensive deliberation. Most of us would like to believe that the decisions we make are rational and well-founded, but research has shown that this is often not the case (cf., e.g., March, 1994; Shapira, 1995; Simon, 1976). We believe that experience, knowledge and intuition can guide us through the decision making process, but our inherent limitation of processing information in an objective manner and to capture the “whole picture” frequently lead us astray.

Our previous experiences help us navigate the increasingly complex environment we exist in. Adopting shortcuts and automated behaviour is often efficient and sometimes necessary, but these mental shortcuts can also often send us off track. The automated responses that are triggered can instead cause us to make faulty judgments, e.g., when it comes to assessing probabilities of certain events occurring, different judgmental heuristics can cause us to over- or underestimate such figures. One such example is the so called *availability* heuristic, which can cause us to overestimate the probability of an event occurring if something similar has happened recently or media reports on the certain type of event have flourished recently, i.e. events that are available or come to mind easily seem more probable, which is not necessarily true. Another example is that the *framing* of a question, i.e. its presentation, has a strong effect on people’s responses, e.g., simply framing something as half-full or half-empty often cause people to make different decisions, although the two are equivalent in terms of quantity. When the decisions we face are of significant importance, making decisions under such premises seem rather negligent, and a more structured ap-

proach is needed in order for a sufficient analysis to be made prior to action. However, as very few people in society or organizations have been educated in making well-deliberated decisions based on formal methods, this causes problems.

The fact that people often have problems making decisions was early noted within a wide range of areas, and decision making has been an issue of concern for quite some time. During the last few decades, research has increased significantly and has shown that the cognitive limitations of the human mind make it difficult to process the complex, large amounts of information intrinsic in many decision making situations. In fact, people seldom talk about possibilities when speaking about decisions to be made, but more or less always use the term decision *problem*. Much of the work that steers progress within organizations relates to decision making and problem solving (Simon et al., 1986), and consequently, there is a great interest in how decisions are made in these settings. Within organization theory (c.f. March and Simon, 1958; Cyert and March, 1963; March and Olsen, 1976), especially strategic management (c.f. Porter, 1980; Hart, 1992), decision making is central, and the tradition of rationality is especially considered important in systematic approaches to management, such as, e.g., planning and processing. In a broad sense of the term, rational behaviour has to do with reasonable and consistent acts, whereas its meaning in the classical economic literature (cf. von Neumann and Morgenstern, 1947) is to maximize and choose the optimal alternative of all those available to us.

Within research, theoretical developments of decision making have traditionally been divided into the normative and descriptive disciplines. Within the normative discipline, the rational model has been prominent (von Neumann and Morgenstern, 1947; Savage, 1954, Luce and Raiffa, 1957), and models belonging here describe how decision-makers *should* make choices when considering risk. The rational model of decision making is essentially based on the notion that managers systematically gather information in order to objectively analyse it before making a decision (Morrow and Hitt, 2000). However, even though rationality is a desirable trait, the rational model has often been criticized over the years in the behavioural literature concerning its inherent assumptions on cognitive and motivational assumptions (Hart, 1992). As a consequence, the descriptive discipline (cf., e.g., Simon, 1957; Kahneman and Tversky, 1979) has evolved, where models describing how people actually *do* make decisions are in focus. Within organizational settings, this has led to the development of other models (Simon, 1945, 1957; March and Simon, 1958; Cyert and March, 1963; March, 1994), where organizational characteristics, like, e.g., context, societal structures, organization, conflicting or unclear goals, and political activities (con-

flict among stakeholders), cause decision-makers in organizations to depart from rational decision making procedures. However, descriptive models mainly account for actual behaviour and do not provide tools for applied decision making and, e.g., Kirkwood (1997) argues that in order to make decisions strategically, it is more or less a requirement to adopt a structured decision making process. Also, although real decision-makers do not behave like the normative models predict, they might still need and want help (Bell et al., 1988). People do not naturally approach problems in a structured fashion, and the amount of information involved in many decision making situations has increased dramatically in recent years, which complicates matters further. Consequently, the fields of Business Intelligence (BI) and Decision Support Systems (DSS) have advanced, where the former is about providing people with the right information, and the latter has to do with computerized information systems which support business and organizational decision making activities. Yet, when it comes to decision making processes, structured methods are not often applied in real settings, and decision-makers often act on rules of thumb, intuition, and experience instead.

During the last few decades, the field of *decision analysis*, the applied form of decision theory (Raiffa, 1968; Keeney and Raiffa, 1976), has developed as a structured approach to formally analyse decision problems. The field is based on research within several disciplines,¹ such as psychology, mathematics, statistics and computer science. Decision analyses are aimed at helping people make better decisions (Keeney, 2004), and over the years, research on quantitative decision making has moved from the study of decision theory founded on single criterion decision making towards decision support for more realistic decision making situations with multiple, often conflicting, criteria. In particular Multi Criteria Decision Analysis (MCDA) stands out as a promising category within the decision support methods. MCDA can provide decision-makers with a better understanding of the trade-offs involved in a decision, e.g., between economic, social and environmental aspects (criteria). After identifying the primary objective(s) or goal(s) of the decision-maker(s) and the different alternatives (the available courses of action), the possible consequences are analysed mathematically on the basis of the provided input data.

Research within the instrumental part of the decision making process as well as means to support it have developed significantly during the last half century. Still, despite the promising solutions offered today, and a belief in their

¹ See, e.g., early work by psychologists Edwards (1954), who introduced Bayesian analyses to psychology, and Ellsberg (1961), and statisticians Wald (1950) and Savage (1954). In economics, Simon (1955), and March and Simon (1958) have had a great deal of influence, and Markowitz (1952) has contributed to decision theoretic research within finance.

potential to support complex decision making, decision analysis tools are seldom utilized to aid decision making processes in most organizations (March, 1994; Shapira, 1995; Brown, 2006), and people rarely perform formal analysis to complex problems (Keeney, 2004). The number of MCDA applications has increased during the last decade, but behavioural issues have not received much attention within this field of research, yet the identification of such problems and the call for research on behavioural issues have been recognized for a long time (Wallenius et al., 2008). Moreover, current applications provide good support for decision analytical calculations, but lack support for the decision making process itself (French and Xu, 2005). French and Xu (2005) suggest that this functionality is something that needs to be included in MCDM packages in further developments, and Banville et al. (1998) claim that regardless of the progress made within the instrumental dimension of multiple criteria approaches, the under- or non-utilization problem will continue until parallel research on the socio-political context in which these MCDA methods are to be applied is emphasized.

1.1 PROBLEM BACKGROUND

There are successive levels of complexity involved in decision making situations, e.g., the fact that decision-makers rarely possess all the relevant information about a situation and have inherent limitations to process the existing information about a situation adequately. Thus, if the aim is to act “rational” and comprehensive, a systematic approach for information processing and analysis of some kind is needed, especially when the problem at hand is complex, non-repetitive and involves uncertainty.² Accordingly, an explicit, structured decision making process in combination with support of some kind in order to handle all the required data sufficiently and to strengthen the process of making informed business decisions could be a means to meet these criteria.

There are several models for how decision processes should be outlined, but in order to be rational (based on decision theory), it should contain a certain number of key components. In Clemen (1996), the standard model of a decision process, which is the most widely used within the field of decision analysis, is described (see more in Chapter 3. Decision Process Outline). In this model, the *elicitation* of input data to the decision model, which is constructed during the

² An effort to avoid uncertainty has been noted in much organizational decision making (Zhou, 1997), but in spite of efforts to exclude it, it is still an implicit factor which can result in unpredictable behavior.

process, is one of the key components during modelling. However, in spite of the fact that it involves quite a few problematic elements and its importance to the quality of the process as a whole, focus within research has been more on the latter components, such as improving algorithms for evaluations and sensitivity analysis.

Since traditional decision analysis tools use formal rules for mathematical calculations, which in turn are based on values that are often, at least partly, elicited from experts and/or decision-makers, the implicit assumption that people are able to assess the values required for analysis accurately seem precarious. Such values, like knowledge, beliefs, and preferences, have more imprecise semantics in people's minds and are not naturally represented in probabilistic form. Several studies have pointed at the difficulties in expressing such values with numeric preciseness (cf., e.g., Shapira, 1995; Corner and Corner, 1995; Riabacke et al., 2006a), and the difficulty of eliciting precise probabilities and utilities in real-world settings is a widely discussed practical difficulty involved in the use of expected utility models for decision making (Fox, 2003; Fischhoff et al. 1983). Numerous empirical studies have shown inconsistencies with the predictions of the rational model in human behaviour (cf., e.g., Ellsberg, 1961; Edwards, 1968; Slovic et al., 1988; Tversky and Kahneman, 1974, 1981, 1986), and, e.g., the fundamental assumption of *procedure invariance* within normative theories, i.e. that normatively equivalent methods for eliciting values should give rise to the same results, does not always hold and elicited values can be heavily dependent on the method of assessment (cf., e.g. Tversky and Kahneman, 1981; Lenert and Treadwell, 1999; Pöyhönen and Hämäläinen, 2001; Pählman and Riabacke, 2005). The applicability of the generated results in decision analysis is often reliant on the quality of the input data, which consequently constitutes a problem if elicited data can vary depending on the method used.

The identified discrepancy between normative and descriptive theories have led to an increased attention for *prescriptive* approaches to decision making (cf., e.g., Bell et al, 1988; French and Rios, 2000; Keeney, 1992a), which is focused on merging the two main disciplines within decision making into a more practically useful approach for handling decision problems and help decision-makers solve real problems. Decision analysts talk of *prescriptive decision analysis* as a more pragmatic approach than what the normative theories suggest, and Brown and Vari (1992, p. 35) state that prescriptive science "supports the development of inference and decision aids which are to be used and useful, i.e. aids which humans can effectively supply the input to and use the output from". In line with prescriptive ideas, there is a need to complement the tools supporting deci-

sion analytical calculations with support for the decision making process itself, i.e. to aid users during the different steps of the decision making process.

Brown (1989) states that the low level of attention given to prescriptive decision support research in real settings has contributed to the limited practical impact that formal decision aid has had on decision making in business and government. The explicit use of quantitative decision models to support and improve decision making activities remains modest in real settings and the current use is often to justify or communicate a choice rather than a support to make the choice in the first place (Brown, 2006). There is reason to believe that without the involvement of actual decision-makers and research on the process surrounding the developed models, strategies and techniques, the utilization of these tools as a self-evident aid in real decision making processes will not increase. Another explanation to their limited usage within organizations today could be that they are too demanding in terms of required time (especially the first time of usage) and effort. As, e.g., Keeney (2004, p. 194) points out, "we all learn decision making by doing it", and he states that almost everyone could make better decisions, but we need training (which is something very few people have ever had). Research has shown that people are more likely to deal with information in a controlled fashion when they are motivated and capable to analyse it carefully (Cialdini, 2009). Thus, one conclusion is that there is a need to better communicate the benefits of using decision support tools. Moreover, many decision problems have large outcome spaces, making the representation and elicitation of preferences and beliefs for all outcomes a costly venture in terms of time and cognitive effort. However, even in situations where the outcome space is manageable, there is a need for elicitation methods better adapted for real usage, since part of the attraction of using a decision analysis tool to support the decision process is reliant on the applicability of generated results. Suggested techniques for elicitation is to a great extent a matter of balancing the retrieved quality of the elicitation with the time and cognitive effort demand on the users for eliciting all the required values. A more prescriptively useful approach to adopt within elicitation would be to allow for greater imprecision when possible, both to accommodate human capabilities as well as to reduce the cognitive effort and time required. Additionally, decision analysis tools deviate too much from current practice within decision making as their use (if employed) is mainly as a tool for calculation and analysis and not as more complete support for decision-makers throughout their decision making processes (French and Xu, 2005).

In summary, there is a need for decision analysis tools that offer support at earlier stages of the decision making process for decision-makers in real-life

situations. A note should be made, that today, the decision support systems that are widely used in the business world are user friendly and often employ simple techniques like spreadsheets (Wallenius et al., 2008).

1.2 PROBLEM FOCUS

Decision making activities in organizations seldom follow systematic approaches for information processing and analysis (March, 1994; Shapira, 1995; Keeney, 2004), although several models for decision making, and support for such processes, have been suggested. A promising approach is decision analysis, which is focused on aiding decision processes by providing models for structuring problems and applying formal models of rational choice in order to guide the processes and improve their outcome. However, current software applications mainly provide good support for decision analytical calculations, and lack support for the decision making process itself (French and Xu, 2005), and the identified gap between ideal and real behaviour (see more in Chapter 2. Decision Theories) has been identified as a potential obstacle to the employment of these models in real settings. Yet, if the rationality trait so often aimed at in business settings is to be aspired to, the use of a more practically oriented prescriptive approach is recommended as more suitable for real settings (French, 1995b; French and Rios Insua, 2000; Keeney, 1992b) to help structure the decision process, bring understanding to the decision faced, keep focus on the key issues involved, and provide more informed decision material to base decisions on.

If we want to employ decision analysis in real settings, there is often a need to elicit input data from experts, decision-makers and/or other stakeholders (cf., e.g., Clemen, 1996; Keeney, 2004), and as pointed at in the previous section, elicitation of such decision data involves quite a few problematic elements. Several components of the standard decision making model (cf., Clemen, 1996) have been quite extensively studied, like the modelling of the problem in decision trees and the development of more sophisticated algorithms for evaluation, but research on the elicitation component within this model has, in comparison, received less attention, and there is still a great need for improvements (cf., e.g., Wallenius et al., 2008). Although elicitation has been an area of concern for quite some time, there are still no generally accepted methods, and the process of eliciting adequate quantitative information from people is still one of the major challenges facing research within this field (Fox, 2003). Without an adequate elicitation component as part of the cyclic sequence of components constituting the decision process model, the subsequent steps in the cycle may not be utilized

sufficiently as the quality of this step is of importance for calculations in later stages. Moreover, Keeney (2004) notes that in many decision problem situations, a complete analysis of the problem is not needed as an adequate elicitation process could be sufficient for addressing the problem.

1.3 RESEARCH FOCUS OF THESIS

In accordance with the issues highlighted in the previous section, this thesis addresses the problem of eliciting adequate input data, such as beliefs and preferences, and the lack of elicitation methods better suited for real decision making processes, i.e. more prescriptively useful. In current decision analysis applications, the specification and execution of the elicitation process is often left to the discretion of the user, which is paradoxical if users are incapable of quantifying their preferences with the degree of precision necessary and providing the input required by these applications. This prevalent requirement for numeric precision within elicitation procedures poses problems in practical applications, but instead of undermining the usefulness of normative theories, e.g., Raiffa (1994) suggests the need to modify them to include cognitive concerns. Many of the current decision analysis tool deficiencies would be alleviated if elicitation of inputs were more effective (Brown, 2006).

Thus, there is a need to allow for a wider spectrum of procedures for handling both qualitative and quantitative aspects within elicitation, and to allow for greater imprecision by the experts and decision-makers as many of the required inputs are subjective in nature and are therefore not naturally represented by the formats most methods are designed for. Subjectivity is inherent in all decision making, especially in the choice of criteria within MCDA and the relative “weight” given to these criteria (Belton and Stewart, 2002). This latter aspect of MCDA has been the focus of extensive debate regarding procedures for eliciting weight information (*ibid.*), and is often considered the most cognitively demanding step of the MCDA process (Morton and Fasolo, 2009). Since most problems in reality are of multi-criteria character, this part of MCDA is of particular interest to study. Consequently, the focus herein is to:

- 1) Discuss elicitation problems and suggest measures to reduce such effects.
- 2) Focus in particular on the elicitation of weights within MAVT/MAUT models (see more in Section 2.2 Multi-Criteria Decision Analysis)

- 3) Allow for more impreciseness and more noisy inputs within elicitation, and
- 4) Propose a weight elicitation method fitting for a prescriptive method, i.e. find a more practically useful method suitable for the employment in a prescriptive decision making process context.

In line with prescriptive ideas, we assume that the decision process starts with the identification of a problem, iterates a cycle of steps (or key components), and eventually results in basic decision data that *can* be used for making a decision, but serves as a guide rather than a definite commitment to action.³ From a prescriptive position, the alternative with the highest expected value must not necessarily be the one chosen (Keeney, 1992b). Keeney (1992b) states that one needs to interpret the implications of the model along with considerations not explicitly in the model, which in turn could result in the choice of another alternative (as this could be considered optimal for the real world choice) than the one with the expected utility. This can be the case in organizational settings, where other factors could cause another alternative to be considered (or chosen) instead.

1.4 METHODOLOGICAL ISSUES

In the previous sections, a number of problematic elements with the application of rational models for decision making have been pointed at, not least the need for more research on prescriptively useful elicitation methods and support for decision-makers during this important step, which is in focus in this thesis. Taking a prescriptive stand when studying decision making processes and decision analysis yields a multidiscipline research field as prescriptive research lies in the borderland of the two traditional approaches of normative and descriptive decision theory, and deals with how to benefit from the application of formal methods of decision making in real settings. Consequently, research within this thesis has been conducted under different scientific paradigms, and a combination of research methods has been used.

At the onset of the research, literature reviews within normative, descriptive, and prescriptive theory were conducted in order to find combined approaches to be used in the research. Initially, a quantitative approach was

³ The traditional definition of a decision process usually encompasses a definite commitment to action.

adopted in order to yield additional information regarding people's choices when presented with the type of questions commonly used within utility elicitation. Qualitative methods were thereafter used to strengthen these results and bring more understanding to the phenomenon and finding solutions to problems observed. Instead of confirming or contradicting hypotheses (as in positivist research), the approach was more in line with the approach of design research, and the findings were instead iterated back to further studies of choice behaviour within this realm resulting in recommendations on how to reduce descriptive violations of normative rules.

At later stages of the research, the findings from quantitative and qualitative studies in combination with theoretical reviews within the domain were used to propose a more prescriptively useful weight elicitation method, which was also tested in case studies. Since few evaluations in real settings are conducted (although prescriptive scientists advocate such an approach, cf., e.g., Brown, 1989), external evaluation of the method was carried out in the case studies.

1.5 OUTLINE

In Chapter 2, the theoretical background of decision theory is accounted for, starting with the normative and descriptive theories, and thereafter the prescriptive approach is described. The prescriptive approach is the main focus of the chapter, since the aim of this thesis is to study elicitation within this realm and propose an elicitation method fitting for a prescriptive method. Also, the decision process is of importance to elicitation as elicitation methods need a context. Therefore the prescriptive decision process outline is illustrated in Chapter 3 and a modified model emphasizing elicitation more explicitly is suggested. Thereafter, in Chapter 4, the theoretical background of elicitation is depicted as well as issues on current elicitation practice and suggestions on ways to handle them in software applications are suggested. In Chapter 5, the methodological approach of decision theory in general, but more specifically, the methods used in this thesis are discussed. Subsequently, the contributions of this thesis are summarized in Chapter 6, and finally, conclusions and further research are discussed in Chapter 7.

CHAPTER 2

2. DECISION THEORIES

This chapter describes a theoretical background of decision theory, with a focus on the prescriptive approach, since the centre of attention of the thesis is on elicitation within this realm. Research on decision theory has produced a body of theories and methods for structured decision making when there is risk and uncertainty involved. These theories have influences from a number of different disciplines, like economics, statistics, philosophy, computer science, psychology, behavioural science and artificial intelligence, resulting in different views on rational behaviour. There is a traditional distinction between *normative* and *descriptive* models of decision making, and the gap between the two approaches has been an issue of contentious debate over the years (cf., e.g., Cohen, 1981; Kahneman et al., 1982; Simon et al., 2011). The classic approach is strict and precise but less intuitive and more demanding to use, whereas the more qualitative ones are ad hoc but can be more flexible and easier to understand. The purpose of an analysis is to bring understanding (French, 1995a), and decision analysis, the applied form of decision theory, is aimed at helping people make better decisions (Keeney, 2004). The conflict between classical probability theory and more qualitative approaches within decision analysis have led to a more practically oriented discipline, namely the prescriptive approach.⁴ Prescription aims to build from strengths of the formal approach while compensating for people's weaknesses.

⁴ A note should be made that in earlier literature the term "prescriptive" was generally used as a synonym for "normative" (French, 1995b).

2.1 NORMATIVE MODELS

Normative decision theory (often referred to as the theory of rational choice) has a long history and is concerned with how rational actors *should* behave, i.e. how to make rational decisions. The normative rule of expected value states that the most rational decision to make is to choose the alternative with the highest expected value. This value is calculated by multiplying the probability of an outcome with the value of the outcome itself. However, most decisions regard a one-time event (or an event that is not repeated many times) and involves risk, and thus, the principle of maximizing the expected value is not enough. The most well-known theory of rational choice behaviour under risk and uncertainty is the *expected utility theory*⁵ (von Neumann and Morgenstern, 1947), where the objective is to maximize the expected utility (EU). The utility can be described as the value or worth a decision-maker relates to a certain outcome, and utility functions are used to describe people's attitude to risk. There are a number of axioms that need to be fulfilled in order to be rational according to the theory (cf., e.g., Luce and Raiffa, 1957, and Fishburn, 1970, for reviews of utility theory and axioms), and a rational decision-maker should always choose the option (among all possible options) with the optimal combination of probability and utility. Thus, these values have to somehow be assessed for each possible option. While probability distributions are collective in a model, utility functions are individual to the user. Normative theories assume that people are able to make the required assessments of probabilities and utilities accurately. However, over the years the axioms of the expected utility theory have often been violated in experimental research and revisions of the original theory have been made, both as extensions as well as alternative theories.

One of the most famous variations of the von Neumann and Morgenstern theory is the subjective utility theory (SEU), introduced by Savage in 1954. In contrast to the original EU theory, the SEU theory uses the decision-maker's personal (subjective) beliefs of the probabilities of the possible outcomes of an uncertain event. Before the SEU theory, probabilities in the expected utility theory had been handled as objective probabilities, based on relative frequency. However, the idea of subjective probabilities had already been published by Keynes (1921) and Ramsey (1931). In Ramsey (1931), a set of axioms for choice under uncertainty in which probabilities were subjective were presented (also expressed by de Finetti, 1931, 1937, in a similar fashion) as well as a

⁵ The expected utility theory is often discussed in the literature as one unified theory, but there are extensions to the original theory as well as options to it, and there is no single accepted utility theory.

demonstration on how to infer subjective probabilities and utilities from preferences between different gambles. The subjective view on probability is particularly important in cases where an objective probability cannot be established in advance or when it is a one-time event (Plous, 1993), and Bell et al. (1988) state that in most cases decision problems cannot be analysed using entirely objective probabilities.

It should be noted that the published work by Daniel Bernoulli in 1738 marked the beginning of the concept of utility. His solution to the St. Petersburg paradox has been scrutinized by numerous researchers over the years (cf., e.g., Samuelson, 1977) and involved two main ideas. Firstly, he argued that the proper decision criterion for an alternative involving risk was the expected utility instead of the expected value, and secondly, that people's utility from wealth can be described by a logarithmic utility function as opposed to a linear, an observation called diminishing marginal utility.

One of the most notable early conflicts between the EU theory and human reasoning to be experimentally exposed was demonstrated by Allais in 1953. His Allais' paradox is a counterexample to the independence axiom, which illustrates how real behaviour contradicts the assumptions made in the EU theory. Many other experiments demonstrating inconsistencies in human behaviour with regard to the assumptions made in the EU theory have been noticed (cf., e.g., Ellsberg, 1961; Edwards, 1968; Kahneman and Tversky, 1979). For a summary of observed experimental violations of the EU theory, cf., e.g., (Machina, 1983). Supporters of the rational model maintain that if people make inaccurate assessments, these are attributed to random (unsystematic) errors.

2.2 MULTI-CRITERIA DECISION ANALYSIS

While solving problems, people often consider different features of the alternatives by taking different aspects of importance to them into account (choice criteria). Humans have inherent limitations when processing much information at one time, and the more criteria to consider, the more difficult. When solving such problems, we make errors as well as use simplifying strategies to adapt the problem to our own capabilities. As a result, research on quantitative decision making has moved from the study of single criterion decision making towards decision support for more realistic decision making situations involving multiple, often conflicting, criteria. Multi-criteria decision analysis (MCDA) is a category within decision support methods and can provide the decision-makers with a better understanding of the trade-offs involved in a decision, e.g., between eco-

nomic, social and environmental aspects (criteria). Current research is mostly concentrated on providing models to support the structuring of the problem in order to increase understanding and identify possible problematic elements. Furthermore, the output from these models should not be interpreted as the solutions to the problems, but rather give a clearer picture of the potential consequences of selecting a certain course of action. In this context, the decision-maker is assumed to be an agent⁶ who chooses one alternative (or a subset of alternatives) from a set of alternatives (typically consisting of collections of choices of moderate size)⁷ that are being evaluated on the basis of more than one criterion.

Multi Attribute Value Theory, MAVT, and Multi Attribute Utility Theory, MAUT (Keeney and Raiffa, 1976; von Winterfeldt and Edwards, 1986) together with AHP (see below) are the most widely used MCDA methods in practical applications. The relative importance of each criterion is assessed as well as value functions characterizing the satisfaction of the alternatives (according to the decision-maker) under each criterion, and thereafter the overall score of each alternative is calculated. The difference between the two is that MAVT is formulated such as to assume that outcomes of the alternatives are known with certainty, whereas MAUT explicitly takes uncertainty (relating to the outcomes) into account (and uses utility functions instead of value functions). However, in many practical situations, it is hard to distinguish between utility and value functions elicited with risky or riskless methods due to factors such as judgmental errors and response mode effects (von Winterfeldt and Edwards, 1986). Moreover, in many applications, using simple value functions in combination with sensitivity analysis provide essentially the same results and insights (Belton and Stewart, 2002). Most MAUT methods contain the following five steps:

1. Define the alternatives and the relevant attributes (criteria).
2. Evaluate each alternative separately on each attribute, i.e. the satisfaction of each alternative under each criterion represented by a value/utility function.
3. Assess the relative importance of each criterion, i.e. assign relative weights to the attributes.

⁶ The decision making agent can be an individual or a group that agrees to act in uniform according to the equivalent rational decision making process as would be followed by an individual (Wallenius et al., 2008).

⁷ This is in contrast to optimization problems where feasible sets of alternatives usually consist of infinitely many alternatives.

4. Calculate the overall score of each alternative by aggregating the weights of the attributes and the single-attribute evaluations of alternatives into an overall evaluation of alternatives.
5. Perform sensitivity analyses on the model and make recommendations.

Examples of other MCDA methods than the MAVT/MAUT approach are the Analytic Hierarchy Process, AHP (Saaty, 1980), which is similar to MAVT, but uses pairwise comparisons of alternatives (utilizing semantic scales) with respect to all criteria, or outranking methods based on partial ordering of alternatives, where the two main approaches are the ELECTRE family of methods (cf., e.g., Roy, 1996), and PROMETHEE (cf., e.g., Brans et al., 1986). Moreover, Fuzzy set theory (introduced by Zadeh in 1965) is an attempt to model human perceptions and preferences more genuinely, but has some practical problems, e.g., with visualizing an operational elicitation process for the required values (Belton and Stewart, 2002). The Measuring Attractiveness by a Categorical Based Evaluation Technique, MACBETH (Bana e Costa and Vansnick, 1994), uses pairwise comparisons (like the AHP method) to express strength of preference (on a semantic scale) for value increments in moving from performance level p to level q .

Different software systems implementing MCDA have been suggested over the years. MAVT techniques have been implemented in, e.g., V.I.S.A (Belton and Vickers, 1988), HiView (Barclay, 1984), which supports the MACBETH pairwise comparison approach to elicitation (Bana e Costa and Vansnick, 1994), DecideIT (Danielson et al., 2003) and GMAA (Jiménez et al., 2006). The latter two both allow the use of interval value and weight statements. The AHP method proposed by Saaty (1980), is implemented in several applications, amongst which EXPERT CHOICE (expert choice, 2010) is probably the most widely used. HIPRE 3+ (Hämäläinen and Lauri, 1995) and Logical Decisions are examples of software packages supporting both MAVT and AHP methodologies. Decision Lab 2000 (cf. Geldermann and Zhang, 2001, for a review) is based on outranking methods, such as PROMETHEE (Brans et al., 1986).

2.3 DESCRIPTIVE MODELS

Over the years, research on decision making has gone back and forth between theory and observation and other, more descriptive models of choice behaviour, i.e. models describing how people *actually* make decisions, have been proposed. Within the psychological discipline, the dominating viewpoint is that people

make decisions not only based on how they judge the available information, but that they are also influenced by more subconscious factors in the interactive process. One of the early critics of the SEU model of rational choice was Simon (1957), who argued that complete rationality was an unrealistic assumption in terms of human judgment. Instead, he proposed a more realistic approach to rationality, called *bounded rationality*, which takes the inherent limitations humans have when processing information into account. The principle of *satisficing* can be applied without highly sophisticated skills in reasoning and evaluation. It proposes that people attempt to find an adequate solution rather than an optimal, and choose the first course of action that is satisfactory on all the important attributes. Simon also coined the terms substantive and procedural rationality, where the former has to do with the rationality of a decision situation, i.e. the rationality of the choice made (which is what economists have focused on), whereas procedural rationality considers the rationality of the procedure used to reach the decision (has been more in focus within psychology).

Prospect theory (Kahneman and Tversky, 1979) is perhaps the most influential of the descriptive models, and can be perceived as an attempt to bring psychological aspects on reasoning into economic theory. In prospect theory, utility is replaced by value (of gains and losses) and deviations from a reference point. The value function is S-shaped and passes through the reference point. It is asymmetric (steeper for losses than for gains) and implies that people are loss averse, i.e. the loss of \$100 has a higher impact than the gain of \$100. Moreover, it suggests that people, in general, are risk averse when it comes to gains and risk seeking when it comes to losses, and systematically overweight small probabilities and underweight large ones. Prospect theory also expects preferences to depend on the framing of the problem, i.e. how the problem is formulated. People are inclined to simplify complex situations, using heuristics and frames when dealing with information (Kahneman et al., 1982). Regret theory (Loomes and Sugden, 1982; Bell, 1982) has been offered as an alternative to prospect theory. In short, regret theory adds the variable regret to the regular utility function and suggests that people avoid decisions that could result in regret. Other problems with the application of normative theories to decision problems and how people actually make judgments have been accounted for by, e.g., March and Olsen (1976), who coined the term *garbage can decision making*, Slovic et al. (1988), and Shapira (1995). The reality of human decision making and the difference (from normative models) in how decision rules are used by real decision-makers have resulted in adaptations of original rational choice theories to the introduction of the limited rationality concept (March, 1994).

Over the last several decades, numerous models of decision making within organizational settings have been proposed from a number of different theoretical perspectives, and Hart (1992, p. 327) describes the result as “a bewildering array of competing or overlapping conceptual models”. In reality, decision making in organizations seldom follow rational decision making processes. March (1997) states that according to rational theory, decision making processes are based on four parts, namely:

- 1) Knowledge of alternatives (a set of alternatives exist).
- 2) Knowledge of consequences (probability distributions of the consequences are known).
- 3) Consistent preference order (the decision-makers’ subjective values of possible consequences are known and are consistent).
- 4) Decision rule (used for selection among the available alternatives based on its consequences for the preferences).

March (1997) also declares that the structure is understandable and that the core ideas are flexible, but that each of these four main parts (and the assumptions made regarding them in the rational model) have problems when applied in organizational settings. Bounded rationality (Simon, 1957) limits the rationality of identifying all possible alternatives as well as all their consequences. Moreover, when considering a series of choices in order to establish preference consistency, research has shown that this has been notoriously hard to determine (March, 1997).

2.4 PRESCRIPTIVE DECISION ANALYSIS

“The art and science of elicitation of values (about consequences) and judgments (about uncertainties) lies at the heart of prescriptive endeavours.” (Bell et al., 1988, p. 24)

Although real decision-makers do not behave like the normative models declare, they might still need and want help (Bell et al., 1988). Thus, if the aim is to act “rational” and comprehensive, a systematic approach for information processing and analysis of some kind is needed, especially when the problem at hand is complex, non-repetitive and involves uncertainty. The identified gap between normative (cf., e.g., von Neumann and Morgenstern, 1947; Luce and Raiffa, 1957) and descriptive (cf., e.g., Simon, 1957; Kahneman and Tversky, 1979) theories,

suggests that a prescriptive approach to a decision making process would be valuable (cf., e.g., Bell et al., 1988; French and Rios Insua, 2000; Keeney, 1992a).

In 1966, Howard coined the term *decision analysis* as a formal procedure for the analysis of decision problems. It is the applied form of decision theory, and is particularly useful for dealing with complex decision making involving risk and uncertainty. Major advances within the discipline have been achieved by Raiffa (1968), and extended to include multiple objectives by Keeney and Raiffa (1976). Decision analysis is a structured way of modelling decision situations in order to explore and increase understanding of the problem, possible problematic elements, and improve the outcome of the decision process. After identifying the primary objective(s) or goal(s) of the decision-maker(s) and the different alternatives (the available courses of action), the possible consequences are analysed mathematically on the basis of the provided input data.

The discrepancy between theory and real behaviour (accounted for in the previous sections) is the very heart of prescriptive interventions (Bell et al., 1988), and decision analysts talk of *prescriptive decision analysis* as a more pragmatic approach than what the normative theories suggest. It can be described as “the application of normative theories, mindful of the descriptive realities, to guide real decision making” (French and Rios Insua, 2000, p. 5). Prescriptive decision analysis is focused on merging the two main disciplines (the normative and the descriptive) within decision making into a more practically useful approach for handling decision problems, and help decision-makers solve real decision problems. The prescriptive approach aims at obtaining the required components for analysis in a structured and systematic way with a great deal of human participation and awareness of the descriptive realities (von Winterfeldt and Edwards, 1986). Brown and Vari (1992) state that much of the work within the descriptive discipline is of substantive importance for prescriptive decision aiding, such as the work on cognitive illusions and human limitations (Kahneman et al., 1982), which can be corrected (or reduced) by decision aids. Moreover, prescriptive analysis can be seen as the application of reason to real-world decision problems, and the employment of an underlying formal model can increase knowledge about the problem at hand and create incentives to acquire as accurate information as possible (Larsson, 2008). In essence, prescriptive decision analysis is about the applicability of decision analysis to real problems in real contexts (and by real decision-makers), and French (1995b, p. 243) brands the term as the usage of “normative models to guide the evolution of the decision-makers’ perceptions in the direction of an ideal, a consistency, to which they aspire, recognizing the (supposed) limitations of their actual cognitive processes.” In the end, the prescriptive approach deals with the tailoring of the decision

analysis process for the specific problem, context and decision-maker(s) at hand, and the theoretical and operational choices made provide the means by which the analyst helps guide the decision-maker(s) through the analysis (Keeney, 1992b). The main criteria for evaluating prescriptive models are usefulness (Keeney, 1992b) and pragmatic value (Bell et al., 1988), and such models should, thus provide decision-makers with suitable assistance in order to improve their decision making.

Keeney (1992b) stresses that, unlike normative and descriptive theories, the focus of prescriptive decision analysis is to address one decision problem at a time, and is not particularly concerned with whether the axioms utilized to support the analysis for the given problem are appropriate for *classes* of problems (typically the focus of descriptive theories) or *all* other problems (the focus of normative theories). Keeney (1992b) distinguishes between decision theories through the following classification (Table 1), where we can also see the differences in criteria for validation.

Theories	Domain	Criterion	Judges of Theories
Normative	All decisions	Correctness	Theoretical sage
Descriptive	Classes of decisions	Empirical validity	Experimental researchers
Prescriptive	Specific decisions	Usefulness	Applied Analysts

Table 1. Keeney's (1992b) classification of the three categories of decision theories.

Fischer (1989) argues that unless a clearly superior alternative to the EU model is available (and consensus is established among decision analysts regarding a new alternative), there is a danger in abandoning it, since the concept of rationality will lose much of its appeal (if rationality becomes a matter of taste) and the field of decision analysis will no longer be coherent. For many decision problems, the expected utility axioms provide a very good basis for decision analysis (Keeney, 1992b; Howard, 1992; Fischer, 1989), but tackling of the unique and complex about a decision problem may require the use of complementary rules as well as a wider spectrum of risk attitude modelling.

Complexities about a problem may involve such factors as significant uncertainties, multiple objectives, multiple stakeholders, and multiple decision-makers. The choice of axioms to guide the prescriptive analysis is a problem facing the analyst trying to help the decision-maker(s), where the overall objective is to provide a foundation for quality analysis (Keeney, 1992b). These axioms should be practical in the sense that it is feasible to conduct an analysis based on them, and the information required to implement them must be attainable and possible to assess in a logically sound and consistent manner. An influential approach to successful prescriptive analysis is the value-focused thinking advocated by Keeney (1992a). Keeney argues that the values of the decision-makers should be understood before the formulation of alternatives takes place in order for the decision-makers to be more creative and think more widely about possible courses of actions. This is in contrast to alternative-focused thinking where you initially find the available alternatives and thereafter evaluate them. However, Keeney recognizes that the ideal of value-focused thinking is hard to achieve, and many decision problems we face initially arise from a set of alternatives from which we must choose (French, 1995b).

Any decision analysis is essentially a model of a specific decision situation, a simplification of reality which includes significant aspects of the problem and lends insights about these aspects (Keeney, 1992b). Phillips (1984) uses the term *requisite model* to describe a judgmental model, which is appropriate when neither normative nor descriptive models are sufficient to capture the value judgments and their relative importance. In a requisite model, everything required to solve the problem is represented in the model or can be simulated by it. Such models are construed through the interaction between specialists, who contribute to the form and assist with the encoding of the content to the appropriate form, and problem owners who provide content. The *form* of the model is decision-theoretic, both in its structure (decision trees, influence diagrams etc.) and its generic components (events, outcomes, consequences, attributes etc.). The *content* of the model (probabilities, utilities, weights etc.) is a result of the participants' understanding of the problem, which evolves through the course of modelling (Phillips, 1984). Phillips mentions three aspects of simplification in such models: 1) elements of reality that are not expected to contribute significantly to the problem solving are omitted, 2) complex relationships among elements of the social reality are approximated, and 3) distinctions in either form or content regarding social reality may be blurred in the model, e.g., one may choose not to make distinctions between present and future worth of a product. The prescriptive decision analysis process (see more in Chapter 3. Decision Process Outline) is cyclic with iterations through the steps of modelling values,

identifying alternatives, evaluating, reflecting and possible remodelling of values, modifying or identifying new alternatives, re-evaluating, and so on (cf., e.g., French, 1995b). During prescriptive decision analysis, perceptions change and evolve, and the representation of these perceptions should not be static (French and Rios Insua, 2000). The perceptions of the decision-maker(s) evolve because of the analysis, and it is important to see the modelling process involved in representing these perceptions as creative, dynamic and cyclic (French and Rios Insua, 2000; French, 1995b). *Requisite modelling* is the term used by Phillips (1984) to describe this approach to modelling and a model is *requisite* when it is sufficient for the inference or decision faced. This is in contrast to the static view, too often taken in decision analysis, where all of the judgments of the decision-maker(s) are taken as fixed and binding from the outset of the analysis (French and Rios Insua, 2000).

The modelling and the selection of the appropriate formal rules are only part of the assumptions necessary to approach the problem prescriptively. An important aspect to consider is how to assess or elicit the required information and values in order to apply the rules in a prescriptive manner. "The art and science of elicitation of values (about consequences) and judgments (about uncertainties) lies at the heart of prescriptive endeavors" (Bell et al. 1988, p. 24). The techniques and methods used for elicitation must be practicable, and should not require too many inputs from the decision-maker(s) (French and Rios Insua, 2000). Fischer (1989) points at three fundamental problems that need to be confronted when attempting to develop prescriptive science: (1) reference effects (which lead to systematic violations of the independence principle of the EU model), e.g., people's tendency to be risk-averse for gains and risk-seeking for losses as well as weigh losses more heavily than gains (Kahneman and Tversky, 1979), (2) framing problems, i.e. that formally equivalent ways of framing (describing) decision problems can highly influence people's choices in reality, and (3) different outcomes resulting from strategically equivalent assessment procedures for eliciting preferences. Prescriptive analysts must, thus, be attentive to the descriptive realities of human behaviour and the common mistakes people make when eliciting decision data as the applicability of generated results often relies on the quality of input data (see more in Chapter 4. Elicitation). The prescriptive analysts must learn how to elicit adequate judgments from decision-makers and make sense out of them (Bell et al., 1988). Moreover, many decision analysts believe that the insights that can come to light during the elicitation process can be more valuable than what is done with the elicited values after elicitation, and it is thus an important ingredient in a successful prescriptive decision analysis process. The specialist (or analyst) has a dual role,

both to facilitate the work and keeping the decision-maker(s) task oriented as well as to contribute to the aspects of the task concerned with model form, but not its content (Phillips, 1984).

Conclusively, when decision analysis applications are used to aid prescriptive decision making processes, additional demands are put on these applications to adapt to the users and context. French and Rios Insua (2000) state that prescriptive methodologies for decision analysis should aim to be satisfactory regarding such aspects as:

- i. *Axiomatic basis.* The axiomatic basis should be acceptable to the users, and they should want their decision making to reflect the ideal behaviour encoded in the set of axioms used for analysis.
- ii. *Feasibility.* The techniques and methods used must be practicable, which suggests that the elicitation of decision data from the users must be feasible (the number of required inputs from the users should be acceptable) and results must be intelligible to the users. As already mentioned in previous sections, the descriptive realities of human behaviour also adds demands on elicitation processes to reduce the cognitive load on decision-makers as well as to aim to eliminate biases that have been documented in behavioural research.
- iii. *Robustness.* The sensitivity to variations in the inputs should be understood, e.g., if the analysis results rely heavily on certain inputs the decision-makers should be aware of this and be able to reconsider judgments made.
- iv. *Transparency to users.* The users must understand the analysis procedure and find it meaningful.
- v. *Compatibility with a wider philosophy.* The model used for analysis must agree with the decision-makers' wider view of the context (and the world). The model must be requisite, thus, the application must provide for interactivity and cyclic modelling possibilities in order to reach the goal of compatibility.

CHAPTER 3

3. DECISION PROCESS OUTLINE

In this chapter, conventional decision processes are presented. Thereafter, elicitation within this context and the need to model decision processes in a way that makes clear the critical role played by elicitation in these processes are discussed.

A decision process can be described as a “set of actions and dynamic factors that begins with the identification of a stimulus for action and ends with the specific commitment to action” (Mintzberg et al. 1976, pp. 246). There are several proposals for how decision processes should be designed. Ideally, a decision process should contain a certain number of key components in order to be called rational (based on decision theory). In Clemen (1996), the standard model of a decision process, which is the most widely used within the field of decision analysis, is outlined as the following cycle of steps (see Figure 1). According to this model, 1) the first step is that the decision-maker identifies the decision situation and thereafter needs to understand his or her objectives in this context. Clemen states that although finding problems to solve is usually not a problem, the specification of the problem at hand may sometimes be inaccurate, thus resulting in “error of the third kind” (treating the wrong problem). The next step of the decision analysis process 2) is the creation of alternatives. Step 3) involves the structuring and modelling of the problem (problems can often be decomposed into more manageable parts). The model of the problem contains a) a model of the problem structure (often in the form of influence diagrams or decision trees), b) a model of uncertainty (using probabilities), and c) a model of preferences (assessments of the decision-maker’s view on the relationship among multiple objectives as well as his or her values of the different outcomes). Step 4) contains the mathematical evaluation of the problem (i.e. finding the best

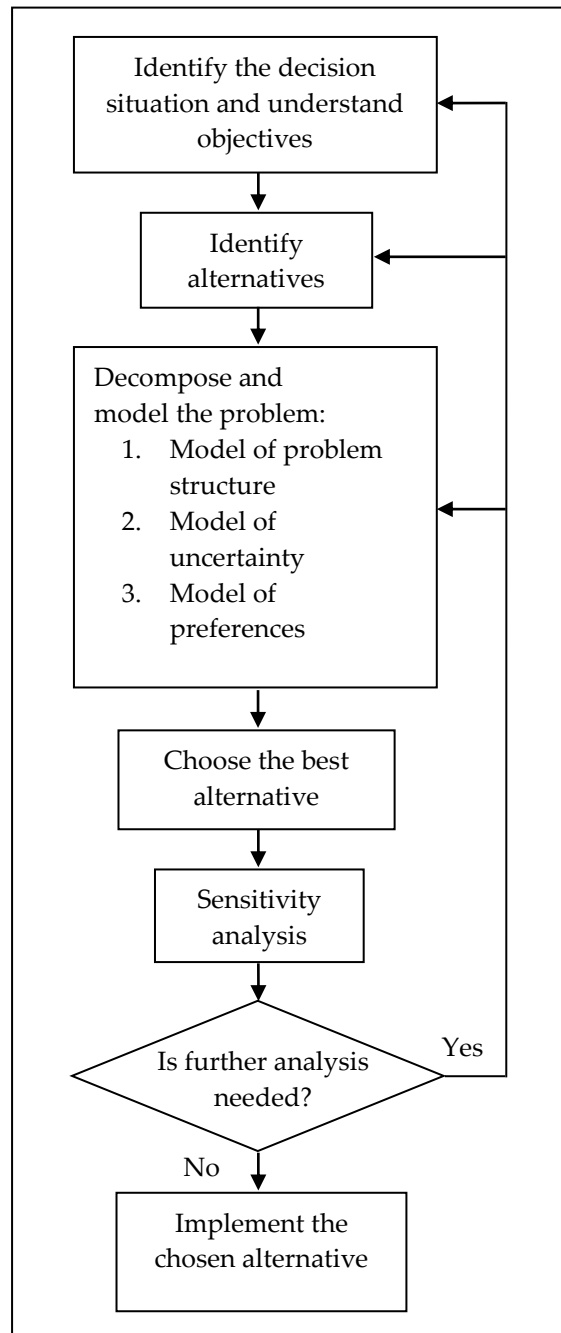


Figure 1. A decision analysis process flowchart. Source: (Clemen, 1996, p.6).

alternative given the input), and step 5) involves sensitivity analysis (shows how stable the input data is). If slight changes in one or more aspects of the model alter the recommended choice, the decision-maker may want to reconsider these sensitive aspects. This cycle may be iterated several times before a satisfactory solution is found (Clemen, 1996).

Keeney's (1992a) value-focused thinking model is an alternative to the standard model, and he argues that decision-makers should be more proactive and reverse the order of the acts that are part of the initial step depicted above. Instead, they should initially figure out their central values and objectives and thereafter look for means to accomplish these objectives. However, these two activities, understanding the decision context (the nature of the decision situation) and the identification of objectives, are closely linked regardless of the order of execution. A note should be made that Keeney's model, however logical it may sound, is still not the most common within the field.

For a prescriptive approach to decision analysis, the standard decision process model is too limiting as important elicitation aspects risk being overlooked. Therefore, there is a need for a model where elicitation aspects are made more explicit (in line with the ideas presented in Section 2.4 Prescriptive Decision Analysis). Such a model is presented in the following section, illustrated in Figure 2. This model builds on the standard model and is further developed to include essential elicitation properties and steps, which support a prescriptive approach.

In order to strive for rationality, there are a number of key components required with some adaptations to descriptive realities in order to fit real settings better. The decision process should be viewed as a process that starts with the identification of a problem and eventually results in basic decision data that *can* be used for making a decision, but serves as a guide rather than a definite commitment to action.⁸ This is more in line with prescriptive ideas, since from a prescriptive viewpoint, the alternative with the highest expected utility must not necessarily be the one chosen (Keeney, 1992b). Keeney states that one needs to interpret the implications of the model along with considerations not explicitly in the model, which in turn could result in the choice of another alternative (as this could be considered best for the real world choice) than the one with the highest expected utility. This could be the case in organizational settings, where other factors affect people's actions, such as rules, routines, traditions, hunches, imitating others, identities and roles (Anderson, 1983; March and Simon, 1993),

⁸ The traditional definition of a decision process usually encompasses a definite commitment to action.

which in turn could cause another alternative to be considered (and chosen) instead. Moreover, sometimes the prescriptive analyst might not be able to employ a complete description of the decision-maker's preference structure to find only one "best" alternative and eliminate the others (Bell et al., 1988). Decision analysis could also be used in situations involving much uncertainty as a guide to identify the data with the most effect on outcomes in order to save time and effort, e.g., if some factors of the constructed model have little effect in total, this data may not need to be so precise, whereas other data could need to be specified in more detail (as, e.g., in Danielson et al., 2010).

In the following, a prescriptive decision process model based on the standard model (cf. Clemen, 1996) with emphasis on elicitation and its role within this context is described (conceptually visualized in Figure 2), which allows for the ideal process components to be adapted to descriptive realities in order to suit the context. A note should be made that regardless of which specific arrangement of activities (which model) that is recommended in a decision process, elicitation remains an important part, and in order to retrieve good quality input data, a more controlled elicitation part is desirable. Thus, a process model is needed as well as an information model. The extraction of information (elicitation) is required for the analysis (problem, structure and solution). In order to have a meaningful elicitation, the process should be supplied with transparent information (and therefore openness is needed).

Decision process outline:

1) Problem identification and definition

What is the problem we face and can it be divided into sub-problems? Require the decision-maker/expert to make explicit what is reasonable to believe about the problem at hand. Moreover, if there is more than one decision-maker - do they have a common view of what they want and need?

2) Problem structuring - Establishing common goals and objectives

Identify and determine the goals (of the process). If more than one decision-maker is involved, a unanimous understanding has to be established. If decision-makers wish to be more proactive, this step

could come first, in line with Keeney's (1992a) ideas. According to Keeney (1992a), identifying values and objectives is preferred to precede the identification of alternatives, but sometimes alternatives already exist etc. causing the identification of alternatives to occur first. However, if there are no alternatives from the outset, identifying the values and objectives early can be a means to open up for more creative thinking in the creation of alternatives, instead of anchoring the generation process on already existing alternatives. An objective is something we wish to achieve, and Keeney (1992a) distinguishes between fundamental and means objectives, which are context-dependent. The former is the reason for interest in a decision situation, whereas the means objectives are ways to measure the achievement of the fundamental objectives. The quantification of objectives can be a way of clarifying their meaning. In an organization, strategic objectives should be because of the analysis, and it is important to see the modelling process involved in representing these perceptions as creative, dynamic and cyclic (French and Rios Insua, 2000; French, 1995b). Consequently, starting the reflection process regarding criteria early on seems reasonable, both due to the dynamic character of perceptions in prescriptive decision analysis as well as to stimulate the decision-maker to increase his/her understanding of the meaning of the criteria in the specific context.

3) Problem structuring - Identify alternatives and their outcomes

Identify all alternatives one can come up with - not all possible alternatives, thus accepting the limits of rationality (c.f. March and Simon, 1993) - and their outcomes. Consequently, aiming to find an alternative "good enough" to satisfy the explicit goals.

4) Modelling

After identifying criteria and alternatives, one can start to build a model in a computerized decision support tool. If the initial preference elicitation step part of the problem structuring is included, the modelling of criteria and preference elicitation regarding these can be carried out

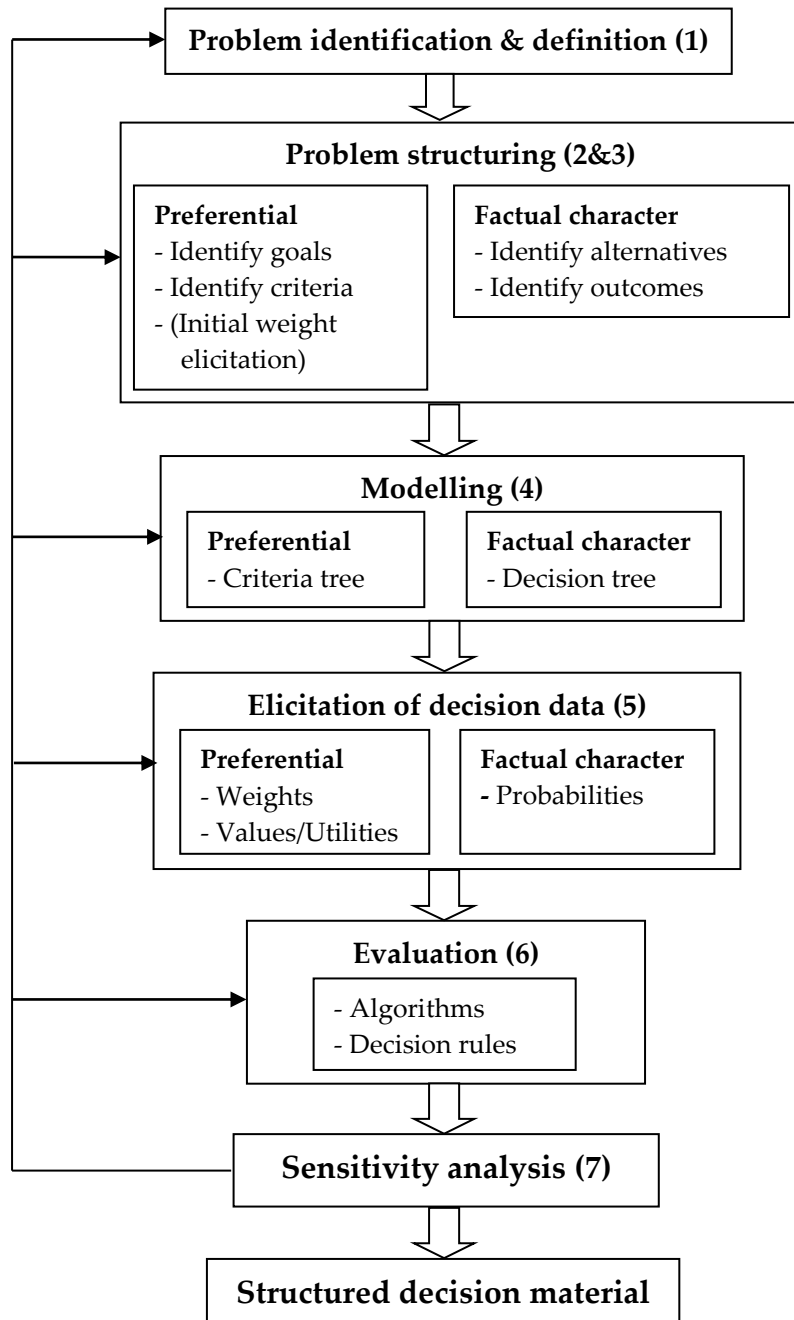


Figure 2. The decision process including prescriptive elicitation steps.

before the alternatives are included in the tool (as in Danielson et al., 2010), provided that the tool used supports this action.

5) Elicitation of decision data (both beliefs and preferences)

The elicitation of decision data is a very important step of the decision process as a whole, but has so far been relatively underrepresented within decision analysis research. After starting the modelling of the decision in a computerized tool, the elicitation of the required decision data values are needed. There is a need for procedures for developing input for the decision analytic framework, i.e. for probabilities regarding the outcomes, and decision-maker preferences. If an initial preference elicitation step regarding the criteria was performed earlier in the process, a re-elicitation of these preferences is now needed. At this stage, the decision-makers should have a better understanding of the problem at hand, the different possible alternatives, and their own beliefs. Even if an early preference elicitation step was not included, the elicitation can still be done twice by letting some time pass in between (as in Riabacke et al., 2011, 2012), where decision-makers have a chance to clarify their views. As already mentioned, this step is one of the more problematic parts of the decision making process (cf., e.g., March, 1997; Fox, 2003; Tversky and Kahneman, 1981; Corner and Corner, 1995; Keeney, 1982; Riabacke, 2007), and in order to accommodate real settings better, there is, e.g., a need to allow for more imprecision within elicitation (see more in Chapter 4. Elicitation).

6) Evaluation

When the model is sufficient for the decision faced, or requisite (Phillips, 1984), the evaluation of the decision at hand can take place. The selection of appropriate formal rules for evaluation have to be decided upon if not already selected (usually one or two of the different MCDA methods is implemented in the decision tool, and rules are implicit). Moreover, if there is more than one decision-maker involved in the decision process and individual preferences have been elicited, the form

of aggregation of preferences has to be decided. The decision-makers have to determine if each involved decision-maker's preferences should be evaluated individually first and thereafter aggregated (and how - are all equally important?) or whether all preferences should be aggregated initially into a common preference function (and if so, how?) and thereafter evaluate the decision.

7) Sensitivity analysis

The stability of the constructed model is tested.

In line with the prescriptive approach, this cycle can be iterated as many times as needed in order to reach approval, and certain steps may take more time. Several components of the outlined process have already been quite extensively studied, like the modelling of the problem in decision trees as well as the development of effective algorithms for evaluation, but within the area of *elicitation of decision data*, there is still a great need for improvements.

The decision process model as illustrated in Figure 2 explicitly deals with elicitation within decision processes. This model could be used instead of the standard model in order to illuminate the importance of elicitation within such processes as well as the need for more structured elicitation of decision data. Here, three of the components of the standard decision making process model have been further divided into elements of preferential and elements of factual character. This division in the process model is made due to differences in the character of the decision data. Probabilities can be learned from previous data regarding outcomes, or elicited from experts. Even though expert elicitation does not produce truly objective probabilities (Bell et al., 1988), the outcome of such processes is regarded as more factual in character and methods for probability elicitation can have a different character than those for preference elicitation. When it comes to criteria weights and values/utilities, such data is in the mind of the decision-maker and is individual, i.e. each decision-maker has his/her own set of preferences regarding certain outcomes and criteria, which in turn has consequences for the validation of data (is there a right or wrong?).

CHAPTER 4

4. ELICITATION

In most decision making situations where decision analysis is used, complete information about the world we seek to depict is unavailable. Any decision analysis situation relies on input of which we are unsure (French, 1995a), and some of the uncertainty relates to judgmental estimates of numerical values, like beliefs or preferences. The models used for computation require probabilistic information to represent uncertainty (in the form of probability distributions) and preferences (in the form of utility functions). In decisions involving multiple objectives, there is also the need to make value trade-offs to indicate the relative desirability of achievement levels on one objective in comparison to the others (represented by criteria weights in MAVT/MAUT methods).

While there has been an increase in research (and an intense debate) on elicitation over the last few decades within several disciplines, such as, e.g., psychology, statistics, and decision and management science, there are still no generally accepted methods and the process of eliciting adequate quantitative information from people is still one of the major challenges facing research within the field of decision analysis (Fox, 2003). Although different research areas have different explanations for elicitation problems, they seem to agree on the fact that in applied contexts we should be concerned not only with *what* we ask experts to assess, but also *how* we ask it (Kynn, 2008). Statistical research on elicitation has been greatly influenced by psychological findings on how people represent uncertain information cognitively, and how they respond to queries regarding that information.

The suggested methods for elicitation have distinct features which all can impact their applicability in practice and need to be addressed more explicitly. Also, both procedural and evaluative elicitation aspects are often discussed

interchangeably. In order to study and analyse elicitation more explicitly, there is a need to categorize methods and the following division of elicitation into three conceptual components is therefore suggested:

1) Extraction

This component deals with how information (probabilities, utilities, weights) is derived through user input.

2) Representation

How to capture the retrieved information in a formal structure, i.e. the internal format used to represent user input.

3) Interpretation

Is dependent on the expressive power of the representation used and how to assign meaning to the captured information in the evaluation of the decision model used.

An analytic division of this kind is not found in literature within the field and is a useful way to analyse elicitation methods in order to recognize their characteristics and identify elements that can impact their applicability in practice.

4.1 PROBABILITY AND UTILITY ELICITATION

In the classic decision analytic framework (cf., e.g., von Winterfeldt and Edwards, 1986), numerical probabilities are assigned to the different events in decision tree⁹ representations of decision problems. The best alternative is the one with the optimal combination of probabilities and utilities corresponding to the possible outcomes associated with each of the possible alternatives. After the process of identifying what aspects of a problem (what parameters) to elicit, which subjects (information sources) to use and possible training for the subject(s), the most crucial part is then to elicit the necessary values from people. Probability information is most commonly elicited from domain experts, and the expert has to express his or her knowledge and beliefs in probabilistic form during the step of extractions. This task often involves a facilitator to assist the expert, as most people are unaccustomed to expressing knowledge in this fashion. Garthwaite et

⁹ A decision tree is a way of structuring the sequences of possible external events and actions consequent to each external event (von Winterfeldt and Edwards, 1986).

al. (2005) conclude that in order for an elicitation to be successful, the values may not be “true” in an objectivist sense (and cannot be judged that way), but are an accurate representation of the expert’s knowledge (regardless of the quality of that knowledge). Moreover, they believe that a reasonable goal for elicitation is often to describe the “big message” in the expert’s opinion. The subjectivist outlook on the information required in decision analysis is shared by others, see, e.g., Keeney (2004) who states that the foundation for decision making must be based on subjective information, although part of the decision analysis discipline still refers to an objective analysis. For a wider discussion concerning objective (classical) and subjective (personal) probabilities, cf., e.g., (de Finetti, 1968; Savage, 1954; Wright and Ayton, 1994). Subjective probability is thus one of the prime numerical inputs, but the meaning of probability depends on the conceptual distinction between single-event probabilities and frequencies. This perception can differ among experts, even among those making assessments regarding the same quantities (cf., e.g., Riabacke, 2010). The elicitation of probabilities has been quite extensively studied, and recommendations as to how to make such assessments and corresponding problems can be reviewed further in, e.g., (Clemen, 1996; Corner and Corner, 1995; Hogarth, 1975; Morgan and Henrion, 1990; Wallsten and Budescu, 1983).

Methods for utility elicitation have many similarities to probability elicitation processes, but are more complex. Probabilities can be elicited from experts (and *should* remain the same regardless of who makes the assessment), but can also be learned from data, whereas utility functions are to accurately represent decision-makers’ *individual* risk attitudes, and are, thus, required for each user. Utility can be described as the value a decision-maker relates to a certain outcome, and in utility elicitation, different methods are used to give the (abstract) concept of preference an empirical interpretation. The elicitation process itself, regardless of the method employed, has proved to be cognitively demanding for people and error prone. Several techniques for utility elicitation have been proposed and used, and in Johnson and Huber (1977) a categorization of these techniques is provided. The category labelled gamble methods contain the most commonly used techniques, where several variations on question design are being used. A broad categorization of standard-gamble methods is given in (Farquhar, 1984), but framing the utility assessment in terms of hypothetical gambles and lotteries may not map people’s behaviour in real situations. Some people have a general aversion towards gambling, and people often overweight certain outcomes in comparison to those that are merely probable (Kahneman and Tversky, 1979), which complicates matters further.

Moreover, the classical theory of preference assumes that normatively equivalent procedures for elicitation should give rise to the same preference order, which is an assumption often violated in empirical studies, cf., e.g. (Tversky and Kahneman, 1981; Lenert and Treadwell, 1999; Pählman and Riabacke, 2005). Lichtenstein and Slovic (2006) state that people do have well-articulated preferences regarding certain matters, but in some settings construct their preferences during the process of elicitation, which is the cause for these violations. They suggest that the need for preference construction often occurs in situations where some of the decision elements are unfamiliar and where there are some types of conflicts among our preferences regarding the choices presented to us. Such circumstances make us more susceptible to be influenced by certain factors during the elicitation process, such as, e.g., framing (see more in Section 4.3 Issues on Elicitation) and could thus explain many of the problems related to elicitation.

4.2 WEIGHT ELICITATION

In multi-criteria decision analysis (MCDA), the relative importance of the different criteria is a central concept. In an additive MAVT/MAUT model, the weights reflect the importance of one dimension relative to others. The weight assigned to a criterion is basically a scaling factor, which associates scores on that criterion to scores on all other criteria. Methods for eliciting criteria weights are compensatory, i.e. the extracted information on the weights' relative importance as assigned by decision-makers implicitly determine trade-offs between the number of units on one criterion they are willing to waive in order to increase the performance on another criterion by one unit. As already mentioned, there are various sources of uncertainty within the application of MCDA methods, and regarding criteria their definition as well as the elicitation of criteria performance values (weights) involve uncertainty.

There are several techniques for deriving weights from preference statements. However, like probability and utility elicitation, the elicitation of weights is a cognitively demanding task (Larichev, 1992; Barron and Barrett, 1996a; Belton and Stewart, 2002), is subject to different biases (cf., e.g., Borcherding et al., 1991), and the elicited values can be heavily dependent on the method of assessment (cf., e.g., Pöyhönen and Hämäläinen, 2001). In the literature, there have been a number of methods suggested for assessing criteria weights, but the suggested methods have distinct features which all can impact their applicability in practice. Weight elicitation methods differ regarding the type of information

they preserve from the decision-maker's judgments in the extraction component to the interpretation component. In practice, the actual usefulness of elicitation methods is determined by procedural aspects (Mustajoki et al. 2005), and therefore elicitation methods with relatively simple extraction components are most common in applied settings. There are several weighting methods that appear to be minor variants of one another, but even small procedural differences have shown to have important effects for inference and decision making (Bottomley and Doyle, 2001). In the following sections, some of the most prominent weight elicitation methods are described (an overview is also given in Table 2, and a summary of their extraction parts is accounted for in Table 3).

4.2.1 Ratio Weight Procedures

Ratio weight procedures maintain ratio scale properties of the decision-maker's judgments from extraction and use exact values for representation and interpretation. Common to all these methods is that the actual attribute weights used for the representation are derived by normalising the sum of the given points (from the extraction phase) to one. Methods adopting this approach range from quite simple rating procedures, like the frequently used direct rating (DR) and point allocation (PA) methods (for a comparison of the two methods, cf., e.g., Bottomley et al., 2000), to somewhat more advanced procedures, such as the often used SMART (Edwards, 1977), SWING (von Winterfeldt and Edwards, 1986) or trade-off (Keeney and Raiffa, 1976) methods.

As already mentioned, these methods all differ in the procedure during the step of extraction. In the DR method, the user is asked to rate each attribute on a scale of 0-100, whereas the user in PA is asked to distribute a total of 100 points among the attributes. Bottomley et al. (2000) concludes that weights derived from DR are more reliable, and perhaps the extra cognitive step of having to keep track of the number of points to distribute in the PA method influences the test-retest reliability.

In SMART, the user is asked to identify the least important criterion, which receives 10 points, and thereafter the user is asked to rate the remaining criteria relative to the least important one by distributing points. Since no upper limit is specified, the rating extracted from the same person can differ quite a bit in the interpretation if the method is applied twice (Riabacke et al., 2012). Consequently, this aspect of the extraction stage of SMART seems like an element that can affect the internal consistency in the interpretational step of the method.

In the SWING method, the decision-maker is asked to consider all criteria at their worst consequence levels, and to identify which criterion whose conse-

quence he or she would prefer most to change from its worst to its best level (swing). This criterion will be given the highest number of points, 100. This procedure is repeated on the remaining set of criteria. First, with the criterion next to the most important swing, where this criterion will be given a value relative to the most important one (thus their points denote their relative importance), and so on. Common to all the methods described so far is also that the number of judgments required by the user during extraction is a minimum of N number of judgments, where N is the number of attributes.

In the trade-off method, the criteria are considered in pairs where two hypothetical alternatives are presented to the decision-maker during extraction. These alternatives differ only in the two criteria under consideration and in the first alternative the performance of the two criteria are set to their worst and best levels respectively and in the second alternative the opposite applies. The decision-maker is asked to choose one of the alternatives, and thereby indicates the more important one. Thereafter (s)he is asked to state how much (s)he would be willing to give up on the most important criterion in order to change the other to its best level, i.e. state the trade-off (s)he is willing to do for certain changes in values between the criteria. The minimum number of judgments is $N-1$, but a consistency check requires considering all possible combinations of criteria, which would result in $N \cdot (N-1)$ comparisons. Consequently, the extraction component of the trade-off method is operationally quite complex and cognitively demanding in practice due to the large number of pairwise comparisons needed when the criteria are more than a few. Moreover, there is a tendency to give greater weight to the most important attribute in comparison to methods like DR and SWING (see, e.g., Fischer, 1995).

Most commonly, the degree of importance of an attribute depends on its spread (the range of the scale of the attribute), and this is why methods like SMART, which do not consider the spread specifically, have been criticized. The SMART and SWING methods were therefore later combined in the SMARTS method (Edwards and Barron, 1994) to explicitly include spread as well during elicitation. Yet, with methods where ranges are explicitly considered during the elicitation of weights, several empirical studies imply that people still do not adjust weight judgments properly when there are changes in the ranges of the attributes (cf., e.g., von Nitzsch and Weber, 1993). In all studies reported in the literature, the range sensitivity principle (measured by the Range Sensitivity Index, RSI, as suggested by von Nitzsch and Weber, 1993) is violated, often significantly (Beattie and Barron, 1991; von Nitzsch and Weber, 1993; Fischer, 1995; Yeung and Soman, 2005). Von Nitzsch and Weber (1993) suggest that during decision-makers' judgment on importance, an intuitive idea of an attribute's

importance (past experience) functions as an anchor that is thereafter adjusted by the range of the attribute in the current choice context. Fischer (1995) hypothesized that methods, which more explicitly focus on what is gained or lost in terms of different objectives, result in assessed values that are more sensitive to the ranges of the consequences. As an alternative explanation to violations of the range sensitivity principle, Monat (2009) claims that the use of local scales may be the problem. Instead, global scales that reflect the best and worst values from the decision-maker's view (not the best and worst from the option set) should be remapped to the best and worst values on the scale (*ibid.*). However, in such a model the problem is instead the difficulty in identifying the extreme values on the global scale. So far, no method has managed to adequately respect the range sensitivity principle in empirical studies.

4.2.2 Imprecise Weight Elicitation

Accurate determinations of attribute weights by using ratio weight procedures are tricky to acquire in practice as assessed weights are always subject to response error (Jia et al., 1998), and some suggest that the attempt of finding precise weights may be an illusion (Barron and Barrett, 1996a). Consequently, suggestions on how to use imprecise weights instead have been proposed. In MCDA, there are different approaches for handling more imprecise preference, mainly outlined as one or more of the following (Belton and Stewart, 2002): (1) Ordinal statements, (2) Classifying outcomes into semantic categories, and (3) Interval assessments of magnitudes using lower and upper bounds.

Rank-order methods belong to the first approach. During the extraction, decision-makers simply rank the different criteria, which are represented by ordinal values. Thereafter, these ordinal values are translated into surrogate (cardinal) weights that are consistent with the supplied rankings in the interpretational step. The conversion from ordinal to cardinal weights is needed in order to employ the principle of maximizing the expected value or any other numerical decision rule in the evaluation. Thus, in these methods ratios among weights are determined by the conversion of ranks into ratios in the interpretational step. Several proposals on how to convert such rankings to numerical weights exist, e.g., rank sum (RS) weights, rank reciprocal (RR) weights (Stillwell et al., 1981), and centroid (ROC) weights (Barron, 1992). Of the conversion methods suggested, ROC has gained most recognition. Edwards and Barron (1994) propose the SMARTER (SMART Exploiting Ranks) method to elicit the ordinal information on importance before being converted to numbers using the ROC method.

However, there is often some weak form of cardinality, e.g., people can be quite confident that some differences in importance are greater than others (Jia et al., 1998), which is totally ignored in rank-order approaches. Thus, although mere ranking alleviates some of the cognitive demands on users, the conversion from ordinal to cardinal weights may produce differences in weights that do not closely reflect what the decision-maker actually means by his/her ranking. A contribution in this thesis is to propose a structured framework for mixing ordinal and cardinal information. In Riabacke et al. (2012), a suggestion on how to complement the supplied ranking with preference relation information without demanding exactness from the decision-maker is given. In the Cardinal Rank Ordering of Criteria (CROC) method (ibid.), the user supplies both ordinal as well as imprecise cardinal relation information on criteria during extraction, which is translated into regions of significance in the interpretational step.

Methods utilizing semantic scales (e.g. very much more important, much more important, moderately more important etc.) for stating importance weights and/or values of alternatives during extraction belong to the second category, like in the AHP method (Saaty, 1980). However, the correctness of the conversion in the interpretational step, from the semantic scale to the numeric scale used by Saaty as a measure for preference strength has been questioned, e.g., by Belton and Stewart (2002). Moreover, the use of verbal terms in general during elicitation have been criticised, since words can have very different meanings for different people and people often assign different numerical probabilities to the same verbal expressions (Merkhofer, 1987; Kirkwood, 1997). Thus, such numerical interpretations of verbally extracted information from people are less common among the imprecise preference methods (except for the AHP method).

In some applications for decision analysis, preferential uncertainties and incomplete information are handled by using intervals (cf., e.g., Walley, 1991; Danielson et al., 2008), where a range of possible values is represented by an interval. Such methods belong to the third approach, and are claimed to put less demands on the decision-maker as well as being suitable for group decision making as individual differences in preferences and judgments can be represented by value intervals (Jiménez et al., 2006). When using interval estimates during extraction, the minimum number of judgments is $2 \cdot (N-1)$, since both the upper and lower bounds are needed for the preference relations.

In the GMAA system (Jiménez et al., 2003, 2006), there are two procedures for assessing weights. Either the extraction is based on trade-offs among the attributes and here the decision-maker is asked to give an interval such that (s)he is indifferent with respect to a lottery and a sure consequence. The authors state that this method is most suitable for low-level criteria, whereas the other

Weight Elicitation Method	EXTRACTION			REPRESENTATION	INTERPRETATION
	Assessment	Input	Min. no. of judgments		
Direct Rating	Cardinal Joint proc.	Precise	N	Point estimates	Normalized Criteria Weights
Point Allocation	Cardinal Joint proc.	Precise	N	Point estimates	Normalized Criteria Weights
SMART	Cardinal Joint Proc.	Precise	N	Point estimates	Normalized Criteria Weights
SWING	Cardinal Joint Proc.	Precise	N	Point estimates	Normalized Criteria Weights
Trade-off	Cardinal Pairwise procedure	Precise	$N \cdot (N-1)$ (with consistency check)	Point estimates Rel. between pairs of criteria	(Combined) Normalized Criteria Weights
Rank-order methods	Ordinal Joint procedure	Rank-order	N	Comparative statements	Surrogate Criteria Weights (translated using a conversion method.
AHP	Cardinal Pairwise procedure	Semantic	$N \cdot (N-1)$ (with consistency check)	Semantic estimates Rel. between pairs of criteria	(Combined) Surrogate Criteria Weights (translated from semantic to exact numerical)
CROC	Ordinal & Cardinal Joint proc.	Rank-order and Imprecise cardinal relation information	N (>N with cardinal input)	Comparative statements + Imprecise cardinal relation information	Surrogate (centroid) Criteria Weights
Interval methods	Normally, a generalized ratio-weight procedure	Interval endpoints (precise)	2 · (min. no. of judgments in employed ratio-weight procedure)	Intervals	Interval Criteria Weights

Table 2. An overview of some of the most prominent weight elicitation methods (N = No. of criteria).¹⁰

¹⁰ To explicitly include spread, the SWING of criteria can be applied to methods, where the initial procedural design does not include criteria range, as in e.g., SMARTS (SMART including SWING).

Weight Elicitation Method	Extraction (assessment procedure)
Direct Rating	Rate each criterion on 0-100 scale.
Point Allocation	Distribute 100 points among criteria.
SMART	1) Identify least important criterion, assign 10 points. 2) Rate remaining criteria relative to least important.
SWING	1) Consider all criteria at their worst consequence level. 2) Identify the criterion most important to change from worst to best level, assign 100 points. 3) Continue steps 1 and 2 with remaining criteria, rate relative to most important.
Trade-off methods	Judge criteria in pairs. 1) Make choice between two alternatives, - alt.1: best consequence level of first criterion and worst of the second, and - alt.2: worst consequence level of first criterion and best of the second. 2) State how much he/she is willing to give up on most important criterion to change the other to its best level. 3) Continue steps 1 and 2 with remaining criteria.
Rank-order methods	Ordinal statements of criteria importance, i.e. rank all criteria from most important to least important.
AHP	Uses a systematic pairwise comparison approach in determining preferences. 1) Make choice between two criteria to determine which is most important. 2) State how much more important the criterion identified in step 1) is in comparison to the second criterion using a semantic scale to express strength of preference. 3) Continue steps 1 and 2 with remaining criteria.
CROC	1) Rank all criteria from most to least important. 2) Most important criterion is given 100 points, user is asked to express the importance of the least important criterion in relation to the most important. 3) Adjust distances between criteria on an analogue visual scale to express cardinal importance information between criteria.
Interval methods	Most commonly, a generalized ratio weight procedure, which employs interval judgments to represent imprecision during extraction instead of point estimates, as in e.g., interval SMART/SWING.

Table 3. A summary of the assessment procedure during the step of extraction of some of the most prominent weight elicitation methods.

extraction approach, direct assignment, is more suitable for the upper level criteria that could be more political. Here, the decision-maker directly assigns a weight interval to the respective criteria. In the interpretational step, the extracted interval values are automatically computed into an average normalized weight (precise) and a normalized weight interval for each attribute. In Mustajoki et al. (2005), the authors propose an Interval SMART/SWING method, in which they generalize the SMART and SWING methods (for point estimates) into a method that allows interval judgements to represent imprecision during extraction. Here, the reference attribute is given a fixed number of points, whereas the decision-maker is allowed to reply with interval assessments to the ratio questions during extraction (to describe possible imprecision in his/her judgments). The extracted weight information is represented by constraints for the attributes' weight ratios, which in addition to the weight normalization constraint determine the feasible region of the weights in the interpretational step.

4.3 ISSUES ON ELICITATION

There is a widely discussed contradiction between the ambiguity of human judgment and the exactness (of elicited values) required by most decision analysis models. People have problems judging exact values (see, e.g., Shapira, 1995), which poses a problem when the required values are point estimates, and some of the deviations from the traditional decision theoretical expectations could probably be attributed to this inability. Using a single number to represent an uncertain quantity can also confuse a person's judgment about uncertainties with the desirability of various outcomes (Kirkwood, 1997). Also, subjects often do not initially reveal consistent preference behaviour in many decision situations (Keeney and Raiffa, 1976; Keeney, 1982; Wehrung et al., 1980), or protect themselves from exposure by obscuring and managing their preferences (March, 1997). Brunsson (1989) argues that organizations continuously work with a two faced perspective and logical approach, where the logical rationality of a decision has to be legitimized, which in turn results in ambiguous preferences. Moreover, in elicitation methods where a risky alternative is compared to a certain outcome, people often overweight the certain outcome - the so called certainty effect (Kahneman and Tversky, 1979). In addition, the conditions for procedure invariance are generally not true, people do not have well-defined values and beliefs in many decision situations where decision analysis is used, and choice is instead contingent or context sensitive (Tversky et al., 1988). People are, further-

more, poor intuitive decision-makers in the sense that our judgments are clearly affected by the frame in which information is presented as well as the context. Decision-makers appear to use only the information that is explicitly presented in the formulation of a problem (Slovic, 1972; Fischhoff et al., 1978), and implicit information that has to be deduced from the display seems to be ignored. The framing (formulation) of the problem strongly affects human reasoning and preferences, even though the objective information remains unchanged (Tversky and Kahneman, 1981, 1986).

The heuristics and biases programme initiated by Tversky and Kahneman (1974) illustrates many of the systematic deviations from traditional theoretical expectations inherent in our ways of reasoning, making judgments and in our memory, which cause problems for elicitation processes. We have, e.g., a tendency to be overconfident in our own judgments, overestimate desirable outcomes and seek confirmation of our preconceptions. Tversky and Kahneman (1974) argued that the processes of human judgment were totally different than what rational models required, and identified a set of general-purpose heuristics that underlie judgment under uncertainty. These heuristics (originally three – availability, representativeness, and anchoring and adjustment) were shown to result in systematic errors (biases), such as the conjunction fallacy and base rate neglect. Over the years, many more such heuristics and biases have been identified. These can be both motivational (due to overconfidence) and cognitive (due to human thought processes). Studies where methods for elicitation have been compared in practice are often inconsistent (cf., e.g., Wang et al., 2002, regarding probabilities, Lenert et al., 2001, concerning preferences, and Pöyhönen and Hämäläinen, 2001, regarding weights), and there is no general agreement on the underlying cognitive processes involved in these assessments. Behavioural concerns are highly relevant to prescriptive decision aiding, especially in identifying where the improvable deficiencies in current practices are as well as in fitting the design of decision aids to the reality of human abilities (Brown and Vari, 1992).

An additional problem in measuring method preciseness for preference elicitation methods occur due to the subjective nature of the elicited values. Even though most people now agree on the fact that assessed probabilities are subjective in nature, they are to represent facts and if experts' values disagree, different methods can be used for combining multiple assessments in order to improve the quality of the final probabilities (in belief that a set of experts can provide more information than one). When combining assessments, the main approaches for doing this are by mathematical aggregations of individual assessments or by obtaining group consensus (Clemen and Winkler, 1999). When it comes to preference elicitation, it is more difficult to determine that the elicited values

correctly represent the preferences held by the decision-maker. Thus, there is a bigger problem with validation in this realm (see more in Chapter 5. Methodology).

4.4 APPROACHING ELICITATION PRESCRIPTIVELY

There is a great deal of uncertainty involved in elicitation and the many reports of the difficulties with eliciting precise values (probabilities, utilities and weights) from people, accounted for in the previous section, suggest that current procedures need to be better adapted to real settings in order to be more practically useful. The human brain is not inherently numeric, and we are not introspective by nature.

Elicitation is an iterative process, where the elicited values may have to be adjusted, due to deviations from theoretical expectations or an increased understanding of the problem and the context by the expert/decision-maker. Coherence in elicited values has to do with how well the values fit together, and models for coherence are mainly focused on probability theory, and compensating for the fact that it often falls short as a model of subjective probability (Kynn, 2008). For example, Tversky and Kahneman have raised the question of whether probability theory should really be thought of as the calculus of human uncertainty in the first place, and Fox (1994, p. 80) states that “mathematical probability has been developed as a *tool for people to use*; a body of concepts and techniques which helps to analyse uncertainty and make predictions in the face of it”, but that a more liberal attitude would allow for a better understanding of human judgment under uncertainty and the development of more sophisticated technologies for aiding such judgment. The prescriptive analysts must learn how to elicit judgments from decision-makers and make sense out of them (Bell et al., 1988).

As already mentioned, prescriptive decision analysis (see more in Section 2.4 Prescriptive Decision Analysis) is an attempt to narrow the gap between research within the normative and descriptive disciplines. It is a more practically useful approach for handling decision problems, still employing a formal model for analysis but better adapted to real decision problems. If we are to adopt the prescriptive approach there is a need for processes, methods, and tools that can handle the inherent uncertainty of the decision-maker more explicitly, and are coherent. In essence, prescription involves making it all more realistically useful, and for tools to provide the good support they are intended to, there is a need to use them realistically. We can support decision-makers in their decision making,

but we cannot change them. Brown and Vari (1992), among others, assert that the behavioural (descriptive) realities are very important in order to design more prescriptive decision aids.

In the literature on elicitation of the inputs required for decision analysis (probabilities, utilities, weights), there is no consensus regarding:

- the exact nature of the identified gap between ideal and real behaviour,
- how to avoid the observed phenomena, and
- how to evaluate whether a method has produced accurate input data.

However, reaching consensus on all aspects within the decision analysis community seems difficult, and perhaps we should instead, as e.g. Pöyhönen (1998) suggests, focus research on how methods are used in practice instead of searching for a superior theoretical base for methods. By dividing elicitation into the three components *Extraction*, *Representation* and *Interpretation* (see more at the beginning of this chapter), the applicability of methods in practice can be addressed more explicitly. In practice, we should strive for finding methods that are less cognitively demanding and less sensitive to noisy input within each component.

The extraction component appears like the most error-prone of the three suggested as it concerns the procedural design of the method, which is more or less cognitively demanding during user interaction. Especially behavioural research has concentrated on the extraction aspect of elicitation, most commonly how different biases occur when people interact with elicitation methods. Within this realm, the interpretational component is mostly discussed during validation as a means for measurement (e.g. illustrating procedure invariance).

In order to reduce the gap between theoretical research and practical needs, there are several aspects of the extraction component that need to be considered. Behavioural aspects, like the heuristics and corresponding biases people use during extraction (cf., e.g., Kahneman et al., 1982), are important to be observant of in order to reduce such effects. In Pählman and Riabacke (2005) and Riabacke et al. (2006a), increased awareness of how presentation formats affect people's choices is suggested to reduce the framing problems (cf., e.g., Tversky and Kahneman, 1981) so often discussed as a hindrance to sufficient elicitation, e.g., such as being aware of people's aversion to losses, tendency to overweight certain outcomes etc. Moreover, relaxation of the precise statements that are commonly required in the extraction component of elicitation methods seems like an advantageous approach to adopt. As already mentioned, there are several ways this can be achieved.

Examples of approaches for eliciting the required information in a less precise fashion are methods based on visual aids or verbal expressions. For example,

the probability wheel (Spetzler and Stael von Holstein, 1975) is a popular visual method for eliciting probabilities (the user indicates his/her belief in probability size on a circle by sizing a pie wedge to match the assessment on that probability). Such methods often use a combinatorial extraction approach, where the user can modify the input both visually and numerically. The representation of visually extracted input is most commonly an exact number, which is then also used in the interpretation. The use of verbal terms during extraction is supposedly more in line with the generally imprecise semantics of people's expressions of preferences and beliefs, but have as already mentioned been criticised for their vagueness, which can cause problems in the interpretational step where the verbal expressions are represented by numbers. Words can have different meanings for different people and people often assign different numerical probabilities to the same verbal expressions (Merkhofer, 1987; Kirkwood, 1997).

Another way of handling preferential uncertainties and incomplete information in a less precise way is by using intervals (cf., e.g., Danielson et al., 2009; Jiménez et al., 2006; Park, 2004), where a range of possible values is represented by an interval. Declared benefits with the interval approach are, e.g., that such representations are more realistic interpretations of people's knowledge, beliefs and preferences, since these elements are not stored with preciseness in our minds. One can also conduct a first analysis of the problem given imprecise statements and test whether the input is sufficient for the final evaluation of alternatives. If not, one can identify the input that needs to be further specified. Other advantages are that methods based on more approximate preferences can lead to a more interactive decision support process as the evolution of the decision-maker's priorities can be computed throughout the process, which in turn could lead to improved decision quality (Salo, 1995). In addition, such methods are especially suitable for group decision making processes as individual preferences can be represented by a union of the group's judgments (ibid.). In the latter case, group members can seek consensus by trying to reduce the width of the intervals and compromise on their individual judgments if needed.

There is relatively little research done regarding people's perception of interval representations of uncertainty. In elicitation methods using intervals, the user still normally has to enter exact numbers to specify the end points of the intervals in the extraction step of the methods (e.g., giving direct statements or answering trade-off questions). In Riabacke et al. (2006b), the use of interval representations of uncertainty were compared to the use of point estimate representations and the study showed that such expressions can be preferred during extraction. However, although methods using the interval representation allow people to be more imprecise in the extraction, this step can still be quite

cognitively demanding, e.g., when the decision-maker needs to express weight interval preferences on each consequence range for the criteria in question. Yet, the representation appears more realistic and less sensitive to noisy input in the extraction (as the interval is more likely to contain the “true” input value) than using an exact numeric representation of the required input values. The interval representation is thereafter used in the interpretational step. Approaches using intervals in the interpretational step do not always require interval input during extraction. Instead, comparative statements on the values of assessment can be made, which are later translated into intervals in the interpretational step as in Danielson and Ekenberg (2007).

For the elicitation of weights, ranking methods using surrogate weights (e.g., ROC weights, Barron, 1992; Barron and Barrett, 1996b) in the interpretational step have been alleged to be less cognitively demanding and advantageous for group consensus (as groups are more likely to agree on ranks than precise weights, Barron and Barrett, 1996a). The input retrieved from the extraction step of elicitation methods adopting this approach is a ranking order of the criteria in question, and thus, the representation is merely ordinal information. The interpretation is the surrogate weights (exact numbers) resulting from the conversion method used.

Looking at weight elicitation, methods differ regarding the type of information they preserve from the decision-maker’s judgments in the extraction component to the interpretation component. The two extremes within weight elicitation approaches are to use either exact values or mere ranking during extraction. In Riabacke et al. (2012), we have acknowledged the problems with using exact values by not forcing the decision-maker to express unrealistic precision, yet not ignoring weaker forms of cardinality which often exist. An initial attempt of such a method was suggested in Riabacke et al. (2009), and the CROC weight elicitation method described in Riabacke et al. (2012) is a further development of this approach. In the CROC method, the user supplies both ordinal as well as imprecise cardinal relation information during extraction by providing a ranking of criteria complemented by imprecise preference relation information (using a graphical method). This information is translated into regions of significance in the interpretational step and the resulting weight distribution can be found by a decision analysis tool such as DecideIT (Danielson et al., 2003). In its representational and interpretational aspects, CROC extends the ROC weight method (Barron, 1992; Barron and Barrett, 1996b) into handling imprecise and cardinal information, and aims to reduce effects of “noisy input” not only in the extraction step, but also in the interpretation. Allowing for more imprecise preference

statements could also be a way to lessen decision-makers' reluctance to reveal their true preferences.

Important for the practical applicability of MCDA methods is the easiness of the method (see, e.g., Stewart, 1992), and simpler tools are easier to use and therefore more likely to be useful. The CROC method for eliciting weights was implemented and tested in practice as part of an MCDA process model used to aid a decision making process (Danielson et al., 2010), where it was shown useful. In this real-life case, elicitation was more explicitly emphasized throughout the decision process in line with the model suggested in Chapter 3 (Figure 2). Moreover, elicitation methods that are more direct are easier and less likely to produce elicitation errors (Edwards and Barron, 1994). In Riabacke et al. (2012), the CROC method was compared to the Point Allocation method (which uses exact values) and shown to be more stable. Some even claim that simpler, fast and frugal methods can produce results that are almost or as good as results attributed to those obtained by more extensive analysis, cf., e.g. (Katsikopoulos and Fasolo, 2006). Larichev et al. (1995), among others, suggest that exactness of results must not be the main aim with decision analysis and that different situations call for different levels of exactness depending on decision-makers' abilities to provide exact judgments. Consequently, simpler approaches, less reliant on great precision from people, are allegedly a better way of accommodating decision-makers in real settings.

4.5 HANDLING ELICITATION IN SOFTWARE

As already mentioned in previous sections, elicitation is highly important to prescriptive endeavours and to decision making processes as a whole. From behavioural research, the need for decision support systems (based on decision analysis) that are easier for people to understand and use has been highlighted, although their application will still require some form of training prior to usage and/or facilitator to assist during the decision making process. Many of the current decision analysis tool deficiencies would be alleviated if elicitation of inputs were effective (Brown, 2006). Decision analysis applications should also agree with and support the different steps of the decision making process (see more in Chapter 3. Decision Process Outline) in a more complete fashion (French and Xu, 2005), and offer more flexibility by being adaptable to different user needs (Riabacke et al., 2006b).

Adopting a prescriptive approach in order to make decision making processes, supported by decision analysis applications, more practically useful is

now more or less an agreement within the decision analysis community. However, there is still a great need to focus on the elicitation of information to use for analysis as the axioms of decision analysis advise us how to analyse decision problems, but do not indicate where the information used for analysis comes from. Within elicitation, a central question is how to elicit knowledge in probabilistic form and to represent people's individual preferences adequately.

In the following, suggestions on measures to reduce some of the problems with elicitation identified in the previous sections and things to consider when designing decision analysis tools are described. The intended context is for application in a decision analysis tool used to support a decision making process. In practice, we should aim for methods that are less cognitively demanding and less sensitive to noisy input.

The extraction component of elicitation methods appears like the most error-prone of the three elicitation components (see the beginning of this chapter) as it concerns the procedural design of the elicitation method. This can be more or less cognitively demanding during the step of extraction, and therefore, there is a need to incorporate support for users during the extraction step of the elicitation method in decision analysis applications. This procedure should preferably be interactive in order to reduce incoherence in values by assisting the expert/decision-maker during extraction in a cyclic fashion until coherence is achieved and satisfaction is reached.

Moreover, the extraction should allow for more imprecision in the decision input data required from users, especially such input that is subjective in nature and/ or is not naturally processed numerically by people during assessments, like beliefs and preferences. Also, sometimes users are unable to provide all the needed input with the exactness required by many current applications. In any case, the preciseness in such fixed assessments is not certain enough, and presents a false impression of accuracy (von Winterfeldt and Edwards, 1986, calls the precision of numbers an illusion). Allowing for more imprecision during extraction could be achieved by using imprecise procedures, where user statements are quantified to imprecise statements in the interpretational step of the approach. In Riabacke et al. (2011, 2012), a formalization of such an imprecise method was realized in two types of software applications, both as a weight elicitation method part of a decision tool as well as a web based weight elicitation method part of a participatory approach to increase democratic decision making. Imprecise extraction can also be achieved by having the user state an imprecise value, e.g., an interval as in software such as PRIME Decisions (Gustafsson et al., 2001), GMAA (Jiménez et al., 2006) and DecideIT (Danielson et al., 2003). In applications supporting an interval approach, the user normally

enters an interval in the interface of the extraction part of the elicitation method, which is later used in the interpretation.

The elicitation method employed could also be more flexible and allow for different types of statements to accommodate the user's preference of method as advocated in Riabacke et al. (2006b). Different methods may suit different types of problems better, e.g., the size of the problem and the time available for analysis could rule out certain methods at the onset as they would be unfitting. However, a note should be made that this could also constitute a problem if more than one individual makes the assessments and they use different methods (due to the frequency of procedure invariance violations noted in descriptive studies, i.e. different methods often yield different results).

When designing elicitation methods, there is a need to understand psychological traps within elicitation, such as framing and heuristics that produce biased assessments in order to apply measures to lessen their effect in the method design. Using a clear terminology is important, such as e.g., explaining the meaning of certain terms in the specific context, thoroughly considering the phrasing of questions, being explicit on whether the required probabilities are single-event probabilities or frequencies (and explain the difference to people unaware of the difference) etc.

Moreover, if MAVT/MAUT is used (as opposed to single criterion decision analysis), the attributes should be well-defined and measures to increase understanding of these and stimulating reflections about personal preferences could be to include two weight elicitation occasions as in the process model suggested in Chapter 3 and employed in Riabacke et al. (2011, 2012). In addition, using a decision process in line with the one described in Chapter 3, where the aim is to identify objectives before the alternative generation step, could reduce anchoring on an already existing alternative when finding new ones.

It is important to understand that the use of formal training can help make people more familiar with the concepts of decision analysis and the value of using this approach (Keeney, 2004). For example, within weather forecasting experts have shown to make better judgements with training and experience. However, in many situations where one would benefit from the use of decision analysis, the values to elicit are one-time events (and thus frequencies cannot be used) or the situation is new, but training in decision analysis and probabilistic thinking could probably still improve assessments made in situations new to us, especially probability assessments. Moreover, motivation to use decision analysis (if we understand the value of it) could increase the usage of decision analysis applications to support decision making processes.

Suggested practical techniques for elicitation are to a great extent a matter of balancing the retrieved quality of the elicitation with the time available and cognitive effort demand on the users for eliciting all the required values. Sensitivity analysis could be used to study the consequential variations in the input provided and identify the information most critical for the results, which may need to be considered and specified more thoroughly. This could save users both time and effort, by making the elicitation step of the decision process simpler and faster as well as reducing the cognitive load.

CHAPTER 5

5. METHODOLOGY

The definition of knowledge about the social world and its nature depends on epistemological considerations, and a central issue within this context is whether or not the social world can and should be studied using a natural science model of the research process (Bryman, 2004). To a large extent the opposing views of positivism and interpretivism affect choices of research strategies (*ibid.*). Positivism affirms the importance of imitating the natural sciences, whereas supporters of the interpretative view argue that reality is indeed a social construct in which the researcher's subjective interpretation of reality provides knowledge of the object under study.

A common distinction is also between deductive and inductive ways of reasoning (Williamson, 2002). In deductive research, the researcher uses already established theories and deduces one or more hypotheses which are subjected to empirical analysis, whereas inductive research, on the other hand, involves drawing generalizable conclusions out of local observations, and progresses from empirical studies to the induction of new theories.

Another division is often made between quantitative and qualitative research. Quantitative research is predominantly influenced by the positivist view on reality and can be viewed as a research strategy that adopts a deductive approach, emphasizing on the testing of theories. The focus is on quantifying the collection and analysis of data. In contrast, qualitative research mainly adopts an inductive approach and focuses on generating theories. However, as e.g. Bryman (2004) argues, the connections between epistemological views and research approaches should be viewed as tendencies rather than definite connections, and the differences between quantitative and qualitative research should not be overdrawn. Furthermore, the combination of quantitative and qualitative methods in

the same study can be used to obtain results that can strengthen and confirm each other respectively (Campbell, 1988; Bryman, 2004; Yin, 2003).

Design research is another approach (complementing the positivist and interpretative perspectives) for analysing the use and performance of designed artifacts. Design research permeates several research disciplines, with origins within engineering, computer science and management, employing a variety of methods and techniques. The focus is on understanding a phenomenon (where all or part of the phenomenon may be created as opposed to naturally occurring) and finding solutions to observed problems. Simon (1996) distinguishes between a natural science, which is a body of knowledge about a class of objects or phenomena describing their behaviour and ways of interaction, as opposed to a science of the artificial, which is a body of knowledge about artificial objects and phenomena designed to meet certain desired goals and to function. March and Smith (1995, p. 253) state that “Whereas natural science tries to understand reality, *design science* attempts to create things that serve human processes.” The design methodology follows a general process design cycle, suggestion for solution (motivated by awareness of a problem), development of an artifact, evaluation of the artifact, and the cycle is iterated until the specific requirements are met (Baskerville et al., 2009). In contrast to positivist research, where the evaluation either confirms or contradicts a hypothesis, initial hypotheses in design research are seldom abandoned, but instead yield additional information regarding the design and/or behaviour of the artifact, resulting in iteration back to the suggestion step of the design process cycle. Research activities can be divided into two design processes, *build* and *evaluate*, where the build activities concern the construction of the artifact and evaluate activities deals with the development of criteria and assessment of artifact performance against these criteria (March and Smith, 1995).

5.1 CONDUCTING AND VALIDATING RESEARCH WITHIN DECISION ANALYSIS

Decision theory and analysis has evolved through research within a number of different disciplines, such as economics, psychology, management science and artificial intelligence to name a few, and theories have thus been developed under different scientific paradigms. The traditional distinction between normative and descriptive choice models also has implications on developmental concerns.

The normative tradition is characterized by ideal inference or decision processes, and is concerned with how idealized people should act in order to be rational, without ensuring that such ideal conditions are met in reality. The decision of what axioms and formal rules to employ are motivated by what some researcher thinks is logical or rational behaviour (Bell et al., 1988). If the mathematical implications of dropping or modifying an axiom in some way are perceived as rewarding by the researcher (e.g., aesthetically pleasing), or if the researcher notices a better correlation between the abstract system and the real world (empirically), or if the researcher merely imagines the suitability of the proposed set of axioms, this motivates development within the normative discipline (ibid.). Consequently, many of the advances within decision analysis research have been theoretically developed and implemented in practical applications without much empirical testing.

The descriptive tradition on the other hand is substantially empirical, and merely studies individual behaviour without trying to modify or influence it with the aim of finding patterns in human decision behaviour (how people actually do infer and decide) in risky choice situations. Descriptive science has a strong tradition within psychological and organizational theory, and research on decision making processes within organizational settings is commonly conducted in the form of case studies.

Within the prescriptive tradition, researchers are concerned with models and decision aids that are useful for real people (and not some idealized person). It can be described as the fusion of normative and descriptive research, which results in a multi-paradigm research community. The diversity of contributors to the prescriptive decision analysis tradition makes it difficult to have a universal agreement on the study of interest and the suitable research methods to employ. Brown (1989) suggests that the earlier tools that were developed in the academic realm within the field of decision analysis were inadequate to help real world decision making. Brown, furthermore, argues that the development of prescriptive research progresses on a more divergent, "follow your nose", basis, which limits its development within the academic realm as more convergent research methods, where the scope can be defined in advance, are more accepted.

The aim of scientists employing design science is to create models, methods and implementations that are useful (March and Smith, 1995), which is in conformity with the aim of prescriptive scientists. March and Smith (1995, p. 254), moreover conclude that "design science offers prescriptions and create artifacts that embody those prescriptions," which is in line with prescriptive decision science. Yet, behavioural science is important too, as such research provides awareness of problems/difficulties that we need to solve. Consequently, both

research paradigms are important to prescriptive decision analysis research as it merges people, technology and organizations, and focuses on identifying problems and finding solutions to these problems through the design of artifacts.

The natural evolution of decision aids comes from the practitioner, who learns from his/her mistakes (Brown and Vari, 1992). Brown and Vari (1992) suggest that the engineering design principle build-test-build-test is a promising research paradigm for prescriptive purposes. In contrast to this principle is the social scientist's aim to establish universal human truths based on empirical testing. Brown and Vari (1992) assert that a fairly standardized research pattern is to use real cases as test beds for prescriptive ideas. The prescriptive advice has to be adapted to the individuals for whom it is intended, since real people are different. The fact that artifact performance is related to the environment in which it operates poses a challenge to the evaluation of prescriptive models, methods, and implementations, as their performance is related to their intended use (March and Smith, 1995). The evaluation criteria for an artefact must be determined for the specific context in which it will be used (*ibid.*). Moreover, design science has a higher tolerance for ambiguity than what is generally acceptable within positivist research (Hevner et al., 2004). Keeney (1992b, p. 58) states that the three traditions within decision theory "are distinct in terms of the breadth of their problem focus, the criterion for appraising appropriate axioms, and the judges who apply the criteria." Bell et al. (1988) present three different criteria used for evaluating decision making models. Normative models are evaluated by their *theoretical adequacy*, i.e. the degree to which they present suitable idealizations or rational choice. Descriptive models are evaluated by their *empirical validity*, i.e. their correspondence to observed choices, the extent to which it accurately characterizes and predicts people's behaviours. Lastly, prescriptive models are evaluated by their *pragmatic value*, i.e. their ability to provide decision-makers with appropriate aid in order to help them make better decisions. All three criteria are difficult to define and evaluate (*ibid.*).

Prescription is closely related to usefulness and prescriptive analysis can, e.g., try to eliminate or reduce descriptive violations of normative rules. Thus, the goals we set out to achieve are what we later can use for evaluation. The requisite model described by Phillips (1984) is the type of model most commonly used for prescriptive decision analysis, where decision theory is its source and a socially-shared understanding of the problem is its subject. An important part of the development of prescriptive decision aid is their evaluation and a relevant practical criterion for evaluation is to evaluate their performance (Brown and Vari, 1992). Moreover, Brown (1989) argues that for prescriptive purposes, there should be a balance between technical soundness (i.e. the technique should

appropriately include all the relevant information and judgment available to the subject that may contribute to the quality of the decision), cost (including delay and cognitive burden), and acceptability (both institutional and psychological), which makes the solution highly situation specific. Technical soundness and the resulting decision quality are the most difficult to determine as it is difficult to assess how “good” a decision is, and it is just as difficult to assess what the use of another decision-aiding technology would have resulted in. Brown (1989) therefore argues that the validation of prescriptive technology should be primarily external, i.e. tools should be tested in real settings. Nonetheless, internal validation, such as testing of logical coherence and consistency with theory, has been the dominant approach within, e.g., management science (ibid.).

5.2 VALIDATING ELICITATION METHODS

An important part of prescriptive research is how best to elicit probabilities, and utilities/values (and weights if an MAUT approach is taken). These elements are the main inputs of decision analysis applications, and as already accounted for, numerous elicitation methods (and remedies to how best to solve related problems) have been suggested. Such methods are concerned with balancing the quality of the elicitation with the time and effort required from the user. The goal of these techniques is to elicit an accurate description of the expert’s and/or decision-maker’s beliefs and preferences. The better the elicited values reflect the expert/decision-maker’s real degree of personal belief or preference, the better the method. Evaluating directly whether an elicitation method has produced accurate input data is a great challenge (showing their closeness to the “true” values), as we would need detailed knowledge about the expert/decision-maker’s knowledge and personal preferences. As it is problematic to conclude that a method is better than another based on objective information, less direct methods for evaluation are used. However, as we are dealing with subjective factors, it is difficult to identify and reach consensus on adequate measures for evaluating elicitation methods in the prescriptive decision analytic context.

According to Roy and Mousseau (1998), there are essentially two (extreme) ways to regard the preference elicitation process, the *descriptivist* and the *constructivist* approaches, although the philosophy most commonly adopted by analysts in real world cases corresponds to a median position. The descriptivist approach views preferences as already existing in the mind of the decision-maker and this pre-existing preference structure is not altered by the elicitation procedure (its role is to “match” already defined existing preferences). Beattie

and Barron (1991) reflect on the fact that such true estimates may remain constant at all times, but can become distorted in the elicitation process. The constructivist approach, on the contrary, does not view preferences as fully established in the mind of the decision-maker, but that values instead reflect statements expressed by the decision-maker during the elicitation process.

When probabilities are assessed by experts/decision-makers, they reflect the degree of personal beliefs held by the assessor. Elicitation methods in probabilistic models are often evaluated using one of three lines of reasoning: (1) expert's preference, (2) benchmark model, and (3) performance measure (Wang et al., 2002). The first approach assumes that the expert's preference for a method will result in better quality probabilities, although preference does not necessarily correlate with precision (*ibid.*). In the benchmark model approach, the elicited probabilities are compared with an existing benchmark model, and the level of precision is assessed by comparing the elicited probabilities with the benchmark model. However, as Wang et al. (2002) emphasize, an underlying assumption is that such a benchmark model is shared by all experts, i.e. presumably the correct set of probabilities. Yet, there are problems with finding such benchmark models as experts frequently disagree with each other (Morgan and Henrion, 1990), and how do we know that we have found the correct model when managing subjective values in the first place. The third approach involves a more pragmatic view and evaluations are here performed using various methods to compare the predictive performance of the derived probabilities. Yet, as Wang et al. (2002) point out, we cannot necessarily assume that the quality of the resulting probabilities is correlated with the accuracy of the elicitation method. In general, a common approach is to set evaluation measures that seem plausible at the onset, and thereafter use them to evaluate the elicited probabilities.

Weight elicitation methods are often measured on convergent validity, internal consistency and sometimes external validity. Convergent validity relates to whether weights derived from different methods agree (numerous studies have shown that this seldom applies). Internal consistency is often measured by eliciting weights on two occasions, a test-retest experiment and study the consistency of the weights. This provides that the subjects have not changed their opinions, which can be the case in real settings, where preferences can change and evolve during the decision process (*cf.*, e.g., Danielson et al., 2010; Riabacke et al., 2011; Riabacke et al., 2012). The measure could also be called reliability (although the reliability of the method is closely entwined with the reliability of the person making the judgments). External validity relates to the comparison of assessed weights with a set of externally elicited weights, but there is very little research regarding this measure of consistency. In Borchering et al. (1991), the authors

admit to the fact that “correct” external weights can be difficult to find, yet suggest that weights elicited from knowledgeable experts in carefully designed interviews could be used for this purpose. Since weights are subjective measures reflecting a person’s personal view on the importance on different criteria, this suggestion for how to derive external weights seems highly questionable. Perhaps this measure should not be used for weight consistency checks, since we cannot really judge the “true” weights of a person.

5.3 METHODOLOGICAL APPROACH OF THESIS

Taking a prescriptive stand when studying decision making processes and decision analysis yields a multidiscipline research field, since prescriptive research merges issues relating to people, technology and organization, and lies in the borderland of normative and descriptive decision theories. Consequently, such research has its origins in and has been developed under different scientific paradigms, and therefore, a diversity of research methods has been used in this thesis.

Prescriptive research is concerned with models and decision aids that are *useful* for guiding real people, and an important step in development is their evaluation. In comparison to the normative tradition, where validation essentially has been a matter of concluding that the model of rational choice is reasonable, and descriptive tradition, where validation is mainly concerned with the extent to which a model accurately characterizes and predicts human behaviour (empirically), relatively little work has been done within the prescriptive domain. Bell et al. (1988) state that the criterion used for evaluating prescriptive decision making models is their pragmatic value, i.e. their ability to provide decision-makers with appropriate aid in order to help them make better decisions (see Table 1, Keeney’s, 1992b, classification of the three categories of decision theories). This criterion is difficult to define and evaluate, and there is no universal set of distinct general prescriptive measures for evaluation of prescriptive decision aid. In accordance with the great variety of objectives, decision models and methods within prescriptive research, it seems adequate to develop and employ more than rational criteria for the various types of problems and settings. As prescription deals with how to benefit from the application of formal methods of decision making in real settings, both normative and descriptive methods for evaluation could be used, and a parallel can be drawn to the evaluation of artifact performance, which March and Smith (1995) conclude is a challenge as the environment is relevant, and consequently, the evaluation of prescriptive models,

methods and implementations is related to their intended use. Thus, criteria for evaluation are context dependent within prescriptive research, and the utility or usefulness of an artifact is assessed by means of measuring performance against the goals or requirements (criteria) initially set. Following the reasoning in the previous sections, the methodological approach of the research conducted herein will be described in more detail in the following.

In paper I, *A Study of Framing Effects in Risk Elicitation*, the main focus was on further exploring framing effects (cf., e.g., Tversky and Kahneman, 1986; Fischer, 1989), an important part of descriptive theory, which cause significant problems to the normative theory of risky choice. Although framing problems are well-known to researchers within the field, they are still not part of normative theory. Yet, knowledge of how to reduce such effects is prescriptively important when using normative principles in decision-making settings. Thus, we chose to combine descriptive and normative theory, and conduct a quantitative study in order to further investigate people's choice behaviour when presented with questions of the character commonly used within utility elicitation. Our hypotheses were based on descriptive theory, but were measured normatively using the strict normative rule of maximizing the expected value. Hence, we tried to find patterns in human behaviour in risky choice situations evaluated against normative criteria for rational behaviour. The normative, fundamental assumption of *procedure invariance*, i.e. that normatively equivalent methods for elicitation should give rise to the same results, was explored empirically. The combinatory research design was motivated by the prescriptive stand taken, and a total of 240 undergraduate students participated in the study. They answered questionnaires with choices of differently framed risky prospects with the aim of studying how different presentation formats of the mathematically equivalent prospects affected their choices.

Paper II, *How Different Choice Strategies Can Affect the Risk Elicitation Process*, is an extension of the study conducted in paper I, where quantitative data was collected from 240 undergraduate students. The findings of the study conducted in paper I gave no further information on the subjects' perception, ways of reasoning or whether they employed any specific strategy in their choices between the risky prospects, why an extension to the study was motivated in order to further study these aspects. In the extended study, qualitative methods were used as a complement to the quantitative approach previously taken. The former study was complemented with an interview study using an additional 12 respondents. Semi-structured interviews of about 30 minutes each were conducted with questions regarding each of their choices (probing was used when needed). The questions were concentrated on choices where we had seen the

strongest tendencies in behaviour to deviate from normative rules. As previously mentioned, the combination of quantitative and qualitative methods can be used to obtain results that can strengthen and confirm each other respectively (Campbell, 1988; Bryman, 2004; Yin, 2003).

Paper III, *Risk Elicitation in Precise and Imprecise Domains - A Comparative Study, Sweden and Brazil*, describes a comparative study between Sweden and Brazil, involving 120 graduate students from both countries. This study was a continuation of the studies conducted in papers I and II. As in paper I, a quantitative approach was employed, where risky prospects were offered to students with the aim of studying whether there are differences in choice behaviours when the uncertainty is expressed as intervals vs. the traditional point estimate representation used in the previous studies as well as when the EMV of prospects are explicit.

In papers IV, *Employing Cardinal Rank Ordering of Criteria in Multi-Criteria Decision Analysis*, and VII, *Cardinal Rank Ordering of Criteria - Addressing Prescription within Weight Elicitation*, the lack of more prescriptively useful weight elicitation is in focus. The approach of design research - to identify problems and find solutions to these problems - reflects the continuation of the research even more. Based on identified problems in existing theory and the findings of papers I-III, a novel method, Cardinal Rank Ordering of Criteria (CROC), for weight elicitation is proposed in line with the design of an artifact.

Based on existing theory and the findings of the studies presented in papers I-III, an implementation of the CROC method's extraction part was designed and tested in an iterative design study (described in paper VII), where the first phase tested its functionality. A challenge to artifact performance is that prescriptive methods are related to the environment in which they operate, and evaluation criteria must be determined for the specific context in which a method will be used (March and Smith, 1995). An important part of prescriptive analysis is, among other things, to try to reduce or eliminate descriptive violations of normative rules in order to increase the usefulness of normative methods in real settings. Thus, the validation of the pragmatic value of the CROC method was initially tested in the second phase of the design study, where subjects participated individually on two occasions with one week in between. On each occasion, an hour long semi-structured interview was conducted, where the subjects used three different weighting methods (CROC as well as two commonly used methods in practice), in order to test whether CROC could reduce some of the identified problems more than the existing methods. Subsequently, the subjects answered questions regarding each of their choices (probing was used when needed). In order to take different types of users into

consideration, the participants were heterogeneous with regard to their educational background, age, and profession.

As already mentioned, many of the advances within decision analysis have been theoretical developments, which have been implemented in practical applications without external validation (cf, e.g., Brown, 1989). Yet, the prescriptive aim is to increase the usefulness of normative methods in real settings and guide decision-makers by recognizing descriptive realities. This motivates external validation as an important aspect to consider, if possible. Brown and Vari (1992) suggest that we need to use real-life test beds in order to retrieve more prescriptive (practically useful) methods, why the CROC elicitation method (which theoretically could reduce many of the major hindrances to real-life usage) was also employed in two real-life cases and tested with real decision-makers. The promising results of the first real-life study (more thoroughly described in paper V) promoted the use of the CROC method for weight elicitation in the second real-life case (described in more detail in paper VI). The results of the first case study were also compared to the employment of the Point Allocation method in an analogous (third) case.

External validation of the CROC method was done both qualitatively and quantitatively. Some of the main aims with the extraction phase in CROC were not to force users to express unrealistic precision, not putting too much cognitive effort on users and not requiring too much time. Such aims are quite difficult to validate adequately as they are measured on a subjective (qualitative) basis and some people may find any effort exerted tiresome whereas others are quite tolerant. Yet, from the results of the case studies, the requirements were reasonably fulfilled. Looking at interpretation, it is reasonable to say that although there are many methods for weight elicitation, it is virtually impossible to find a method, which always generates the same weights when employed on more than one occasion even though the decision-maker has unchanged views. The CROC method was employed twice and in the analysis, the changes in user inputs between the two occasions were studied and how the interpretational part of the CROC method handled variations quantitatively.

In paper V, *Transparent Public Decision Making - Discussion and Case Study in Sweden*, the decision process, which included the employment of the proposed CROC weight elicitation method (described more thoroughly in paper IV), was validated in a real-life case. As already mentioned, external validation is something prescriptive scientists highlight as an important aspect of the evaluation of prescriptive decision process/decision analysis development, and something that is still quite rare within decision analysis research. Brown and Vari (1992, p. 41) state that evaluating the performance of prescriptive decision aid can be based

on “process evaluation focusing on the characteristics of the decision-making process aided by a technology, e.g., use of information, goal-centeredness, flexibility”.

In the Örebro case, the decision-makers employed a prescriptive decision process aided by a decision-analysis tool, which included an implementation of the CROC method for eliciting criteria weights, throughout their work. The validation of the case involved collecting both the opinions of the decision-makers and the experts involved in the assessments of values. One can differentiate between *useful* systems and *used* systems, and in this case, both criteria applied as they both found the methodology useful and used it throughout their work.

Paper VI, *E-participation galore? Extending Multi-Criteria Decision Analysis to the Public*, focused on further investigating the use and usefulness of the individual weight elicitation method (CROC) employed in paper V as part of a new approach to increase participation in democratic decision making. A lack of tools for problem understanding and scaling up participation to large scale use within the field of eParticipation initially promoted the study described in paper VI.

In this study, 90 students at upper secondary school, aged 17-19, participated in the study. The students were guided to a wiki web site with information about the decision problem and structure, the necessary steps of the decision process model, a link to the interactive elicitation method (the students were guided through each step of this process), and a link to a forum (for those part of the second group described in the following). The participants used the elicitation method on two occasions and the participants were divided into three groups with different forms of discussion and deliberation in between. One group discussed the problem presented and their different views in class for about an hour, the second group communicated about the problem in a web-based forum (mainly simultaneously for about an hour), and the third group were a control group with no explicit instructions to discuss the problem and their preferences (but had no restrictions not to). After the second occasion, they were asked to fill out a questionnaire. The two research questions, or evaluation criteria, initially determined:

1. If the elicitation tool was easy to use, individually and remotely as well as understandable within a reasonable amount of time (requirements for scalability), and
2. If the students' use of it was “mature”, i.e. if they understood the problem and options well enough to take part in a well-informed discussion (requirements for decision quality),

were motivated from a very practical point of view as mass participation usually focuses on one view of a problem and the aim here was to introduce “rational” participation considering and analysing different views in parallel. Thus, these two questions were the initial goals/requirements set for validation, and were partly answered through the questionnaires, but also from comparisons between the quantitative data collected on the two occasions.

CHAPTER 6

6. CONTRIBUTIONS

This thesis highlights the importance of adopting a prescriptive approach to eliciting decision information. In many decision making situations, where the use of decision analysis can be advantageous, there is a lack of precise objective data and, thus, the main decision data inputs are subjective in nature as they, in most cases, are elicited from experts with knowledge (in the form of belief statements) and decision-makers declaring what is important to them through preference assessments. Several models for decision making processes have been proposed in the literature, but there is still

1. A need for an extended elicitation model, and
2. A lack of methods for dealing with uncertainty.

The process model proposed in Chapter 3 is more fitting in practice than traditional models as it emphasizes the importance of structured elicitation of adequate input data throughout decision making processes. Moreover, observed problems with existing elicitation methods and the lack of methods better suited for real decision making processes motivated the work described in this thesis.

In order to further define the problematic aspects of elicitation, the empirical studies in papers I-III were conducted, where problems from existing literature were further studied. The problems with eliciting precise decision data suggests that using imprecise values within elicitation is a more realistic and useful approach to strive for, and perhaps even more important within preference elicitation. Imprecise domains were already in focus in paper III, and based on theory and the findings of the studies presented in papers I-III, an elicitation method for imprecise statements and noisy input was formalized in Riabacke et al. (2009). The work progressed and the method was further refined into a

weight elicitation method, Cardinal Rank Ordering of Criteria (CROC), presented in paper VII. The CROC method is both compatible with an adapted prescriptive decision making model, focused on a more structured elicitation component as well as algorithms for dealing with such data. The CROC method was employed and validated in two real-life cases (more thoroughly described in papers V and VI), which is not so common within decision analysis research, i.e. there are relatively few examples of real-life cases employing such techniques. More specifically, the individual contributions in each paper are as follows:

- In paper I (*A Study on Framing Effects in Risk Elicitation*), the importance of considering framing as one of the obstacles to sufficient elicitation is affirmed, and increased awareness of how presentation formats affect people's choices is suggested as one way of reducing such effects. The study generated information on how different forms of framing can yield different results in the elicitation process when using exact values during extraction, and awareness of how to reduce such effects during this process. The prospects in the negative domain, i.e. dealing with losses, were shown even more sensitive to framing, and here a decreasing probability order was shown to deviate less from normative rules. In the positive domain, an increasing probability order was shown preferable.
- In paper II (*How Different Choice Strategies Can Affect the Risk Elicitation Process*), the aim was to further investigate perceptions, ways of reasoning and whether any specific choice strategies were employed in choices between the risky prospects studied in paper I. Three main strategies that the subjects used when choosing among the alternatives in the offered prospects could be identified. Behaviours differed in these categories depending on whether probabilities were high or low, chance or risk domains, and the order of the probabilities. Some interesting characteristics that could be identified were that many of the respondents expressed their evaluation strategy in terms of converting probabilities expressed as percentages into frequencies, and that some respondents felt a tendency to overweight "familiar" probabilities, such as 0.25 and 0.75, in choice alternatives. Moreover, some of the respondents intuitively recalculated the risk of losing into a chance of not losing at all, and there was a tendency to value certain monetary differences as equal to an increase or decrease in probability irrespective of the sizes of the values. Such biased judgment due to inappropriate scale/value trade-offs is an important observation of erratic behaviour as the impact of an increased/decreased chance or risk of 0.1 differs considerably depending

on the sizes of the values. The results of this complementary study indicate that different people may prefer alternative presentation formats than those traditionally employed within elicitation using precise estimates to represent uncertainty. Moreover, results confirm that there is a lack of elicitation methods useful for different user strategies. The findings of the study in combination with the fact that even experienced statisticians make mistakes when reasoning about probabilities and the lack of precise data, which is often the case in real settings, motivated further studies on the use of alternative representations of uncertainty (such as intervals). This initiated thoughts of continuing the initial two studies with the study described in Paper III.

- In paper III (*Risk Elicitation in Precise and Imprecise Domains - A Comparative Study, Sweden and Brazil*), the results indicate that using an interval representation of uncertainty can be functional in the extraction step of the elicitation process. As available information is often imprecise and even conflicting, and preferences are vague and can be inconsistent in real settings, studying people's perception of prospects where uncertainty is expressed as intervals is highly relevant to prescriptive elicitation. There is relatively little research done regarding people's perception of imprecise representations of uncertainty, and therefore this aspect was interesting to compare to precise estimates. Also, the most common approach in decision making situations using decision analysis is to look at the expected monetary value (EMV) of each alternative in the analysis, why this was important to include in the study as well. In addition, as most tool development is carried out in the Western parts of the world, and the prescriptive view should take contextual factors into concern as well, the possible cultural difference was of interest to study. Differences in cultural context were, therefore, considered as the participants (who were similar regarding parameters such as age, and economic and educational background) came from both Sweden and Brazil. The interval representations generated similar behaviours from subjects in comparison to point estimate representations when using the mid-point of the interval in the interpretation. Moreover, the study showed that the Brazilian group were more affected by the EMV information than the Swedish participants during choices between risky prospects. Accordingly, the findings of this study confirmed previous results of the need for more flexibility within decision analysis tool development. This is highly relevant for elicitation as the demand for precise estimates poses problems as previously mentioned, and using less precise data,

like intervals in this case, could be positively perceived by respondents during extraction. Consequently, a more multidisciplinary approach within decision analysis application development seems necessary to improve the practical usefulness of these tools as different cultural aspects can affect ways of reasoning and what information people consider important for their choice.

- Based on the findings in papers I-III and already existing literature, a novel (imprecise) weight elicitation method was initially formulated in Riabacke et al. (2009) and subsequently refined into the CROC method presented in papers IV (*Employing Cardinal Rank Ordering of Criteria in Multi-Criteria Decision Analysis*) and VII (*Cardinal Rank Ordering of Criteria – Addressing Prescription within Weight Elicitation*). Classical models for weight elicitation can broadly be described to fall into one of two categories, either a category of methods using precise values during extraction, or rank-order methods extracting ordinal statements only. As discussed in this thesis, methods belonging to the first category demand too much from users, whereas methods belonging to the second do not utilize weaker forms of cardinality that may exist. In papers IV and VII, the problems arising from the discrepancy between the current treatment of elicitation within decision analysis applications and the ability of real decision-makers to provide what is required is highlighted. In particular the following identified problematic elements were of interest to tackle: 1) reduce “noise” in input, and 2) reduce cognitive demand. Both existing empirical studies as well as the studies presented in papers I and II show the vulnerability in using exact values during extraction as such inputs are more susceptible to “noise”. Not only framing constitutes a problem, but also the numeric order of presentation (paper I). In paper II, biased judgment due to inappropriate scale/value trade-offs could be identified. The employment of intervals to represent uncertainty during extraction was tested in paper III as an attempt to avoid making exact judgments during extraction. However, users still had to make value judgments on the interval values, and views on whether it was easier or harder than judging exact estimates differed. Consequently, trying to reduce the cognitive demand on decision-makers further, the concept of using graphical impreciseness was tested in papers IV and VII. More specifically, the use of a graphical procedure that did not demand exactness in decision-maker judgments during extraction as well as the use of a calibrated scale were some of the characteristics derived necessary from previous studies. The CROC

method with the following characteristics is proposed: 1) The use of imprecise graphical extraction of both ordinal and cardinal criteria, 2) Cardinal comparisons are used for representation, and 3) The extracted information is interpreted as centroids. An implementation of the method is tested in two real-life studies, which both included two elicitation occasions (in line with the process model described in chapter 3), and the results of the first case were compared to the employment of the Point Allocation method in an analogous third case. In both real-life cases employing an implementation of CROC, the method's extraction part worked well during the user interaction with the tool. Moreover, the outcome of the comparison between the first case employing CROC and the case employing PA indicates that CROC is more robust and persistent to noisy input. In all three real-life case studies, people's preferences between some of the criteria were shown to be partly dynamic between the two elicitation occasions and the use of the preferred representation and interpretation (as described in the paper) of importance statements is a way to cover the dynamics of preference entities. Slight changes between occasions (most likely unintended) are represented in the constraint set generated by the proposed interpretation of extracted information. As there appears to be a prominent vagueness even in the ordinal weight information provided by the decision-makers (one step up or down in the priority order of the criteria is perceived as unchanged views), the proposed interpretation of the users' interaction with the tool within such contexts seemed conformable to their preferences. Both case studies employing CROC are more thoroughly described in papers V and VI.

- In paper V (*Transparent Public Decision Making – Discussion and Case Study in Sweden*), the CROC weight elicitation method was employed as part of a prescriptive transparent decision process, where the decision-makers were aided by a decision tool that incorporated the method. The decision making process, the decision tool, and the weight elicitation method were used and considered useful by the decision-makers. Their decision making process became more ordered than usual within their domain, since they followed a structured decision process supported by the tool at the different steps. The initial difference in view of what constituted the problem at hand (this difference was not explicit at the beginning) clarified during the identification of objectives (at the onset, they had different views on the definition of their decision) as well as during the weight elicitation. The employment of the CROC method on

two occasions in order to initiate reflections on their preferences early on and to provide time to contemplate, was shown useful and helpful in the sense that their understanding about criteria, the problem as well as their preferences increased. Moreover, as already mentioned, the dynamicity of in-between criteria in the preference order was confirmed, and strengthens the imprecise interpretation approach (as opposed to precise numeric representations) of preference statements. The process, also helped with the generation of alternatives, and the result of the analysis was perceived as valuable to their final decision.

- In paper VI (*E-participation galore? Extending Multi-Criteria Decision Analysis to the Public*), the use and usefulness of the weight elicitation method developed and employed in papers IV and V, was further explored. In existing literature on (e-)democratic decision making, two major problems have been highlighted. Deliberation tools discussed under the label of “eParticipation” typically yet have to find ways to scale up to cater for very large amounts of users so as to be generally useful for large scale participation. Rational decision making tools discussed under the label of decision support systems can technically cater for large scale use, but typically require users with some analytical inclination, which is obviously a serious restriction. In this paper, the CROC method was adapted and tested as part of an innovative approach to achieve a balance between participation and expertise with the purpose of increasing and scaling up participation in democratic decision making. The study tested whether the model could be realistically applied to large scale democratic decision making, i.e. involving non-sophisticated users and very limited time; two criteria which are both realistic when involving the public in decision making activities regarding complex issues, and also economical in view of large scale use. The elicitation tool was found easy to use individually and remotely as well as understandable within a reasonable amount of time (requirements for scalability). Moreover, the new approach to large scale participation was also shown “rational” in the sense that the decision process was communicated to and employed by the participants; the subjects understood the problem presented well enough to take part in discussion as well as their task of contributing with their preferences regarding the identified evaluation criteria by using the elicitation tool. Moreover, a majority felt that expressing preferences regarding criteria was much easier than taking a stand directly on different alternatives. Consequently, the novel ap-

proach to increase participation was shown promising and practically useful.

CHAPTER 7

7. CONCLUDING REMARKS

In spite of their potential, the employment of decision analytic applications to support decision making activities in businesses and organizations is still very limited in practice. Yet, as highlighted in this thesis, the more complex a decision gets, the more these tools could be of benefit in the structuring and analysis of the decision faced in order to retrieve more clearly structured and comprehensible decision material to use as a basis for the decision. However, existing tools are too incomplete to support decision processes sufficiently (cf., e.g., French and Xu, 2005), and especially the step of extracting information during these processes is not supported adequately. The current use of decision analysis is mainly as expert tools for analysis of decision problems, whereas we believe that their use would be of greater benefit in organizations and businesses if they were used to support the decision making process at an earlier stage (and throughout the process), and not only by a small group of experts. At present, there is little guidance to decision-makers and experts in how to elicit the required decision data, i.e. there is an assumption that users are able to assess the required information adequately. Yet, the elicitation of information (extraction, representation and interpretation) is an important part of a decision process, and there is a need for more prescriptive elicitation methods (realistic and useful) than what is offered today. In turn, we would attain the potential for better decision quality with a more clear motivation of decisions made.

7.1 SUMMARY

This thesis stresses the importance of adapting decision making processes to include elicitation more explicitly in line with the model proposed in Chapter 3. Moreover, problematic elements with current elicitation models as well as important issues to address have been identified. Based on recognised difficulties in the literature as well as results from the empirical studies presented in papers I-III, means to reduce problems relating to elicitation are suggested. Moreover, elicitation is divided into three conceptual components in order to categorize methods and analyse elicitation more explicitly, namely:

1) Extraction

This component deals with how information (probabilities, utilities, weights) is derived through user input.

2) Representation

How to capture the retrieved information in a formal structure, i.e. the internal format used to represent user input.

3) Interpretation

Is dependent on the expressive power of the representation used and how to assign meaning to the captured information in the evaluation of the decision model used.

An analytic division of this kind is not found in literature within the field and is a useful way to examine elicitation methods in order to recognize their characteristics and identify elements that can impact their applicability in practice. In practice, we should strive for finding methods that are less cognitively demanding and less sensitive to noisy input within each component.

An important impediment of current elicitation methods for eliciting experts' and decision-makers' beliefs and preferences is that they are highly vulnerable to fundamental assumptions within normative decision theory. Behavioural aspects, like the heuristics and corresponding biases people use during extraction (cf., e.g., Kahneman et al., 1982), are important to be observant of in order to reduce deviations from normative assumptions. We believe that such elicitation problems could partly be attributed to the use of exact values during extraction as such inputs are sensitive to small variations in statements (noisy input) and people have difficulties formulating their beliefs and preferences with numeric exactness. The studies presented in papers I-III, confirm the vulnerability to framing effects when employing precise numeric expressions during extraction, and results add support to adopting an imprecise elicitation

approach in real settings. Moreover, regarding preferences, one can question whether exactness of such values should be used at all as human preferences are not naturally represented with the exactness required in our minds, their character is not static and we cannot really know the “true” values. Thus, the problems with eliciting precise decision data suggests that using imprecise values within elicitation is a more realistic and useful approach to strive for, and perhaps even more important within preference elicitation.

However, adopting an imprecise approach need not overlook the extraction of cardinal input totally as there often exist weaker forms of cardinality, although not precise. The CROC weight elicitation model suggested in paper VII is such an approach as it uses both ordinal and imprecise cardinal input during extraction and interprets input statements as regions of significance. The CROC approach was also implemented and validated in decision analytical tools in papers V and VI. In paper V, the decision analytical tool was used throughout the different steps of the decision making process, and the CROC method was employed twice in order for preferences to stabilise. The decision-makers found the method practically useful. Post validation was used to test whether input data was sufficient enough for evaluating the decision, i.e. whether the output was satisfactory to base the decision on. Such post validation can be a useful approach in real decision making processes as the time and effort of elicitation can be reduced by an initial analysis of a decision implicating if there is need to specify input (and if so, what input) in more detail for satisfactory output. This practice was appreciated by both decision-makers and experts involved in the assessments of input data.

The focus of paper VI was to further investigate the use and usefulness of the individual weight elicitation method (CROC) employed in paper V as part of an innovative approach to increase public participation in democratic decision making. Two particular major challenges associated with this promoted this study namely, a lack of tools for scaling up participation to large numbers of users and the balancing of participation and expertise. A web-based implementation of the CROC method interactively guided the users throughout the step of elicitation to support their individual and remote use. The findings suggest that the method can be used for large scale participation during a decision making process, but also that a participatory process is improved by lengthier deliberation and more than one point of measurement so that opinions can stabilize.

To conclude, the CROC method seems like a promising approach to more pragmatic weight elicitation based on problematic elements identified in the literature as well as the findings of the empirical studies presented in papers I-III.

A big challenge has been to merge theory with practical findings and needs. In the cases employing CROC, the method proved to be both usable and useful to support the elicitation component of the decision making process. Users were able to employ the method in practice and were supported in expressing their preferences without forcing them to express unrealistic precision or to state more than they were able to. However, there is still a lot of research needed in order to further reduce the gap between decision analysis tools and their users, both for single decision-maker situations as well as more complex situations involving multiple heterogeneous decision-makers.

7.2 FURTHER RESEARCH

The research presented in this thesis highlights the importance of developing and employing more prescriptively useful elicitation methods to improve prescriptive decision processes supported by decision analysis tools. In this particular context, there is a need to adapt elicitation methods to realistic usage, both regarding people's abilities as well as needs.

The inability of people to quantify beliefs, and perhaps even more so, their preferences, during the step of extraction suggests that a promising line of direction to move in is to allow for less precise inputs while still conforming to a rational model of choice. In order to do so, there is a need for adapting models for decision making processes to vaguer forms of expressions, which can be supported by a decision analysis tool throughout the different stages of the process in order to attain improved decision material. In addition, as the extraction of information (elicitation) is required for the analysis, an information model describing how to retrieve and represent the required decision data inputs sufficiently is needed. The information model employed is dependent on the expressive power of the analysis method employed, i.e. what type of input it permits and can handle in the evaluation.

Moreover, there is a need to develop elicitation methods that are more flexible to various user needs. Especially within preference elicitation the trade-off between obtaining more and better quality information and the cost of further elicitation effort is a challenge demanding further research. Elicitation costs can entail such aspects as cognitive load, understanding, exerted effort and time consumption on behalf of decision-makers and users, and should not outweigh the worth of obtaining the information, i.e. the improved output produced by obtaining it.

Furthermore, for the prescriptive decision analysis approach, there is a need to balance theoretical soundness (research) with applied practice. The CROC weight elicitation method suggested in paper VII is such an approach as it is based on existing theories and empirical findings from the studies presented in papers I-III. The CROC method is both compatible with an adapted prescriptive decision making model, focused on a more structured elicitation component as well as algorithms for dealing with such data. CROC adopts an imprecise approach to elicitation and is sufficiently effective for supervised use, but unsupervised use of CROC needs to be studied further in order to identify possible obstacles and add sufficient support regarding:

- **Extraction**

There is a need to identify whether its extraction phase needs refinements by further studies of the intended domain. Thus, the use of real-life cases to increase its prescriptive usefulness and promote its individual use by decision analytical novices is required. The social interaction part of the decision making process is important to decision-makers to assure that they have acted properly and just (March, 1994). Consequently, the use of real-life cases as test-beds for prescriptive progress is also important in order to include this dimension of the decision making process during prescriptive development.

- **Interpretation**

There is a need to further study how to approach decision situations where the interpretational step generates data which is unsatisfactory for the decision evaluation, i.e. where more precision is needed in order to generate more specified results.

Another area where prescriptive elicitation can be useful is within participatory approaches as suggested in the study presented in paper VI. The findings of this study, indicate that the novel approach to increase public participation in democratic decision making, which included an implementation of CROC, was a successful initiative. Although results were promising, further studies still need to be done for more definite conclusions regarding its usefulness within this realm. The proposed setup needs testing in other contexts, both regarding different types of decision making problems as well as cultures, and with larger, heterogeneous samples of the population. There may be a need for different types of extraction procedures depending on such aspects as problem size, time requirements, user groups, and not least the intended usage of the elicited data.

As described in this thesis, post validation can be a way to balance the trade-off between information quality and elicitation costs. However, as em-

ployed in the case described here (paper V), such post validation demands extensive knowledge of how to conduct sensitivity analysis and identify the inputs necessary to specify further. Such requirements limit its practical usefulness as an expert is needed during such elicitation. Yet, to include post validation within elicitation seems like a promising line to pursue in order to balance elicitation costs with output quality, and therefore, further studies of how to include it within elicitation is an important aspect to address during the development of prescriptive decision analysis methods.

REFERENCES

- Allais, M. (1953). Fondements d'une Théorie positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'École Américaine, translated into The Foundations of a Positive Theory of Choice involving Risk and a Criticism of the Postulates and Axioms of the American School, in *Expected Utility Hypothesis and the Allais' Paradox* (1979). D. Reidel Publishing Company, the Netherlands.
- Anderson, J. 1983. *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Bana e Costa, C.A. and Vansnick, J.C. 1994. MACBETH - an Interactive Path Towards the Construction of Cardinal Value Functions. *International Transactions in Operational Research* **1**(4):489-500.
- Banville, C., Landry, M., Martel, J-M, and Boulaire, C. 1998. A Stakeholder Approach to MCDA. *Systems Research and Behavioral Science* **15**(1):15-32.
- Barclay, S. 1984. *Hiview Software Package*. Decision Analysis Unit, London School of Economics.
- Barron, F.H. 1992. Selecting a Best Multiattribute Alternative with Partial Information About Attribute Weights. *Acta Psychologica* **80**(1-3):91-103.
- Barron, F.H. and Barrett, B.E. 1996a. The Efficacy of SMARTER: Simple Multi-Attribute Rating Technique Extended to Ranking. *Acta Psychologica* **93**(1):23-36.
- Barron, F.H. and Barrett, B.E. 1996b. Decision Quality Using Ranked Attribute Weights. *Management Science* **42**(11):1515-1523.
- Baskerville, R., Pries-Heje, J., and Venable, J. 2009. Soft Design Methodology. In *Proceedings of the 4th International Conference on Design Research in Information Systems and Technology (DESRIST'09)*.
- Beattie, J. and Barron, J. 1991. Investigating the effect of stimulus range on attribute weight. *Journal of Experimental Psychology: Human perception and performance* **17**(2):571-585.

- Bell, D. 1982. Regret in Decision Making under Uncertainty. *Operations Research* **30**(5):961-981.
- Bell, D., Raiffa, H., and Tversky, A. 1988. *Decision Making – Descriptive, Normative and Prescriptive Interactions*. Cambridge University Press, UK.
- Belton, V. and Stewart, T. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, UK.
- Belton, V. and Vickers, S. 1988. *VISA Software Package*. Department of Management Science, University of Strathclyde.
- Bernoulli, D. 1738/1954. Exposition of a New Theory on the Measurement of Risk. *Econometrica* **22**(1):503-546.
- Borcherding, K., Eppel, T., and von Winterfeldt, D. 1991. Comparison of Weighting Judgments in Multiattribute Utility Measurement. *Management Science* **37**(12):1603-1619.
- Bottomley, P.A., Doyle, J.R., and Green, R.H. 2000. Testing the Reliability of Weight Elicitation Methods: Direct Rating versus Point Allocation. *Journal of Marketing Research* **37**(4):508-513.
- Bottomley, P.A. and Doyle, J.R. 2001. A Comparison of Three Weight Elicitation Methods: Good, Better and Best. *Omega* **29**(6):553-560.
- Brans, J.P., Vincke, P.H., and Mareschal, B. 1986. How to Select and How to Rank Projects: the PROMETHEE Method. *European Journal of Operational Research* **24**(2):228-238.
- Brown, R. 1989. Toward a Prescriptive Science and Technology of Decision Aiding. *Annals of Operations Research* **19**:467-483.
- Brown, R. 2006. Making Decision Research Useful - Not Just Rewarding. *Judgment and Decision Making* **1**(2):162-173.
- Brown, R. and Vari, A. 1992. Towards a Research Agenda for Prescriptive Decision Science: The Normative Tempered by the Descriptive. *Acta Psychologica* **80**:33-47.

- Brunsson, N. 1989. *The Organization of Hipocracy: Talk, Decisions and Actions in Organizations*. Copenhagen Business School Press, Denmark.
- Bryman, A. 2004. *Social Research Methods* (2nd ed.). Oxford University Press, Oxford.
- Campbell, D. 1988. *Methodology and Epistemology for Social Sciences: Selected Papers*, Overman, E.S. (ed.). University of Chicago Press.
- Cialdini, R. 2009. *Influence: Science and Practice.*, 5th ed. Pearson Education Inc., US.
- Clemen, R.T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis*, 2nd ed. Duxbury Press, US.
- Clemen, R.T. and Winkler, R. 1999. Combining Probability Distributions From Experts in Risk Analysis. *Risk Analysis* 19(2):187-203.
- Cohen, L.J. 1981. Can Human Irrationality be Experimentally Demonstrated? *The Behavioural and Brain Sciences* 4(3):317-331.
- Corner, J. and Corner, P. 1995. Characteristics of Decisions in Decision Analysis Practice. *Journal of the Operational Research Society* 46(3):304-314.
- Cyert, R. and March, J. 1963. *A Behavioural Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Danielson, M. and Ekenberg, L. 2007. Computing Upper and Lower Bounds in Interval Decision Trees. *European Journal of Operational Research* 181(2):808–816.
- Danielson, M., Ekenberg, L., Ekengren, A., Hökby, T., and Lidén, J. 2008. A Process for Participatory Democracy in Electronic Government. *Journal of Multi-Criteria Decision Analysis* 15:15-30.
- Danielson, M., Ekenberg, L., Johansson, J., and Larsson, A. 2003. The DecideIT decision tool. In *Proceedings of ISIPTA'03*, pp. 204-217, Bernard, J-M., Seidenfeld, T., Zaffalon, M. (eds.). Carleton Scientific: Waterloo, Canada.
- Danielson, M., Ekenberg, L., and Riabacke, A. 2009. A Prescriptive Approach to Elicitation of Decision Data. *Journal of Statistical Theory and Practice* 3(1):157-168.

- Danielson, M., Ekenberg, L., Larsson, A., and Riabacke, M. 2010. Transparent Public Decision Making - Discussion and Case Study in Sweden. *E-Democracy: A Group Decision and Negotiation Perspective*, D. R. Insua and S. French (eds.). Springer.
- de Finetti, B. 1931. Sul significato suggestivo della probabilità. *Fundamenta mathematicae* 17, pp. 298-329. English translation in *Induction and Probability*, Monari, P. and Cocchi, D. (eds.), Clueb, Bologna (1993).
- de Finetti, B. 1937. La prévision: ses lois logiques ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7, pp. 1-68. English translation in *Studies in Subjective Probability*, Kyburg, H.E. and Smokler, H.E. (eds.). Wiley, New York (1964).
- de Finetti, B. 1968. Probability: Interpretations. In *International Encyclopedia of the Social Sciences*, pp. 496-505. Macmillan, New York.
- Edwards, W. 1954. The Theory of Decision Making. *Psychological Bulletin* 51(4):380-417.
- Edwards, W. 1968. Conservatism in Human Information Processing. In *Formal Representation of Human Judgment*, Kleinmuntz, B. (ed.). Wiley, New York.
- Edwards, W. 1977. How to Use Multiattribute Utility Measurement for Social Decisionmaking. *IEEE Transactions on Systems, Man, and Cybernetics* 7(5):326-340.
- Edwards, W. and Barron, F.H. 1994. SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement. *Organizational Behavior and Human Decision Processes* 60(3):306-325.
- Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75(4):643-669.
- Expert choice. 2010. <http://www.expertchoice.com> (2010-02-09).
- Farquhar, P. 1984. Utility Assessment Methods. *Management Science* 30(11):1283-1300.
- Fischer, G.W. 1989. Prescriptive Decision Science: Problems and Opportunities. *Annals of Operations Research* 19: 489-497.

- Fischer, G.W. 1995. Range Sensitivity of Attribute Weights in Multiattribute Value Models. *Organizational behaviour and human decision processes* **62**(3):252-266.
- Fischhoff, B., Goitein, B., and Shapira, Z. 1983. Subjective Expected Utility: A Model of Decision Making. In *Decision Making under Uncertainty*, Scholz, R.W. (ed.). Elsevier Science Publisher B. V., Holland.
- Fischhoff, B., Slovic, P., and Lichtenstein, S. 1978. Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation. *Journal of Experimental Psychology: Human Perception and Performance* **4**(2): 330-344.
- Fishburn, P. 1970. *Utility Theory for Decision Making*. Wiley, New York.
- Fox, J. 1994. On the necessity of probability: Reasons to believe and grounds for doubt. In *Subjective Probability*, Wright, G. and Ayton P. (eds.). John Wiley & Sons Ltd., NY.
- Fox, J. 2003. Probability, Logic and the cognitive foundations of rational belief. *Journal of Applied Logic* **1**(3-4):197-224.
- French, S. 1995a. Uncertainty and Imprecision: Modelling and Analysis. *The Journal of the Operational Research Society* **46**(1):70-79.
- French, S. 1995b. An Introduction to Decision Theory and Prescriptive Decision Analysis. *IMA Journal of Management Mathematics* **6**(2):239-247.
- French, S. and Rios Insua, D. 2000. *Statistical Decision Theory*. Oxford University Press Inc., NY.
- French, S. and Xu, D.-L. 2005. Comparison Study of Multi-attribute Decision Analytic Software. *Journal of Multi-Criteria Decision Analysis* **13**(2-3):65-80.
- Garthwaite, P.H., Kadane, J.B., and O'Hagan, A. 2005. Statistical Methods for Eliciting Probability Distributions. *Journal of the American Statistical Association* **100**(470):680-701.
- Geldermann, J. and Zhang, K. 2001. Software Review: "Decision Lab 2000". *Journal of Multi-Criteria Decision Analysis* **10**(6):317-323.

- Gustafsson, J., Salo, A., and Gustafsson, T. 2001. PRIME Decisions: An Interactive Tool for Value Tree Analysis. In *Multiple Criteria Decision Making in the New Millennium*, M. Köksalan and S. Zionts (eds.), Lecture Notes in Economics and Mathematical Systems 507. Springer-Verlag, Berlin.
- Hart, S. 1992. An Integrative Framework for Strategy-Making Processes. *Academy of Management Review* 17(2):327-351.
- Hevner, A., March, S., Park, J., and Ram, S. 2004. Design Science in Information Systems Research. *MIS Quarterly* 28(1):75-105.
- Hogarth, R.M. 1975. Cognitive Processes and the Assessment of Subjective Probability Distributions. *Journal of the American Statistical Association* 70:271-289.
- Howard, R.A. 1966. Decision Analysis: Applied Decision Theory. In *Proceedings of the Fourth International Conference on Operational Methods*, Herty, D.B. and Melese, J. (eds.). Wiley-Interscience, New York.
- Howard, R.A. 1992. In Praise of the Old Time Religion. In *Utility Theories: Measurement and Applications*, Edwards, W. (ed.). Kluwer Academic Publishers, Boston.
- Hämäläinen, R. and Lauri, H. 1995. HIPRE 3+. Computer Software, Systems Analysis Laboratory, Helsinki University of Technology.
<http://www.sal.hut.fi/Downloadables/> (2010-02-09)
- Jia, J., Fischer, G.W., and Dyer, J. 1998. Attribute Weighting Methods and Decision Quality in the Presence of Response Error: A Simulation Study. *Journal of Behavioral Decision Making* 11(2):85-105.
- Jiménez, A., Ríos-Insua, S., and Mateos, A. 2003. A Decision Support System for Multiattribute Utility Evaluation Based on Imprecise Assignments. *Decision Support Systems* 36(1):65-79.
- Jiménez, A., Ríos-Insua, S., and Mateos, A. 2006. A Generic Multi-Attribute Analysis System. *Computers & Operations Research* 33(4):1081-1101.
- Johnson, E. and Huber, G.P. 1977. The technology of Utility Assessment. *IEE Transactions on Systems, Man and Cybernetics* 7(5):311-325.

- Kahneman, D. and Tversky, A. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* **47**(2):263-292.
- Kahneman, D., Slovic, P., and Tversky, A. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Katsikopoulos, K. and Fasolo, B. 2006. New Tools for Decision Analysts. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans* **36**(5):960-967.
- Keeney, R. 1982. Decision Analysis: An Overview. *Operational Research* **30**(5):803-838.
- Keeney, R. 1992a. *Value-Focused Thinking – A Path to Creative Decisionmaking*. Harvard University Press, Cambridge.
- Keeney, R. 1992b. On the Foundations of Prescriptive Decision Analysis. In *Utility Theories: Measurement and Applications*, Edwards, W. (ed.). Kluwer Academic Publishers, Boston.
- Keeney, R. 2004. Making Better Decision Makers. *Decision Analysis* **1**(4):193-204.
- Keeney, R. and Raiffa, H. 1976. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. John Wiley, New York.
- Keynes, J.M. 1921. *A Treatise on Probability*. MacMillan, New York.
- Kirkwood, C. 1997. *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*. Duxbury Press, CA, USA.
- Kynn, M. 2008. The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society, Series A*, **171**(1):239-264.
- Larichev O. 1992. Cognitive Validity in Design of Decision-Aiding Techniques. *Journal of Multi-Criteria Decision Analysis* **1**(3):127-128.
- Larichev, O., Olson, D., Moshkovich, H., and Mechitov, A. 1995. *Organizational Behavior and Human Decision Processes* **64**(1):9-21.

- Larsson, A. 2008. *Essays on Theories and Applications of Decision Analysis in Imprecise Domains*. Doctoral thesis no. 61, ISSN 1652-893X, Mid Sweden University.
- Lenert, L. and Treadwell, J. 1999. Effects on Preferences of Violations of Procedural Invariance. *Medical Decision Making* **19**(4):473-481.
- Lenert, L., Sherbourne, C., and Reyna, V. 2001. Utility Elicitation Using Single-Item Questions Compared with a Computerized Interview. *Medical Decision Making* **21**(2):97-104.
- Lichtenstein, S. and Slovic, P. (eds.) 2006. *The Construction of Preference*. Cambridge University Press.
- Loomes, G. and Sugden, R. 1982. Regret Theory: An Alternative Theory of Rational Choice under Uncertainty. *Economic Journal* **92**(368):805-824.
- Luce, R.D. and Raiffa, H. 1957. *Games and Decisions: Introduction and Critical Survey*. John Wiley and Sons, New York.
- Machina, M. 1983. Generalized expected utility analysis and the nature of observed violations of the independence axiom. In *Foundations of Utility and Risk Theory with Applications*, Stigum, B. and Wenstøp, F. (eds.). D. Reidel Publishing Company, the Netherlands.
- March, J.G. 1994. *A Primer on Decision Making, How Decisions Happen*. The Free Press, New York.
- March, J.G. 1997. Understanding How Decisions Happen in Organizations. In *Organizational Decision Making*, Shapira, Z. (ed.). Cambridge University Press, USA.
- March, J.G. and Olsen, J. 1976. *Ambiguity and Choice in Organizations*. Universitetsforlaget, Bergen.
- March, J.G. and Simon, H. 1958. *Organizations*. Wiley & Sons, New York.
- March, J.G. and Simon, H. 1993. *Organizations*, 2nd ed. Blackwell, Oxford.

- March, S.T. and Smith, G.F. 1995. Design and Natural Science Research on Information Technology. *Decision Support Systems* **15**:251-266.
- Markowitz, H. 1952. The Utility of Wealth. *Journal of Political Economy* **60**(2):151-158.
- Merkhofer, M.W. 1987. Quantifying Judgmental Uncertainty: Methodology, Experiences, and Insights. *IEEE Transactions on Systems, Man and Cybernetics* **17**(5):741-752.
- Mintzberg, H., Raisinghani, D., and Théorêt, A. 1976. The Structure of "Unstructured" Decision Processes. *Administrative Science Quarterly* **21**(2):246-275.
- Monat, J. 2009. The Benefits of Global Scaling in Multi-Criteria Decision Analysis. *Judgment and Decision Making* **4**(6):492-508.
- Morgan, M.G. and Henrion, M. 1990. *Uncertainty – A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, New York.
- Morton, A. and Fasolo, B. 2009. Behavioural Decision Theory for Multi-Criteria Decision Analysis: A Guided Tour. *Journal of the Operational Research Society* **60**:268-275.
- Morrow, J. and Hitt, M. 2000. Rational and Political Models of Strategic Decision Making. In *Handbook of Strategic Management*, Rabin, J., Miller, G., and Hildreth, W.B. (eds.). CRC Press, US.
- Mustajoki, J., Hämäläinen, R., and Salo, A. 2005. Decision Support by Interval SMART/SWING - Incorporating Imprecision in the SMART and SWING methods. *Decision Sciences* **36**(2):317-339.
- Park, K.S. 2004. Mathematical Programming Models for Characterizing Dominance and Potential Optimality When Multicriteria Alternative Values and Weights are Simultaneously Incomplete. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* **34**(5):601-614.
- Phillips, L.D. 1984. A Theory of Requisite Decision Models. *Acta Psychologica* **56**(1-3):29-48.

- Plous, S. 1993. *The Psychology of Judgment and Decision Making*. McGraw-Hill, US.
- Porter, M. 1980. *Competitive Strategy*. Free Press, NY.
- Påhlman, M. and Riabacke, A. 2005. A Study on Framing Effects in Risk Elicitation. In *Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA-IAWTIC'05)*.
- Pöyhönen, M. 1998. On Attribute Weighting in Value Trees. Doctoral thesis, ISSN 0782-2030, Helsinki University of Technology.
- Pöyhönen, M. and Hämäläinen, R. 2001. On the Convergence of Multiattribute Weighting Methods. *European Journal of Operational Research* **129**(3):569-585.
- Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Random House, New York.
- Raiffa, H. 1994. The Prescriptive Orientation of Decision Making: A Synthesis of Decision Analysis, Behavioural Decision Making, and Game Theory. In *Decision Theory and Decision Analysis: Trends and Challenges*, Ríos, S. (ed.). Kluwer Academic Publishers, US.
- Ramsey, F.P. 1931. Truth and Probability. In *The Foundations of Mathematics and Other Logical Essays*, Braithwaite, R.B. (ed.). Routledge and Keegan Paul, London.
- Riabacke, A. 2007. *Development of Elicitation Methods for Managerial Support*. Doctoral thesis no. 24, ISSN 1652-893X, Mid Sweden University.
- Riabacke, M. 2010. *A Study on Probability Elicitation among Highly Qualified Experts/Decision-makers*. Working Paper no. 10-004. Stockholm University.
- Riabacke, A., Pålman, M., and Larsson, A. 2006a. How Different Choice Strategies Can Affect the Risk Elicitation Process. *IAENG International Journal of Computer Science* **32**(4):460-465.
- Riabacke, A., Pålman, M., and Baidya, T. 2006b. Risk Elicitation in Precise and Imprecise Domains - A Comparative Study, Sweden and Brazil. In *Proceedings of the 2006 International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA-IAWTIC'06)*.

- Riabacke, M., Danielson, M., Ekenberg, L., and Larsson, A. 2009. A Prescriptive Approach for Eliciting Imprecise Weight Statements in an MCDA Process. In *Proceedings of the 1st International Conference on Algorithmic Decision Theory (ADT), Venice, Oct. 21-23, 2009*.
- Riabacke, M., Danielson, M., Ekenberg, L., and Larsson, A. 2012. Cardinal Rank Ordering of Criteria - Addressing Prescription within Weight Elicitation. In last revision stage, *International Journal of IT and Decision Making*.
- Riabacke, M., Åström J., and Grönlund, Å. 2011. E-participation galore? Extending Multi-Criteria Decision Analysis to the Public. *International Journal of Public Information Systems, IJPIS* 7(2):79-99.
- Roy, B. 1996. *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers, Dordrecht.
- Roy, B. and Mousseau, V. 1998. A Theoretical Framework for Analysing the Notion of Relative Importance of Criteria. *Journal of Multi-Criteria Decision Analysis* 5(2):145-159.
- Saaty, T.L. 1980. *The Analytic Hierarchy Process*. McGraw-Hill.
- Salo, A. 1995. Interactive decision aiding for group decision support. *European Journal of Operational Research* 84:134-149.
- Samuelson, P.A. 1977. St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described. *Journal of Economic Literature* 15(1):24-55.
- Savage, L. J. 1954. *The Foundations of Statistics*. Wiley, New York.
- Shapira, Z. 1995. *Risk Taking: A Managerial Perspective*. Russel Sage Foundation, New York.
- Simon, H. 1945. *Administrative Behaviour: A Study of Decision Making Processes in Administrative Organizations*. Free Press, New York.
- Simon, H. 1955. A Behavioural Model of Rational Choice. *Quarterly Journal of Economics* 69(1):99-118.

- Simon, H. 1957. *Models of Man: Social and Rational*. Wiley, New York.
- Simon, H. 1976. *Administrative Behaviour*, 3rd ed. Free Press, New York.
- Simon, H. 1996. *The Sciences of the Artificial*, 3rd ed. MIT Press.
- Simon, H. and associates. 1986. Decision Making and Problem Solving, in *Research Briefings 1986: Report of the Research Briefing Panel on Decision Making and Problem Solving* by the National Academy of Sciences, National Academy Press, Washington, DC. <http://dieoff.org/page163.htm> (2009-08-24).
- Simon J., Wang, Y., and Keller, R. 2011. Paradoxes and Violations of Normative Decision Theory. In *Encyclopedia of Operations Research and Management Science*, Hoboken (NJ), Wiley.
- Slovic, P. 1972. Information Processing, Situation Specificity, and the Generality of Risk-Taking Behaviour. *Journal of Personality and Social Psychology* **22**(1):128-134.
- Slovic, P., Fischhoff, B., and Lichtenstein, S. 1988. Response Mode, Framing, and Information-Processing Effects in Risk Assessment. In *Decision making – Descriptive, Normative, and Prescriptive Interactions*, Bell, D., Raiffa, H., and Tversky, A. (eds.). Cambridge University Press.
- Spetzler, C.S. and Stael von Holstein, C.-A. 1975. Probability Encoding in Decision Analysis. *Management Science* **22**(3):340-358.
- Stewart, T.J. 1992. A critical survey on the status of multiple criteria decision making theory and practice. *OMEGA International Journal of Management Science* **20**(5-6):569-586.
- Stillwell, W.G., Seaver, D.A., and Edwards, W. 1981. A Comparison of Weight Approximation Techniques in Multiattribute Utility Decision Making. *Organizational Behavior and Human Performance* **28**(1):62-77.
- Tversky, A. and Kahneman, D. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* **185**(4157):1124-1131.

- Tversky, A. and Kahneman, D. 1981. The Framing of Decisions and the Psychology of Choice. *Science* **211**(4481):453-458.
- Tversky, A. and Kahneman, D. 1986. Rational Choice and Framing of Decisions. *Journal of Business* **59**(4):252-278.
- Tversky, A., Sattath, S., and Slovic, P. 1988. Contingent Weighting in Judgment and Choice. *Psychological Review* **95**(3)371-384.
- Von Neumann, J. and Morgenstern, O. 1947. *Theory of Games and Economic Behaviour*, 2nd Ed. Princeton University Press.
- Von Nitzsch, R. and Weber, M. 1993. The Effect of Attribute Ranges on Weights in Multiattribute Utility Measurements. *Management Science* **39**(8):937-943.
- Von Winterfeldt, D. and Edwards, W. 1986. *Decision Analysis and Behavioural Research*. Cambridge University Press.
- Wald, A. 1950. *Statistical Decision Functions*. Wiley & Sons, New York.
- Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., and Deb, K. 2008. Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead. *Management Science* **54**(7):1336-1349.
- Walley, P. 1991. *Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
- Wallsten, T.S. and Budescu, D.V. 1983. Encoding Subjective Probabilities: A Psychological and Psychometric Review. *Management Science* **29**(2):151-173.
- Wang, H., Dash, D., and Druzdzal, M. J. 2002. A method for evaluating elicitation schemes for probabilities. *IEEE Transactions on Systems, Man, and Cybernetics* **32**(1):38-43.
- Wehrung, D.A., MacCrimmon, K.R., and Brothers, K.M. 1980. Utility Measures: Comparisons of Domain Stability, and Equivalence Procedures. Working Paper no. 603. Faculty of Commerce and Business Administration, University of British Columbia, Canada.

Williamson, K. 2002. *Research Methods for Students, Academics and Professionals*, 2nd ed. Australia: Center for Information Studies.

Wright, G. and Ayton P. (eds.) 1994. *Subjective Probability*. John Wiley & Sons Ltd., NY.

Yeung, C. and Soman, D. 2005. Attribute Evaluability and the Range Effect. *Journal of Consumer Research* **32**:363-369.

Yin, R. 2003. *Case Study Research*. Sage Publications, US.

Zadeh, L.A. 1965. Fuzzy sets. *Information and Control* **8**(3):338-353.

Zhou, X. 1997. Organizational Decision Making as Rule Following. In *Organizational Decision Making*, Shapira, Z. (ed.). Cambridge University Press, US.