

Review Article

State-of-the-Art Prescriptive Criteria Weight Elicitation

Mona Riabacke, Mats Danielson, and Love Ekenberg

Department of Computer and Systems Sciences, Stockholm University, Forum 100, 164 40 Kista, Sweden

Correspondence should be addressed to Love Ekenberg, lovek@dsv.su.se

Received 24 March 2012; Revised 24 October 2012; Accepted 13 November 2012

Academic Editor: Graham Wood

Copyright © 2012 Mona Riabacke et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Comparatively few of the vast amounts of decision analytical methods suggested have been widely spread in actual practice. Some approaches have nevertheless been more successful in this respect than others. Quantitative decision making has moved from the study of decision theory founded on a single criterion towards decision support for more realistic decision-making situations with multiple, often conflicting, criteria. Furthermore, the identified gap between normative and descriptive theories seems to suggest a shift to more prescriptive approaches. However, 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 the context. In particular, the issue of weight elicitation is crucial. There are several techniques for deriving criteria weights from preference statements. This is a cognitively demanding task, subject to different biases, and the elicited values can be heavily dependent on the method of assessment. There have been a number of methods suggested for assessing criteria weights, but these methods have properties which impact their applicability in practice. This paper provides a survey of state-of-the-art weight elicitation methods in a prescriptive setting.

1. Introduction

The fact that people often have problems making decisions was early noted in a wide range of areas, and decision making has been an issue of concern for quite some time [1–3]. It has become more obvious that cognitive limitations of the human mind make it difficult to process the large amounts of complex information intrinsic in many decision making situations. People seldom talk about possibilities when speaking about decisions to be made, but more commonly use the term decision problem. During the last decades, the field of decision analysis (the applied form of decision theory [4, 5]) has developed as a structured approach to formally analyse decision situations and is based on research within several disciplines such as psychology, mathematics, statistics, and computer science. Much of the work within organizations relates to acts of decision making and problem solving [6], and consequently, there is a

great interest in how decisions are actually made in these settings. Within organization theory (c.f. [7–9]), especially strategic management (c.f. [10, 11]), decision making is central, and the tradition of rationality is considered especially important in systematic approaches to management, such as 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 [12] is to maximize and choose the optimal alternative of all those available to us.

Theoretical developments in decision making have traditionally been divided into normative and descriptive disciplines. Within the normative discipline, the rational model has been prominent and elaborated upon by numerous authors, such as [12–14] to mention some early ones. One feature of these is that they describe how decision-makers should make choices when considering risk. The rational model of decision making is essentially based on the notion that decision-makers systematically gather information in order to objectively analyse it before making a decision [15]. However, even though rationality is a desirable trait, the rational model has been criticized over the years in the behavioural literature concerning its inherent assumptions on cognitive and motivational assumptions [11]. As a consequence, the descriptive discipline (cf., e.g., [16, 17]) 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 [1, 7, 8, 16, 18], where organizational characteristics, such as context, societal structures, organization, conflicting or unclear goals, and political activities (conflict 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 guidelines or tools for applied decision making, and Kirkwood [19] argues that in order to make decisions strategically, it is more or less a requirement to adopt a structured decision making process. Although real decision-makers do not behave like the normative models predict, they might still need and want decision support [20]. Yet, when it comes to decision making processes, structured methods are still seldom applied in real situations, and decision-makers often act on rules of thumb, intuition, or experience instead.

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 complex decision making situations with multiple, often conflicting, criteria. In particular, Multicriteria Decision Analysis (MCDA) has emerged as a promising discipline within decision support methods that can provide decision-makers with a better understanding of the trade-offs involved in a decision, for example, between economic, social, and environmental aspects (criteria). Despite the number of MCDA applications having increased during the last decades, behavioural issues have not received much attention within this field of research. The identification of such problems and the call for research on behavioural issues have been recognized for a while [21]. Moreover, current software applications provide relatively good support for decision analytical calculations but less support for the decision making process itself [22]. French and Xu [22] suggest that this functionality is something that needs to be included in further developments of MCDA methods, and Banville et al. [23] point out 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 sociopolitical context in which these MCDA methods are to be applied is emphasized.

This paper provides an overview of the most common MCDA methods in a prescriptive setting and discusses these from the perspective of reasonably applicable weight elicitation. The next section discusses some fundamental aspects of decision analysis in

general and MCDA methods in particular. It starts with descriptive theories and continues with prescriptive approaches. Section 3 presents state-of-the-art elicitation methods in this setting, beginning with a general discussion and continuing with weights in MCDA methods. This section is then followed by a discussion of the usefulness of prescription in Section 4. Finally, concluding remarks are given in Section 5.

2. Decision Analysis

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 potential to support complex decision making, decision analysis tools are rarely utilized to aid decision making processes [1, 2, 24]. Decision-makers rarely perform decision analyses of complex problems [25]. Some authors, such as Brown [26], claim that the low level of attention given to prescriptive decision support research in real settings has contributed to the limited practical impact that decision analytical aids have had on decision making in business. The explicit use of quantitative decision models to support and improve decision making activities remains modest in real decision situations [24]. Theories and tools still deviate too much from reality requirements and there is reason to believe that without the involvement of actual decision-makers and research on the processes surrounding the developed models, strategies and techniques, the utilization of these tools as aids in real decision making processes will not substantially increase. Another explanation to their limited usage within businesses today is the fact that they are too demanding in terms of required time (especially first time use) and effort. As Keeney [25, page 194] points out, “we all learn decision making by doing it.” 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 to real-life usage, since part of the attraction of using a decision analysis tool to support the decision process is reliant on the applicability of the 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 the required information.

2.1. 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, that is, models describing how people actually make decisions, have been proposed. Within psychology, a dominating viewpoint has been that people make decisions not only based on how they judge the available information, but also influenced by more subconscious factors in the interactive process. One of the early critics of the subjective expected utility model of rational choice was Simon [16], 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 important attributes. Simon also coined the terms substantive and procedural rationality, where the former has to do with the rationality of a decision situation, that is, 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 (which has been more in focus within psychology).

Prospect theory [27] is one of 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 a reference point. It is asymmetric (steeper for losses than for gains) and implies that people are loss averse, that is, the loss of \$1000 has a higher impact than the gain of \$1000. Moreover, it suggests that decision-makers in general are risk averse when it comes to gains and risk seeking when it comes to losses, and systematically overemphasize small probabilities and underemphasize large ones. Prospect theory also expects preferences to depend on the framing of the problem, that is, how the problem is formulated. People are inclined to simplify complex situations, using heuristics and frames when dealing with information [28]. Regret theory [29, 30] 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 March and Olsen [9] (who coined the term garbage can decision making), Slovic et al. [31], and Shapira [2], among others. 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 concept of limited rationality [1].

Over the last decades, numerous models of decision making within organizational settings have been proposed from a number of different theoretical perspectives, and Hart [11, page 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 [32] states that according to rational theory, decision making processes are based on four parts.

- (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 (ibid.) notes that the parts themselves are 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 [16] 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.

2.2. Prescriptive Decision Analysis

If the aim is to act rationally and in a comprehensive way, a systematic approach for information processing and analysis of some kind is required, especially when the problem at hand is complex, nonrepetitive, and involves uncertainty. The identified gap between normative and descriptive theories (cf. [33]) suggests that another approach to a decision making process, such as the one outlined in this section, would be valuable.

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 it is particularly useful for dealing with complex decision making involving risk and uncertainty. Some early results within the prescriptive field were made by Raiffa [4], extended to include multiple objectives by Keeney and Raiffa [5]. Decision analysis is a structured way of modelling decision situations in order to explore and increase understanding of the problem and of possible problematic elements, and to 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 formally on the basis of the provided input data.

The discrepancy between theory and real behaviour is at the very heart of prescriptive interventions [20], and prescriptive decision analysis is conceived as a more pragmatic approach than the normative approach. It has been described as “the application of normative theories, mindful of the descriptive realities, to guide real decision making” [34, page 5]. Prescriptive decision analysis is focused on merging the two classic disciplines (the normative and the descriptive) within decision making into a more practically useful approach for handling decision problems, and aid decision-makers in solving real-life decision problems. The prescriptive approach aims at obtaining components required for analysis in a structured and systematic way, emphasizing human participation and awareness of descriptive realities [35]. Brown and Vari [36] point out that some 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 [28], which can be rectified (or at least reduced) by employing decision aids. Moreover, the employment of an underlying structured model can increase knowledge about the problem at hand and create incentives to acquire as accurate information as possible. 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 [37, page 243] coins 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.” Thus, the prescriptive approach deals with the tailoring of decision analysis processes for specific problems, contexts, and decision-makers. The theoretical and operational choices made provide the means by which the process helps guide the decision-makers through the analyses [33]. The main criteria for evaluating prescriptive models are usefulness (*ibid.*) and pragmatic value [20], and such models should provide decision-makers with suitable assistance in order to improve their decision making.

Keeney [33] further 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). On the other hand, Fischer [38] argues that unless a clearly superior alternative to the expected utility 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 good basis for decision analysis (cf., e.g., [38]), but tackling of the unique and complex aspects of a decision problem may require the use of complementary rules [39] as well as a wider spectrum of risk attitude modelling.

Complex aspects of 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 process designer in trying to aid the decision-maker(s), where the overall objective is to provide a foundation for high quality analyses [33]. These axioms should be practical in the sense that it would be 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 value-focused thinking, advocated by Keeney [40]. He 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 broader about possible courses of action. This is in contrast to the more prevalent alternative-focused thinking where the decision-maker initially finds the available alternatives and thereafter evaluates them. However, Keeney recognizes that the ideal of value-focused thinking is hard to achieve, and many decision problems faced initially arise from a set of alternatives which is to be chosen from [37].

Any decision analysis model is essentially a model of a specific decision situation, a simplification of a reality which includes significant aspects of the problem and lends insights about these aspects [33]. The prescriptive decision analysis process is iterative with iterations through the steps of modelling values; identifying alternatives; evaluating, reflecting, and possibly remodelling of values; modifying or identifying new alternatives; and re-evaluating (see, e.g., [37]). During prescriptive decision analysis, perceptions change and evolve, and the representation of these perceptions should not be static [34]. Requisite modelling is the term used by Phillips [41] to describe this approach to modelling, and a model is requisite when it is sufficient for the decision situation faced. This is in contrast to the static view, often taken in classical decision analysis, where all of the judgments of the decision-maker(s) are taken as fixed and binding from the outset of the analysis [34].

The modelling and selection of the appropriate problem description 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 decision rules in a prescriptive manner. Bell et al. state that "the art and science of elicitation of values (about consequences) and judgments (about uncertainties) lies at the heart of prescriptive endeavours" [20, page 24]. The techniques and methods used for elicitation must be practical and should not require too many inputs from the decision-maker(s). Fischer [38] points at three fundamental problems that need to be confronted when attempting to develop prescriptive models: (1) reference effects (which lead to systematic violations of the independence principle of the expected utility model), such as people's tendency to be risk-averse for gains and risk-seeking for losses as well as weigh losses more heavily than gains [27]; (2) framing problems, that is, that formally equivalent ways of describing (framing) decision problems can highly influence people's choices; and (3) different outcomes resulting from strategically equivalent assessment procedures for eliciting preferences. Prescriptive processes must, thus, be attentive to the descriptive realities of human behaviour and

common mistakes people make when eliciting decision data as the applicability of generated results often relies on the quality of input data. The prescriptive processes must contain procedures for how to elicit adequate judgments from decision-makers and make sense out of them [20]. Moreover, many researchers believe that the insights that can be attained during the elicitation process can be as valuable as what is obtained during processing of the elicited values after elicitation, and it is thus an important ingredient in a prescriptive decision analysis process. The process has a dual role, both to facilitate the work and to keep the decision-maker(s) task oriented as well as to contribute to the task concerned with modelling form rather than content [41].

When decision analysis procedures are employed to aid prescriptive decision making processes, additional demands are put on these procedures to adapt to the users and the context. French and Rios Insua [34] conclude that prescriptive methodologies for decision analysis should aim to be satisfactory with respect to the following aspects.

- (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 practical, 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 the results must be intelligible to the users. The descriptive realities of human behaviour also add demands to elicitation processes to reduce the cognitive load on decision-makers as well as to aim at eliminating biases that have been documented in behavioural research.
- (iii) *Robustness*. The sensitivity to variations in the inputs should be understood. For example, if the analysis results rely heavily on specific 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. The model must be requisite, that is, the application must provide for interactivity and cyclic modelling possibilities in order to reach the goal of compatibility.

2.3. Multicriteria Decision Aids

From having been focused on analyses of a set of alternatives, current research within MCDA is more focused on providing models to support the structuring of problems in order to increase understanding and identify possibly problematic elements. Furthermore, the output from these models should not be interpreted as the solutions to the problems, but rather offer a clearer picture of the potential consequences of selecting a certain course of action. In a more prescriptive context, the decision-maker is assumed to be an agent (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 [21]) who chooses one alternative (or a subset of alternatives) from a set of alternatives (typically consisting of collections of choices of moderate size) (this is in contrast to optimization problems where

feasible sets of alternatives usually consist of infinitely many alternatives) that are being evaluated on the basis of more than one criterion.

Multiattribute Value Theory, MAVT, and Multiattribute Utility Theory, MAUT [5] 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 level of satisfaction by each alternative (according to the decision-maker) under each criterion. Thereafter, the overall score of each alternative is calculated. The main difference between the two is that MAVT is formulated to assume that outcomes of the alternatives are known with certainty, whereas MAUT explicitly takes uncertainty (relating to the outcomes) into account (and thus 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 [35]. Moreover, in many applications, using simple value functions in combination with sensitivity analyses could provide essentially the same results and insights [42]. Basically, MAUT methods contain the following five steps.

- (1) Define the alternatives and the relevant attributes (criteria).
- (2) Evaluate each alternative separately on each attribute, that is, the satisfaction of each alternative under each criterion represented by a value/utility function.
- (3) Assess the relative importance of each criterion, that is, assign relative weights to the attributes.
- (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 MCDA methods other than the classical MAVT/MAUT approach include the Analytic Hierarchy Process, AHP [43], which is similar to MAVT but uses pairwise comparisons of alternatives (utilizing semantic scales) with respect to all criteria, and outranking methods based on partial ordering of alternatives, where the two main approaches are the ELECTRE family of methods (cf., e.g., [44]), and PROMETHEE (cf., e.g., [45]). Moreover, fuzzy set theory (introduced by Zadeh [46]) is an attempt to model human perceptions and preferences, but has some practical elicitation problems, for example, in visualizing an operational elicitation process for the required values [42]. The Measuring Attractiveness by a Categorical Based Evaluation TecHnique, MACBETH [47], 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, for example, V.I.S.A. [48], HiView [49], which supports the MACBETH pairwise comparison approach to elicitation [47], DecideIT [50] and GMAA [51], the latter two allowing the use of interval value and weight statements. The AHP method is implemented in several applications, of which EXPERT CHOICE [52] is among the most widely used. HIPRE 3+ [53] and Logical Decisions are examples of software packages supporting both MAVT and AHP methodologies. Decision Lab 2000 [54] is based on outranking methods such as PROMETHEE [45].

Independent of the approach chosen, one of the most challenging problems is that complete information about the situation to be modelled is unavailable. Most decision analysis situations rely on numerical input of which the decision-maker is inherently unsure,

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 a need to make value trade-offs to indicate the relative desirability of achievement levels on each objective in comparison to the others (represented by criteria weights in MAVT/MAUT methods).

3. Elicitation

While there has been an increase in research (and an intense debate) on elicitation over the last decades within several disciplines such as psychology, statistics, and decision and management science, there are still no generally accepted methods available and the process of eliciting adequate quantitative information from people is still one of the major challenges facing research and applications within the field of decision analysis [55]. Although different research areas have different accounts of elicitation problems, they do agree on the fact that in applied contexts decision-makers and analysts should be concerned not only with what experts are asked to assess, but also how they are asked. 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.

Methods suggested in the literature for elicitation have distinct features which 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 suggested elicitation methods more explicitly, there is a need to categorize them and the following division of the elicitation process into three conceptual components is made in this paper.

- (1) *Extraction*. This component deals with how information (probabilities, utilities, weights) is derived through user input.
- (2) *Representation*. This component deals with how to capture the retrieved information in a structure, that is, the format used to represent the user input.
- (3) *Interpretation*. This component deals with the expressive power of the representation used and how to assign meaning to the captured information in the evaluation of the decision model used.

These categories will in the following be used to analyse elicitation methods in order to discuss their characteristics and identify elements that can impact their practical applicability.

3.1. Probability and Utility Elicitation

In a classical decision analytic framework (cf., e.g., [35]), numerical probabilities are assigned to the different events in tree 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 (parameters) to elicit, which subjects (information sources) to use, and possible training for the subject(s), the most crucial part is to elicit the necessary values from people. Probability information is most commonly elicited from domain experts, and the experts have to express their knowledge and beliefs in probabilistic form during the extraction. This task sometimes involves a facilitator to assist the expert, as many people are

unaccustomed to expressing knowledge in this fashion. Garthwaite et al. [56] conclude that in order for an elicitation process to be successful, the values need not be “true” in an objectivist sense (and cannot be judged that way), but should be an accurate representation of the expert’s present knowledge (regardless of the quality of that knowledge). Moreover, Garthwaite et al. conclude that a reasonable goal for elicitation is 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, for example, Keeney [25] 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 broader discussion concerning objective (classical) and subjective (personal) probabilities, that is, for example, [13, 57, 58]. Subjective probability is thus one of the prime numerical inputs in current extraction procedures, but the meaning of probabilities depends on the perceptual distinction between single-event probabilities and frequencies. This perception can differ among experts, even among those making assessments regarding the same quantities. The elicitation of probabilities has been quite extensively studied, and recommendations as to how to make such assessments and the corresponding problems are studied further in, for example, [59–63].

Methods for utility elicitation have many similarities to probability elicitation processes, but are in a sense 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 decision-maker. Utility can be seen 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 proven to be error prone and cognitively demanding for people. Several techniques for utility elicitation have been proposed and used, and in Johnson and Huber [64] a categorization of these techniques is provided. The category of gamble methods contains the most commonly used techniques, where several variations on question design are being used. A broad discussion on standard gamble methods is found in [65], but capturing utility assessments in terms of hypothetical gambles and lotteries may not successfully map people’s behaviour in all real situations. Some people have a general aversion towards gambling, and people often overweigh 100% certain outcomes in comparison to those that are merely probable (<100%) which complicates matters further [27].

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, that is, for example, [66, 67]. Lichtenstein and Slovic [68] state that people do have well-articulated and preconceived preferences regarding some matters, but in other settings construct their preferences during the process of elicitation, which is one 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 conflicts among the preferences regarding the choices presented. Such circumstances make decision-makers more susceptible to the influence by factors such as framing during the elicitation process, and could explain some of the problems related to extraction.

3.2. Weight Elicitation

In multicriteria decision making, 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. But the concepts of weights and scoring scales are often considered as separate by decision-makers, giving rise to a perceived weight/scale duality. The weight assigned to a criterion is basically a scaling factor which associates scores for that criterion to scores for all other criteria. Methods for eliciting criteria weights are compensatory, that is, the extracted information on the weights' relative importance as assigned by decision-makers implicitly determines trade-offs between the number of units of one criterion they are willing to waive in order to increase the performance of another criterion by one unit.

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 [42, 69, 70]. The task is subject to different biases (cf., e.g., [71]), and the elicited numbers can be heavily dependent on the method of assessment (cf., e.g., [72]). In the literature, there have been a number of methods suggested for assessing criteria weights, and the methods have different features which 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 [73], and therefore elicitation methods with relatively simple extraction components are most common in applied settings. There are several weighting methods that are minor variants of one another, but even small procedural differences have been shown sometimes to have important effects for inference and decision making [74].

In the following sections, some of the most prominent weight elicitation methods are discussed.

3.2.1. 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 given points (from the extraction) 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., [75]), to somewhat more advanced procedures, such as the often used SMART [76], SWING [35], and trade-off [5] methods. As already mentioned, these methods differ in procedures during the extraction. In the DR method, the user is asked to rate each attribute on a scale from 0 to 100, whereas the user in PA is asked to distribute a total of 100 points among the attributes. Bottomley et al. [75] conclude that weights derived from DR are more reliable. The extra cognitive step of having to keep track of the remaining number of points to distribute in the PA method influences the test-retest reliability, that is, how the decision-maker performs on two separate but identical occasions.

In SMART, the user is asked to identify the least important criterion, which receives for example 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 substantially in the interpretation if the method is applied twice. Consequently, this aspect of the extraction in SMART can affect the internal consistency in the interpretational step of the method. In the SWING method, the decision-makers are asked to consider their worst consequence in each criterion and to identify which criterion they would prefer most to change from its worst outcome to its best outcome (the swing). This criterion will be given the highest number of points, for example 100, and is

excluded from the repeated process. The procedure is then repeated with the remaining criteria. The next criterion with the most important swing will be assigned a number relative to the most important one (thus their points denote their relative importance), and so on. Common to all methods described so far is that the number of judgments required by the user during extraction is a minimum of N , where N is the number of attributes.

In trade-off methods, 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. In the first hypothetical alternative the performance of the two criteria are set to their worst and best consequences respectively and in the second alternative the opposite is applied. The decision-maker is asked to choose one of the alternatives, thereby indicating 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 consequence, that is, state the trade-off (s)he is willing to do for certain changes in outcomes 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 more complex and more cognitively demanding in practice due to the large number of pairwise comparisons required. 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., [77]).

The degree of influence of an attribute (criterion) 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 into the SMARTS method [78] to explicitly include spread during extraction. Several empirical studies with methods where ranges are explicitly considered during the extraction of weights reveal that people still do not adjust weight judgments properly when there are changes in the ranges of the attributes (cf., e.g., [79]). In many studies reported in the literature, the range sensitivity principle (e.g., measured by the Range Sensitivity Index, RSI, as suggested in [79]) is violated, often significantly [77, 80]. von Nitzsch and Weber [79] suggest that during decision-makers' judgments on importance, an intuitive idea of an attribute's importance (past experience) functions as an anchor, which is thereafter adjusted by the range of the attribute in the current choice context. Fischer [77] hypothesizes that methods which more explicitly focus on gains or losses in terms of different objectives result in assessed values that are more sensitive to the ranges of the consequences. As an alternative explanation of violations of the range sensitivity principle, Monat [81] claims that the use of local scales may be a problem. As a remedy, global scales that reflect the best and worst values from the decision-maker's view (not the best and worst from the option set) could 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 a global scale. So far, no method has managed to adequately adhere to the range sensitivity principle in empirical studies.

3.2.2. *Imprecise Weight Elicitation*

Accurate determinations of attribute weights by using ratio weight procedures are often hard to obtain in practice since assessed weights are subject to response error [82], and some researchers suggest that attempts to find precise weights may rest on an illusion [70]. Consequently, suggestions on how to use imprecise weights instead have been proposed. In MCDA, there are different approaches to handling more imprecise preference, mainly along one or more of the following sets of approaches [42]: (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 set of approaches. During extraction, decision-makers simply rank the different criteria which are represented by ordinal values. Thereafter, these ordinal values are translated into surrogate (cardinal) weights 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. Several proposals on how to convert such rankings to numerical weights exist, the most prominent being rank sum (RS) weights, rank reciprocal (RR) weights [83], and centroid (ROC) weights [84]. Of the conversion methods suggested, ROC has gained most recognition. Edwards and Barron [78] propose the SMARTER (SMART Exploiting Ranks) method to extract the ordinal information on importance before being interpreted (converted) into numbers using the ROC method.

However, decision data is seldom purely ordinal. There is often some weak form of cardinality present in the information. A decision-maker may be quite confident that some differences in importance are greater than others [82], which is ignored in rank-order (and other ordinal) approaches. Thus, although mere ranking alleviates some of the cognitive demands on users, the conversion from ordinal to cardinal weights produces differences in weights that do not closely reflect what the decision-maker actually means by the ordinal ranking. Ordinal and cardinal information can also be mixed, as in Riabacke et al. [85] where the supplied ranking is complemented with preference relation information without demanding any precision from the decision-maker. In the 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 or values of alternatives during extraction belong to the second set of approaches above, like in the AHP method [43]. However, the correctness of the conversion in the interpretational step, from the semantic scale to the numeric scale used by Saaty [43] as a measure for preference strength, has been questioned by many, for example Belton and Stewart [42]. Moreover, the use of verbal terms in general during elicitation have been criticised, since words can have very different meanings for different people and they often assign different numerical probabilities to the same verbal expressions [19, 86]. 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, preferential uncertainties and incomplete information are handled by using intervals (cf., e.g., [87, 88]), where a range of possible values is represented by an interval. Such methods belong to the third set of approaches, 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 [51]. 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 [89], there are two procedures for assessing weights. Either the extraction is based on trade-offs among the attributes and 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 extraction approach, direct assignment, is more suitable for upper level criteria that could be more

political. The decision-maker directly assigns weight intervals to the respective criteria. In the interpretational step, the extracted interval numbers are automatically computed into an average normalized weight (precise) and a normalized weight interval for each attribute. To explicitly include spread, the SWING approach can be applied to methods where the initial procedural design does not include criteria ranges. In Mustajoki et al. [73], 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, whereafter the decision-maker replies with interval assessments to ratio questions during extraction (to describe possible imprecision in his/her judgments). The extracted weight information is represented by constraints on the attributes' weight ratios, which in addition to the weight normalization constraint determine the feasible region of the weights in the interpretational step.

3.2.3. *Summary of Methods*

The different methods for weight elicitation discussed above are summarised in Tables 1 and 2.

4. **Approaching Elicitation Prescriptively**

Using a single number to represent an uncertain quantity can confuse a decision-maker's judgment about uncertainties with the desirability of various outcomes [19]. Also, subjects often do not initially reveal consistent preference behaviour in many decision situations [5, 90, 91] or protect themselves from exposure by obscuring and managing their preferences [32]. Brunsson [92] 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 100% certain outcome, people often overweigh the certain outcome—the so-called certainty effect [27]. 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 [93]. People are, furthermore, poor intuitive decision-makers in the sense that judgments are affected by the frame in which information is presented as well as by the context. Decision-makers appear to use only the information explicitly presented in the formulation of a problem [94, 95], 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 [66, 96].

The heuristics and biases programme initiated by Tversky and Kahneman [17] illustrates many of the systematic deviations from traditional theoretical expectations inherent in human ways of reasoning, making judgments and in human memory, which cause problems for elicitation processes. Decision-makers have a tendency to be overconfident in their judgments, overestimate desirable outcomes, and seek confirmation of preconceptions. Tversky and Kahneman [17] argue that the processes of human judgment are totally different from what rational models require and identify 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

Table 1: An overview of some of the most prominent weight elicitation methods (N = no. of criteria).

Weight elicitation method	Assessment	Extraction Input	Min. no. of judgments	Representation	Interpretation
Direct Rating	Cardinal Joint procedure	Precise	N	Point estimates	Normalized criteria weights
Point Allocation	Cardinal Joint procedure	Precise	N	Point estimates	Normalized criteria weights
SMART	Cardinal Joint procedure	Precise	N	Point estimates	Normalized criteria weights
SWING	Cardinal Joint procedure	Precise	N	Point estimates	Normalized criteria weights
Trade-off	Cardinal Pairwise procedure	Precise	$N \cdot (N - 1)$ (with consistency check)	Point estimates Relative 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 Relative between pairs of criteria	(Combined) surrogate criteria weights (translated from semantic to exact numerical)
CROC	Ordinal and 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)	$2 \cdot$ (min. no. of judgments in employed ratio-weight procedure)	Intervals	Interval criteria weights

Table 2: A summary of assessment procedures during the extraction step of some of the most prominent weight elicitation methods.

Weight elicitation method	Extraction (assessment procedure)
Direct Rating	Rate each criterion on a 0–100 scale
Point Allocation	Distribute 100 points among the criteria
SMART	(1) Identify the least important criterion, assign 10 points to it (2) Rate the remaining criteria relative to the least important one
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 to it (3) Continue with steps 1 and 2 with the remaining criteria, rate relative to the most important
Trade-off methods	Judge criteria in pairs (1) Make a choice between two alternatives alt.1: the best consequence level of the first criterion and the worst of the second alt.2: the worst consequence level of the first criterion and the best of the second (2) State how much the decision-maker is willing to give up on the most important criterion in order to change the other one to its best level (3) Continue with steps 1 and 2 with the remaining criteria
Rank-order methods	Ordinal statements of criteria importance, that is rank all criteria from the most important to the least important
AHP	Use a systematic pairwise comparison approach in determining preferences (1) Make a choice between two criteria to determine which is the 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 with steps 1 and 2 with the remaining criteria
CROC	(1) Rank all criteria from the most to the least important (2) The most important criterion is given 100 points. The decision-maker is asked to express the importance of the least important criterion in relation to the most important (3) Adjust the distances between the criteria on an analogue visual scale to express the cardinal importance information between the criteria
Interval methods	Generalized ratio weight procedures which employ interval judgments to represent imprecision during extraction instead of point estimates, as in for example, interval SMART/SWING

(biases) such as 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., [97] regarding probabilities, [98] concerning preferences, and [72] regarding weights), and there is no general agreement on the nature of 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 [36].

An additional problem in measuring method preciseness for preference elicitation methods occurs due to the subjective nature of the elicited values. Even though most researchers now agree on the fact that assessed probabilities are subjective in nature, the assessments are intended to represent facts and if experts' assessments disagree, different methods can be used to combine multiple assessments in order to improve the quality of the final estimates. When combining assessments, the main approaches are by mathematical aggregations of individual assessments or by obtaining group consensus [99]. When it comes to preference extraction, 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. There is a great deal of uncertainty involved in elicitation and many reports on the difficulties with extracting precise numbers (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. Thus, there is a need for prescriptive approaches to elicitation.

Elicitation should be an iterative process, where the elicited values may have to be adjusted due to deviations from theoretical expectations or to 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 of coherence are mainly focused on probability theory, compensating for the fact that it often falls short as a model of subjective probability [56]. For example, Tversky and Kahneman have raised the question of whether probability theory should really be thought of as a calculus of human uncertainty in the first place, and Fox [100, page 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 judgement under uncertainty and the development of more sophisticated technologies for aiding such judgement. Prescriptive analyses must include how to elicit judgements from decision-makers and make sense out of them [20]. Prescriptive decision analysis is an attempt to narrow the gap between research within the normative and descriptive disciplines while being rooted in both traditions. It is a more practical approach to handling real-life decision problems, still employing a structured model for analysis. Brown and Vari [36], among others, assert that the behavioural (descriptive) realities are very important in order to design more prescriptive decision aids.

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

- (i) the exact nature of the identified gap between ideal and real behaviour,
- (ii) how to avoid the observed extraction complications, or
- (iii) how to evaluate whether a method has produced accurate input data.

Reaching consensus on these aspects within the decision analysis community is difficult. Pöyhönen [101] suggests focusing research on how methods are used in practice instead of searching for an inclusive theoretical base for all methods. As a guideline, prescriptive research should strive for finding methods that are less cognitively demanding and less sensitive to noisy input within each component. The extraction component is the most error-prone as it concerns the procedural design of the method which could be cognitively demanding during user interaction. Behavioural research has been concentrated on the

extraction component of elicitation, most commonly on 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).

One trend in approaches for extracting the required information in a less precise fashion is methods based on visual aids or verbal expressions. For example, the probability wheel [102] is a popular visual method for eliciting probabilities (the user indicates his/her belief in probability on a circle by sizing a pie wedge to match the assessment on that probability). Such methods often use a combined 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 [19, 86].

Another trend in handling preferential uncertainties and incomplete information in a less precise way is by using intervals as representation (cf., e.g., [51, 88, 103]), where a range of possible values is represented by an interval. Potential benefits with an interval approach include that such representations could facilitate more realistic interpretations of decision-makers' knowledge, beliefs, and preferences, since these elements are not stored with preciseness in human minds. A first analysis of a decision problem can be made using imprecise statements followed by a test whether the input is sufficient for the evaluation of alternatives. If not, the input that needs to be further specified can be identified. Other advantages include that methods based on more approximate preference representations can lead to a more interactive decision support process as the evolution of the decision-maker's priorities can be calculated throughout the process, which in turn could lead to improved decision quality [104]. 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 by compromising their individual judgments if necessary.

For the elicitation of weights, ranking methods using surrogate weights in the interpretational step (e.g., ROC weights [84, 105]) are claimed to be less cognitively demanding and advantageous for group consensus (as groups are more likely to agree on ranks than on precise weights [70]). 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.

Methods for weight elicitation 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 are to use either exact values or mere ranking during extraction. In the CROC method [85], 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 is obtained by calculations. In its representational and interpretational aspects, CROC extends

the ROC weight method [84, 105] 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 is also a way to lessen the decision-makers' reluctance to reveal their true preferences.

The interest of these matters have been around for a while [106], but most important for the practical applicability of MCDA methods is the easiness of employment of the method (see, e.g., [107, 108]). Simpler tools are often easier to use and therefore more likely to be useful. Moreover, elicitation methods that are more direct are easier and less likely to produce elicitation errors [78]. Some even claim that simpler, fast, and frugal methods can produce results that are almost as good as results attributed to those obtained by more extensive analysis, that is, for example, [109]. Larichev et al. [110], among others, suggest that exactness of results should not be the main aim with decision analysis and that different situations call for different levels of exactness depending on the decision-makers' contextual abilities to provide exact judgments. Others have argued for an even more natural way of modelling decision problems and how this relate to prescriptive or normative decision making [111] and a spectrum of new application domains are still emerging [112–114]. Considering the development outlined in Table 1, these trends point to more practically useful elicitation methods (but disregarding the complexities involved with arguing and consensus building [115]) and thus to decision methods that could resolve the apparent gorge between descriptive and normative approaches to decision science. In prescriptive decision science, descriptive and normative approaches are not seen as opposing but rather as two perspectives guiding the design of practically useful and effective decision methods. Especially the two latter extraction methods in Table 2 point to methods with potential to increase the effectiveness. Advances in prescriptive weight elicitation will eventually produce more effective MCDA methods and contribute to a more widespread use of decision methods.

5. Concluding Remarks

This paper has discussed some pragmatic aspects of decision analysis with a focus on weight elicitation models in a prescriptive setting. It discusses these from the perspective of reasonably realistic weight elicitation. As a foundation, the paper discusses some fundamental aspects of decision analysis, specifically MCDA methods. It covers descriptive, normative, and prescriptive theories. The focus is on state-of-the-art elicitation methods, in particular weights in MCDA methods. Elicitation is important to prescriptive approaches and to decision making processes in general. From a behavioural perspective, the need for decision support systems (based on decision analysis) that are easier for decision-makers 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.

As has been discussed above, there are several issues involved. When designing elicitation methods, there is a need to understand psychological traps within extraction, such as framing and heuristics that produce biased assessments in order to apply measures to lessen their effect in method design. Using a clear terminology is important, such as explaining the meaning of specific terms in the 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), and so forth. In order to reduce the gap between theoretical research and practical needs, there are aspects of the extraction component that need to be considered. Behavioural aspects, like the heuristics and corresponding biases people use during extraction, are important to be observant of

in order to reduce such effects. Increased awareness of how presentation formats affect decision-makers' choices is called for to reduce the well-known framing problems. Such problems are often considered a hindrance to sufficient extraction, for example, being aware of the decision-makers' aversion to losses, tendency to overweigh 100% certain outcomes, and so forth.

Moreover, relaxation of the precise statements that are commonly required in the extraction and representation components of elicitation methods could be advantageous. There is a contradiction between the ambiguity of human judgement and the exactness (of elicited values) required by most current decision analysis models. People have problems judging exact values, which poses a problem when the required values are point estimates, and some of the deviations from the traditional decision theoretical expectations could be attributed to this inability.

One must also keep in mind that practical techniques for elicitation are to a great extent a matter of balancing the obtained quality of elicitation with the time available and cognitive effort demand on the users for extracting all the required information. Sensitivity analyses could be used to monitor 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. Decision methods must be used with caution regarding when they should be meaningfully utilised. Advances in prescriptive weight elicitation are one key component to design more effective MCDA methods and contribute to a more widespread use of multicriteria decision methods.

References

- [1] J. G. March, *A Primer on Decision Making, How Decisions Happen*, The Free Press, New York, NY, USA, 1994.
- [2] Z. Shapira, *Risk Taking: A Managerial Perspective*, Russel Sage Foundation, New York, NY, USA, 1995.
- [3] H. Simon, *Administrative Behaviour*, Free Press, New York, NY, USA, 3rd edition, 1976.
- [4] H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*, Random House, New York, NY, USA, 1968.
- [5] R. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*, John Wiley & Sons, New York, NY, USA, 1976.
- [6] H. Simon, "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, USA, 1986.
- [7] J. G. March and H. Simon, *Organizations*, John Wiley & Sons, New York, NY, USA, 1958.
- [8] R. Cyert and J. March, *A Behavioural Theory of the Firm*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1963.
- [9] J. G. March and J. Olsen, *Ambiguity and Choice in Organizations*, Universitetsforlaget, Bergen, Norway, 1976.
- [10] M. Porter, *Competitive Strategy*, Free Press, New York, NY, USA, 1980.
- [11] S. Hart, "An integrative framework for strategy-making processes," *Academy of Management Review*, vol. 17, no. 2, pp. 327–351, 1992.
- [12] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton University Press, 2nd edition, 1947.
- [13] L. J. Savage, *The Foundations of Statistics*, John Wiley & Sons, New York, NY, USA, 1954.
- [14] R. D. Luce and H. Raiffa, *Games and Decisions: Introduction and Critical Survey*, John Wiley & Sons, New York, NY, USA, 1957.
- [15] J. Morrow and M. Hitt, "Rational and political models of strategic decision making," in *Handbook of Strategic Management*, J. Rabin, G. Miller, and W. B. Hildreth, Eds., CRC Press, Boca Raton, Fla, USA, 2000.
- [16] H. Simon, *Models of Man: Social and Rational*, John Wiley & Sons, New York, NY, USA, 1957.

- [17] A. Tversky and D. Kahneman, "Judgment under uncertainty: heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty," *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974.
- [18] H. Simon, *Administrative Behaviour: A Study of Decision Making Processes in Administrative Organizations*, Free Press, New York, NY, USA, 1945.
- [19] C. Kirkwood, *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*, Duxbury Press, Pacific Grove, Calif, USA, 1997.
- [20] D. Bell, H. Raiffa, and A. Tversky, *Decision Making—Descriptive, Normative and Prescriptive Interactions*, Cambridge University Press, Cambridge, UK, 1988.
- [21] J. Wallenius, J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, and K. Deb, "Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead," *Management Science*, vol. 54, no. 7, pp. 1336–1349, 2008.
- [22] S. French and D. L. Xu, "Comparison study of multi-attribute decision analytic software," *Journal of Multi-Criteria Decision Analysis*, vol. 13, no. 2-3, pp. 65–80, 2005.
- [23] C. Banville, M. Landry, J. M. Martel, and C. Boulaire, "A stakeholder approach to MCDA," *Systems Research and Behavioral Science*, vol. 15, no. 1, pp. 15–32, 1998.
- [24] R. Brown, "Making decision research useful—not just rewarding," *Judgment and Decision Making*, vol. 1, no. 2, pp. 162–173, 2006.
- [25] R. Keeney, "Making better decision makers," *Decision Analysis*, vol. 1, no. 4, pp. 193–204, 2004.
- [26] R. V. Brown, "Toward a prescriptive science and technology of decision aiding," *Annals of Operations Research*, vol. 19, no. 1, pp. 465–483, 1989.
- [27] D. Kahneman and A. Tversky, "Prospect Theory: an analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–292, 1979.
- [28] D. Kahneman, P. Slovic, and A. Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, 1982.
- [29] G. Loomes and R. Sugden, "Regret Theory: an Alternative Theory of Rational Choice under Uncertainty," *Economic Journal*, vol. 92, no. 368, pp. 805–824, 1982.
- [30] D. Bell, "Regret in Decision Making under Uncertainty," *Operations Research*, no. 5, pp. 961–981, 1982.
- [31] P. Slovic, B. Fischhoff, and S. Lichtenstein, "Response mode, framing, and information-processing effects in risk assessment," in *Decision Making—Descriptive, Normative, and Prescriptive Interactions*, D. Bell, H. Raiffa, and A. Tversky, Eds., Cambridge University Press, 1988.
- [32] J. G. March, "Understanding how decisions happen in organizations," in *Organizational Decision Making*, Z. Shapira, Ed., Cambridge University Press, Cambridge, Mass, USA, 1997.
- [33] R. Keeney, "On the foundations of prescriptive decision analysis," in *Utility Theories: Measurement and Applications*, W. Edwards, Ed., Kluwer Academic Publishers, Boston, Mass, USA, 1992.
- [34] S. French and D. Rios Insua, *Statistical Decision Theory*, Oxford University Press, New York, NY, USA, 2000.
- [35] D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioural Research*, Cambridge University Press, 1986.
- [36] R. Brown and A. Vari, "Towards a research agenda for prescriptive decision science: the normative tempered by the descriptive," *Acta Psychologica*, vol. 80, no. 1–3, pp. 33–47, 1992.
- [37] S. French, "An introduction to decision theory and prescriptive decision analysis," *IMA Journal of Management Mathematics*, vol. 6, no. 2, pp. 239–247, 1995.
- [38] G. W. Fischer, "Prescriptive decision science: problems and opportunities," *Annals of Operations Research*, vol. 19, no. 1, pp. 489–497, 1989.
- [39] M. Danielson, "Generalized evaluation in decision analysis," *European Journal of Operational Research*, vol. 162, no. 2, pp. 442–449, 2005.
- [40] R. Keeney, *Value-Focused Thinking—A Path to Creative Decisionmaking*, Harvard University Press, Cambridge, UK, 1992.
- [41] L. D. Phillips, "A theory of requisite decision models," *Acta Psychologica*, vol. 56, no. 1–3, pp. 29–48, 1984.
- [42] V. Belton and T. Stewart, *Multiple Criteria Decision Analysis: An Integrated Approach*, Kluwer Academic Publishers, Lancaster, UK, 2002.
- [43] T. L. Saaty, *The Analytic Hierarchy Process*, McGraw-Hill, 1980.
- [44] B. Roy, *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [45] J. P. Brans, P. Vincke, and B. Mareschal, "How to select and how to rank projects: the Promethee method," *European Journal of Operational Research*, vol. 24, no. 2, pp. 228–238, 1986.

- [46] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [47] C. A. Bana e Costa and J. C. Vansnick, "MACBETH—an interactive path towards the construction of cardinal value functions," *International Transactions in Operational Research*, vol. 1, no. 4, pp. 489–500, 1994.
- [48] V. Belton and S. Vickers, *VISA Software Package*, Department of Management Science, University of Strathclyde, 1988.
- [49] S. Barclay, *Hiview Software Package*, Decision Analysis Unit, London School of Economics, 1984.
- [50] M. Danielson, L. Ekenberg, J. Johansson, and A. Larsson, "The DecideIT decision tool," in *Proceedings of the 3rd International Symposium on Imprecise Probabilities and Their Applications (ISIPTA '03)*, Bernard, J.-M., T. Seidenfeld, and M. Zaffalon, Eds., pp. 204–217, Carleton Scientific, Lugano, Switzerland, 2003.
- [51] A. Jiménez, S. Ríos-Insua, and A. Mateos, "A generic multi-attribute analysis system," *Computers and Operations Research*, vol. 33, no. 4, pp. 1081–1101, 2006.
- [52] "Expert choice," 2010, <http://www.expertchoice.com>.
- [53] R. Hämmäläinen and H. Lauri, *HIPRE 3+ Computer Software*, Systems Analysis Laboratory, Helsinki University of Technology, 1995.
- [54] J. Geldermann and K. Zhang, "Software review: Decision Lab 2000," *Journal of Multi-Criteria Decision Analysis*, vol. 10, no. 6, pp. 317–323, 2001.
- [55] J. Fox, "Probability, logic and the cognitive foundations of rational belief," *Journal of Applied Logic*, vol. 1, no. 3–4, pp. 197–224, 2003.
- [56] P. H. Garthwaite, J. B. Kadane, and A. O'Hagan, "Statistical methods for eliciting probability distributions," *Journal of the American Statistical Association*, vol. 100, no. 470, pp. 680–700, 2005.
- [57] B. de Finetti, "Probability: interpretations," in *International Encyclopedia of the Social Sciences*, pp. 496–505, Macmillan, New York, NY, USA, 1968.
- [58] G. Wright and P. Ayton, Eds., *Subjective Probability*, John Wiley & Sons, New York, NY, USA, 1994.
- [59] R. T. Clemen, *Making Hard Decisions: An Introduction to Decision Analysis*, Duxbury Press, Pacific Grove, Calif, USA, 2nd edition, 1996.
- [60] J. L. Corner and P. D. Corner, "Characteristics of decisions in decision analysis practice," *Journal of the Operational Research Society*, vol. 46, no. 3, pp. 304–314, 1995.
- [61] R. M. Hogarth, "Cognitive processes and the assessment of subjective probability distributions," *Journal of the American Statistical Association*, vol. 70, pp. 271–289, 1975.
- [62] M. G. Morgan and M. Henrion, *Uncertainty—A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, New York, NY, USA, 1990.
- [63] T. S. Wallsten and D. V. Budescu, "Encoding subjective probabilities: a psychological and psychometric review," *Management Science*, vol. 29, no. 2, pp. 151–173, 1983.
- [64] E. M. Johnson and G. P. Huber, "The technology of Utility Assessment," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 7, no. 5, pp. 311–325, 1977.
- [65] P. H. Farquhar, "Utility assessment methods," *Management Science*, vol. 30, no. 11, pp. 1283–1300, 1984.
- [66] A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.
- [67] L. A. Lenert and J. R. Treadwell, "Effects on preferences of violations of procedural invariance," *Medical Decision Making*, vol. 19, no. 4, pp. 473–481, 1999.
- [68] S. Lichtenstein and P. Slovic, Eds., *The Construction of Preference*, Cambridge University Press, 2006.
- [69] O. Larichev, "Cognitive validity in design of decision-aiding techniques," *Journal of Multi-Criteria Decision Analysis*, vol. 1, no. 3, pp. 127–128, 1992.
- [70] F. H. Barron and B. E. Barrett, "The efficacy of SMARTER—simple multi-attribute rating technique extended to ranking," *Acta Psychologica*, vol. 93, no. 1–3, pp. 23–26, 1996.
- [71] K. Borchering, T. Eppel, and D. von Winterfeldt, "Comparison of weighting judgments in multiattribute utility measurement," *Management Science*, vol. 37, no. 12, pp. 1603–1619, 1991.
- [72] M. Pöyhönen and R. P. Hämmäläinen, "On the convergence of multiattribute weighting methods," *European Journal of Operational Research*, vol. 129, no. 3, pp. 569–585, 2001.
- [73] J. Mustajoki, R. P. Hämmäläinen, and A. Salo, "Decision support by interval SMART/SWING—incorporating imprecision in the SMART and SWING methods," *Decision Sciences*, vol. 36, no. 2, pp. 317–339, 2005.
- [74] P. A. Bottomley and J. R. Doyle, "A comparison of three weight elicitation methods: good, better, and best," *Omega*, vol. 29, no. 6, pp. 553–560, 2001.
- [75] P. A. Bottomley, J. R. Doyle, and R. H. Green, "Testing the reliability of weight elicitation methods: direct rating versus point allocation," *Journal of Marketing Research*, vol. 37, no. 4, pp. 508–513, 2000.

- [76] W. Edwards, "How to use multiattribute utility measurement for social decisionmaking," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 7, no. 5, pp. 326–340, 1977.
- [77] G. W. Fischer, "Range sensitivity of attribute weights in multiattribute value models," *Organizational Behavior and Human Decision Processes*, vol. 62, no. 3, pp. 252–266, 1995.
- [78] W. Edwards and F. H. Barron, "SMARTS and SMARTER: improved simple methods for multiattribute utility measurement," *Organizational Behavior and Human Decision Processes*, vol. 60, no. 3, pp. 306–325, 1994.
- [79] R. von Nitzsch and M. Weber, "The effect of attribute ranges on weights in multiattribute utility measurements," *Management Science*, vol. 39, no. 8, pp. 937–943, 1993.
- [80] J. Beattie and J. Baron, "Investigating the effect of stimulus range on attribute weight," *Journal of Experimental Psychology*, vol. 17, no. 2, pp. 571–585, 1991.
- [81] J. P. Monat, "The benefits of global scaling in multi-criteria decision analysis," *Judgment and Decision Making*, vol. 4, no. 6, pp. 492–508, 2009.
- [82] J. Jia, G. W. Fischer, and J. S. Dyer, "Attribute weighting methods and decision quality in the presence of response error: a simulation study," *Journal of Behavioral Decision Making*, vol. 11, no. 2, pp. 85–105, 1998.
- [83] W. G. Stillwell, D. A. Seaver, and W. Edwards, "A comparison of weight approximation techniques in multiattribute utility decision making," *Organizational Behavior and Human Performance*, vol. 28, no. 1, pp. 62–77, 1981.
- [84] F. H. Barron, "Selecting a best multiattribute alternative with partial information about attribute weights," *Acta Psychologica*, vol. 80, no. 1–3, pp. 91–103, 1992.
- [85] M. Riabacke, M. Danielson, L. Ekenberg, and A. Larsson, "A prescriptive approach for eliciting imprecise weight statements in an MCDA process," in *Proceedings of the 1st International Conference on Algorithmic Decision Theory (ADT '09)*, Venice, Italy, October 2009.
- [86] M. W. Merkhofer, "Quantifying judgmental uncertainty: methodology, experiences, and insights," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 17, no. 5, pp. 741–752, 1987.
- [87] P. Walley, *Reasoning with Imprecise Probabilities*, Chapman and Hall, London, UK, 1991.
- [88] M. Danielson and L. Ekenberg, "Computing upper and lower bounds in interval decision trees," *European Journal of Operational Research*, vol. 181, no. 2, pp. 808–816, 2007.
- [89] A. Jiménez, S. Ríos-Insua, and A. Mateos, "A decision support system for multiattribute utility evaluation based on imprecise assignments," *Decision Support Systems*, vol. 36, no. 1, pp. 65–79, 2003.
- [90] R. L. Keeney, "Decision analysis: an overview," *Operational Research*, vol. 30, no. 5, pp. 803–838, 1982.
- [91] D. A. Wehrung, K. R. MacCrimmon, and K. M. Brothers, "Utility measures: comparisons of domain stability, and equivalence procedures," Tech. Rep. 603, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, Canada, 1980.
- [92] N. Brunsson, *The Organization of Hipocracy: Talk, Decisions and Actions in Organizations*, Copenhagen Business School Press, Copenhagen, Denmark, 1989.
- [93] A. Tversky, S. Sattath, and P. Slovic, "Contingent Weighting in Judgment and Choice," *Psychological Review*, vol. 95, no. 3, pp. 371–384, 1988.
- [94] P. Slovic, "Information processing, situation specificity, and the generality of risk-taking behavior," *Journal of Personality and Social Psychology*, vol. 22, no. 1, pp. 128–134, 1972.
- [95] B. Fischhoff, P. Slovic, and S. Lichtenstein, "Fault trees: Sensitivity of estimated failure probabilities to problem representation," *Journal of Experimental Psychology*, vol. 4, no. 2, pp. 330–344, 1978.
- [96] A. Tversky and D. Kahneman, "Rational choice and framing of decisions," *Journal of Business*, vol. 59, no. 4, pp. 252–278, 1986.
- [97] H. Wang, D. Dash, and M. J. Druzdzel, "A method for evaluating elicitation schemes for probabilistic models," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 32, no. 1, pp. 38–43, 2002.
- [98] L. A. Lenert, C. D. Sherbourne, and V. Reyna, "Utility elicitation using single-item questions compared with a computerized interview," *Medical Decision Making*, vol. 21, no. 2, pp. 97–104, 2001.
- [99] R. T. Clemen and R. L. Winkler, "Combining probability distributions from experts in risk analysis," *Risk Analysis*, vol. 19, no. 2, pp. 187–203, 1999.
- [100] J. Fox, "On the necessity of probability: reasons to believe and grounds for doubt," in *Subjective Probability*, G. Wright and P. Ayton, Eds., John Wiley & Sons, New York, NY, USA, 1994.
- [101] M. Pöyhönen, *On attribute weighting in value trees [doctoral thesis]*, Helsinki University of Technology, 1998.
- [102] C. S. Spetzler and C. A. S. Stael von Holstein, "Probability encoding in decision analysis," *Management Science*, vol. 22, no. 3, pp. 340–358, 1975.

- [103] K. S. Park, "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 A*, vol. 34, no. 5, pp. 601–614, 2004.
- [104] A. A. Salo, "Interactive decision aiding for group decision support," *European Journal of Operational Research*, vol. 84, no. 1, pp. 134–149, 1995.
- [105] F. H. Barron and B. E. Barrett, "Decision quality using ranked attribute weights," *Management Science*, vol. 42, no. 11, pp. 1515–1523, 1996.
- [106] K. Katsikopoulos and C. H. Lan, "Herbert Simon's spell on judgment and decision making," *Judgment and Decision Making*, vol. 6, no. 8, pp. 722–732, 2011.
- [107] T. J. Stewart, "A critical survey on the status of Multiple Criteria Decision Making theory and practice," *Omega*, vol. 20, no. 5-6, pp. 569–586, 1992.
- [108] J. Hülle, R. Kaspar, and K. Möller, "Multiple criteria decision-making in management accounting and control-state of the art and research perspectives based on a Bibliometric Study," *Journal of Multi-Criteria Decision Analysis*, vol. 18, no. 5-6, pp. 253–265, 2011.
- [109] K. Katsikopoulos and B. Fasolo, "New tools for decision analysts," *IEEE Transactions on Systems, Man and Cybernetics A*, vol. 36, no. 5, pp. 960–967, 2006.
- [110] O. I. Larichev, D. L. Olson, H. M. Moshkovich, and A. J. Mechitov, "Numerical vs cardinal measurements in multiattribute decision making: how exact is enough?" *Organizational Behavior and Human Decision Processes*, vol. 64, no. 1, pp. 9–21, 1995.
- [111] C. Nemeth and G. Klein, "The naturalistic decision making perspective," *Wiley Encyclopedia of Operations Research and Management Science*. In press.
- [112] A. Posch, A. Agarwal, and P. Strachan, "Managing industrial symbiosis (IS) networks," *Business Strategy and the Environment*, vol. 20, no. 7, pp. 421–427, 2011.
- [113] H. Kunz and T. Schaaf, "General and specific formalization approach for a Balanced Scorecard: an expert system with application in health care," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1947–1955, 2011.
- [114] R. Brown, *Environmental Regulation: Developments in How to Set Requirements and Verify Compliance*, George Mason University School of Public Policy, Research Paper no. 2012-01, 2012.
- [115] Weide and T. L. van der, *Arguing to Motivate Decisions*, SIKS Dissertation Series, 2011.

