Biases and Debiasing in Multi-Criteria Decision Analysis

Gilberto Montibeller  
Department of Management  
London School of Economics and Political Science  
Houghton Street, London WC2A 2AE

Detlof von Winterfeldt  
Department of Industrial and Systems Engineering  
University of Southern California  
University Park, Los Angeles, CA 90089

Abstract

Developing models and estimating parameters for multi-criteria decision analysis requires judgments by experts and decision makers. These judgments are subject to biases, which can reduce the quality of the analysis. Some of these biases are due to faulty cognitive processes; some are due to motivations for preferred analysis outcomes. We describe these biases, how they affect multi-criteria decision analysis and discuss some debiasing techniques.

1. Introduction

The methodologies for improving decision making with multiple objectives have evolved over the past decades into a set of mature and widely applied model and techniques, which we refer to jointly as multicriteria decision analysis (MCDA, for a summary see Belton and Stewart [7]). These include multiattribute utility functions [44], measurable value functions [15], multi-criteria programming methods [17], outranking methods [79], and the Analytic Hierarchy Process [80].

All MCDA methodologies start by developing a list of alternatives and objectives. The alternatives are options to be evaluated like cars, houses, job applicants, technologies, business strategies, or R&D portfolios. The objectives define what the decision maker wants to achieve. In the context of buying a home, the objectives would likely include, size, location, view, proximity to good schools and shopping, etc. Objectives are operationalized by measurable criteria (also called “performance measures” or “attributes”). For example, the objective “size” of a home can be operationalized by square footage of living space. All MCDA approaches involve some system for scoring the performance of alternatives on the criteria and most of them assign weights to the criteria. Some methods require rescaling performance measures as value or utility functions. Given performance assessments, values or utility functions and weights, alternatives are usually evaluated by calculating the weighted average of the scores on the performance criteria.

While we focus here on approaches based on multi-attribute utility and value function, many of the issues discussed in the paper also apply for alternative methodologies.

All MCDA methodologies require judgmental inputs from decision makers and stakeholders, regarding alternatives, and objectives, attributes, value functions, and weights. It is by now well known that obtaining these judgments can be subject to biases, which in turn can lead to poor analysis results.

Despite the relevance of biases for MCDA, there are few articles that cover the topic from an analytic perspective: von Winterfeldt [91] identified several cognitive biases when discussing the implications of behavioral research for decision analysis; Weber and Borcherding [87] examined biases in multi-attribute weight assessment; Morton & Fasolo [67] reviewed the implications of biases for multi-criteria decision analysis modeling; and Fasolo et al. [21] for resource allocation models. Larrick [54] discusses motivational issues related to decision making performance.

2. Biases and Debiasing

We distinguish between cognitive and motivational biases. A cognitive bias is a systematic discrepancy between the “correct” answer in a judgmental task, given by a formal normative rule, and the decision maker’s or expert’s actual answer to such a task [89]. There is a vast literature on cognitive biases and excellent compilations of papers are provided in Kahneman et al. [40] and Gilovich et al. [30]. Many of these biases concern judgments about facts, events, and uncertainties. In contrast this paper focuses on biases in value assessment.

We define motivational biases as those in which judgments are influenced by the desirability or undesirability of events, consequences, outcomes, or choices (see also Kunda [53], von Winterfeldt [91] and Molden & Higgins [63]). An example of a motivational bias is the deliberate attempt to assign higher weights to an attribute which favors a preferred alternative.

Debiasing refers to attempts to eliminate, or at least reduce, cognitive or motivational biases. The small literature on debiasing has focused on cognitive biases; early attempts showed the limited efficiency of debiasing tools [3,24,41], but more recent papers
are slightly more optimistic about overcoming biases, particularly with the use of adequate tools [4,54,61]. In addition to the behavioral literature on debiasing, we also draw on applied experiences of practitioners of MCDA. Some of these have been subjected to experimental tests (e.g., Seaver et al. [82] and Abbas et al. [1] show how the overconfidence bias can be reduced by choice of an appropriate elicitation technique), others have been described in applied research papers (e.g., Dillon et al. [14] describe attempts to reduce the anchoring and overconfidence biases of engineering cost estimators).

Most MCDA methodologies involve five judgment tasks and one aggregation task (see also Keeney and Raiffa [44] and Keeney [46]):
1. Generation of alternatives and objectives
2. Development of attributes for the objectives
3. Assessment of the performance of the alternatives on the attributes
4. Elicitation of utility or value functions over attribute levels
5. Elicitation of weights for attributes
6. Aggregation

When considering biases in MCDA, we are primarily concerned with steps 1, 2, 4, and 5, which involve judgments by decision makers and stakeholders. Performance assessment (task 3) usually consists of data collection and sometimes involves expert judgments, including uncertainty assessments. There is a vast literature on biases in this task that we do not cover in this paper – examples are the overconfidence bias and the anchoring bias (see Kahneman et al. [40] and Gilovich et al. [30]). Aggregation is a mechanical task, which does not involve human judgments and therefore is not subject to biases.

3. Biases When Generating Alternatives and Objectives

The process of generating alternatives and objectives often involves going back and forth between objectives and alternatives with the intent to develop a complete set that captures the main features of the decision problem under investigation. Identifying potentially good alternatives is crucial in decision making [46], but individuals and organizations often consider only one alternative [18,69] defining the problem as a binary choice between this alternative and the status quo. Limiting the set of alternatives based on narrowly set goals and constraints is another common mistake [47]. Objectives are either directly elicited in interviews with the decision makers, or constructed from multiple interviews with stakeholders [8,10].

**Biases.** Perhaps the most prevalent bias when generating alternatives and objectives is the omission of important items in this process. This omission bias can lead to the failure to include important alternatives that may turn out to be contenders or winners in an evaluation [36]. Evidence for the omission of good alternatives is also presented in Pitz et al. [74] and Jungermann et al. [39]. Similarly, several studies have shown that people fail to generate a complete set of objectives in many decision problems [9,10].

This omission bias may lead to poor recommendations [71], as some important consequences are completely disregarded in the analysis. The simulation performed by Fry et al. [28] assessed the impact of omissions of objectives and shows that it tends to increase with the rise of both the number of objectives and the number of missing objectives. Identifying and structuring objectives rely heavily on decision makers’ mental models [38]. Research shows that the myopic problem representation bias tends to generate an incomplete problem descriptions due to over-simplified mental models [55,56,71].

Other biases that play a role in the generation of alternatives are anchoring, when alternatives are anchored on an initial set [46] and availability, when the existence of one alternative prevents the generation of other ones [62]. In addition, the desirability bias may lead to the exclusion of alternatives that compete with the preferred one. Bond et al. [9,10] shows that subjects find it difficult to generate a comprehensive set of objectives.

**Debiasing.** Empirical research has shown that presenting one objective at a time and asking respondents to generate alternatives that meet this objective generates more alternatives than when no objectives or all objectives together are presented [39,74]. More recently, Butler & Scherer [12] have shown that presenting objectives leads not only to more, but also to better alternatives. Keeney [46], Gregory & Keeney [31], Keeney [48], and Keller & Ho [49] suggest an extensive list of strategies to help generating decision alternatives. These can be classified into the following categories [49]: objective-based strategies (e.g. presenting one objective at a time and asking for high value achieving alternatives, designing options that perform well on high-weighted objectives, etc.), state-based strategies (e.g. presenting possible states one at a time and asking for high value achieving alternatives in that future state, etc.), and alternative-based (e.g. imagining an ideal option and designing alternatives from it, using existing options to generate new ones, etc.). Farquhar & Pratkanis [19] also mention the use of phantom alternatives as a way of stimulating
creativity, e.g. the inclusion of an unfeasible “ideal” alternative helping decision-makers to create new options, as described by Phillips [72]. Tools such as cognitive maps [5,65] and strategy-generation tables [34] can also be used to develop alternatives.

Bond et al. [10] found that the use of generic categories and, even more effective, the challenge to increase the number of objectives, increased the number of objectives generated by the subjects. Leon [57] discovered that value-focused thinking helped in eliciting not only more objectives, but also objectives that were perceived to have better features as evaluation criteria. Keeney [46] suggests several probes to help decision makers in generating objectives, including writing a wish list, thinking about features of good (and bad) alternatives, imagining consequences of actions, considering goals and constraints, and adopting other stakeholders’ perspectives that help in reducing the myopic problem representation bias. Other tools to identify objectives are the use of cognitive maps [16], networks of ideas with a means-end structure [5,64], or affinity diagrams, where objectives are elicited and clustered [70].

To obtain a comprehensive set of objectives, practitioners often interview multiple stakeholders [90]. Creating a comprehensive list of objectives from multiple inputs is usually uncontroversial, because the decision maker(s) can always zero out selected objectives in the weighting process (see below). Another way of obtaining multiple perspectives is to elicit the objectives in groups, using decision conferencing supported by a facilitator [27,73].

4. Biases When Defining Attributes

The choice of an attribute (criterion, performance measure) is a very important task in MCDA, because it prescribes what data on performance should be collected, how it should be collected, whether expert judgment is needed, and to what degree uncertainty should be incorporated in these judgments. While the process of defining attributes is not biased per se, a poorly defined attribute can lead to biases in the subsequent judgments of performance, value or utility functions, and weights.

Biases. Research on scaling biases [75,76], a family of biases which occur when stimulus and response scales are mismatched, is relevant in the definition of attributes. This research shows that different ways of presenting and scaling an attribute, as well as the definition of upper and lower limits of the attribute scale, are the main causes of bias. Five biases are encompassed by this family: contraction bias (underestimating large sizes/differences and overestimating small/size differences); logarithmic response bias (using step-changes in the number of digits used in the response, which fit a log scale); range equalizing bias (using most of the range of response whatever is the size of the range of the stimuli); centering bias (producing a symmetric distribution of responses centered on the midpoint of the range of stimuli); and equal frequency bias (using equally all parts of the response scale).

Studies on attribute framing effects [58,59] are also relevant in the definition of attributes, as they show that the gain-loss bias may occur when an attribute has a positive or negative connotation (e.g. whether assessing the degree of success, or instead, failure of a decision alternative). Poulton [76:12] suggests some generic ways of dealing with each magnitude judgment bias, and Levin et al. [59] mentions in which situations the gain-loss bias are more prevalent.

Proxy attributes are often used in multi-criteria analysis, when fundamental attributes are hard to measure. For example, it is often easier to measure the amounts of pollutants emitted per year by a power plant than to determine the health effects that result from the pollution. Fischer et al. (1987) have shown that using proxy attributes instead of direct measures of fundamental objectives leads to the proxy bias - distortion in weights in multiattribute utility models.

Debiasing. Whenever possible the attribute scales should use natural units (such as square footage to measure the size of a house), making sure that the range of the scale matches the spread of performances of the alternatives. When natural scales are not available, constructed attributes should be used with special attention to steps of the scale and its end points [91]. Care should also be taken in considering whether the attribute has a positive or negative frame in assessing performances. From a broader perspective, Keeney [46] emphasizes the importance of the selection of appropriate attributes, and Keeney & Gregory [43] provide excellent guidelines on how to choose and build an appropriate attribute. The analyst must ensure that the attributes are unambiguous for the assessment of consequences, comprehensive in covering the range of consequences, measure as directly as possible the consequences, and is understandable by the decision makers.

5. Biases When Eliciting Value and Utility Functions

Value functions express the decision maker’s strengths of preference for evaluating alternatives in the absence of risk. Utility functions express both risk attitude and strengths of preference, for decisions
under uncertainty. There are several elicitation procedures for both value and utility functions [20,89], with the former requiring judgments about preferences and strengths of preferences among riskless outcomes, and the latter requiring choices among gambles.

**Biases.** Several studies show that the results of an elicitation of utility functions depend on the design of stimuli and responses [37,81]. In addition to random noise [33,50,84], both the anchoring bias [13], and the gain-loss bias [59] have been identified in this context. Another bias that affects utility assessment is the certainty effect [2,42], in which people prefer sure things to gambles with similar expected utilities, and discount the utility of sure things dramatically when they are even a slight degree of uncertainty [32,81]. In addition the desirability bias [91] might distort the utility function in a direction that favors a preferred alternative.

Examples of the impact of the gain-loss bias are the special role that the status quo plays in utility assessment [25], or the influence of the elicitation procedure employed (certainty-equivalent or probability-equivalent) on the shape of the function [32]. Another example is the impact that presenting a gamble in terms of gains or losses has on the utility function being elicited [32]. They may be mitigated by Arkes [4]'s suggestions on how to reduce psychophysically based errors. In terms of anchoring, Chapman and Johnson [13] have shown that value judgments are influenced by irrelevant starting points, but found out that prompting the subjects to consider reasons different than the anchor has alleviated the bias.

**Debiasing.** Many practitioners adopt simplified forms of elicitation and representation of partial values, given the noise associated with these elicitation procedures and the dependency of the responses on the framing of stimuli [91]. In many ways, value and utility functions are more “constructed” than “elicited” [83]. These simplifications include using value functions as approximations of utility functions, as advocated by von Winterfeldt & Edwards [89]; deriving utility functions from value functions [45]; or using standardized shapes for utility functions, such as linear value functions [91] or exponential utility function [60]. If utility functions are elicited using gambles, the analyst should avoid sure things in their elicitation to avoid the certainty effect.

Often multi-criteria models are created to support group decision making using decision conferences [73], with the decision analyst as a facilitator [27]. This opens up the issue on how individual value assessments should be combined (see Belton & Pictet [6]) and biases in groups. There is evidence that the degree of shared mental models by group members increases the effectiveness in reaching a decision and satisfaction with the decision making process [86]; and that the aggregation of preferences that are perceived by the group as procedurally fair can increase satisfaction with and legitimacy of decision making [52]. However, groups are more confident than individuals [52], sometimes showing overconfidence [51] and, in addition, they may polarize, thus exacerbating cognitive and motivational biases.

6. **Biases When Eliciting Attribute Weights**

Attribute weights define the tradeoffs between possibly conflicting objectives. They are scaling constants that are used to aggregate single attribute value or utility functions [44]. There are many common mistakes in defining weights [47,87], and several protocols for eliciting weights in an appropriate way [89].

**Biases.** Research has identified a family of biases affecting the elicitation of weights. According to the splitting bias objectives that are defined in more detail receive a larger portion of the weights than objectives that are defined in less detail [11,78,88] (but see some criticisms about the experimental settings of these studies in Pöyhönen & Hämäläinen [77]). With the equalizing bias decision makers tend to allocate similar weights to all objectives [26,35]. The gain-loss bias may also affect weights, for instance if trade-offs are elicited considering relative improvements or degradations of performances [87].

Two other biases occurring in this step are the proxy bias and the range-insensitivity bias. According to the proxy bias objectives are overweighted when measured by a proxy attribute instead of by an attribute that directly measures a fundamental objective [22]. According to the range insensitivity bias, weights are insensitive to the range of attribute values [29,68]. Because weights are scaling constants that should depend on attribute ranges, this insensitivity can lead to highly distorted weight judgments. Finally, the desirability bias [91] may lead to the over/under weighting of attributes to favor a preferred alternative.

**Debiasing.** Elicitation procedures that ask for direct assessments of importance should not be used [68] at all, but even methods that explicitly make decision makers consider the range of attributes, such as swing-weights and the trade-off method may suffer from range-insensitivity bias [23]. In practice most decision analysts use simple methods, such as swing weights, cross-checked with selected trade-offs [91], and they consider the weighting process as an
interactive and constructive process rather than as one of discovery [71]. To reduce the splitting bias one should avoid excessive detail in some objectives and little detail in others. This can often be achieved by obtaining objectives and attributes from multiple stakeholders, which provide different degrees of detail to different parts for the value tree (e.g., environmentalists provide detail about environmental objectives and engineers provide detail about cost and performance). To reduce the equalizing bias, one can set up the lower and upper anchors of each attribute in a way that they indeed allow similar weights for all objectives (as in the case study described by Morton et al. [66]). Alternatively, one can use ranking and ratio weighting methods, coupled with hierarchical weighting, which generally produce steeper weights [85]. Another way of dealing with the joint effects of the splitting bias and the equalizing bias is the calibration method proposed by Jacobi & Hobbs [35]. Finally, the use of either natural or constructed attributes for fundamental objectives, as recommended by Keeney & Gregory [43], avoids the proxy bias.

7. Conclusions and a Research Agenda

In this article, we reviewed cognitive and motivational biases that can occur in Multi-Criteria Decision Analysis (MCDA). Considering the importance of eliciting judgments as inputs into an MCDA (alternatives and objectives, attributes, value or utility functions, weights), it is somewhat surprising that not more attention has been paid to the possible distortions of an analysis due to these biases. In conclusion, we suggest a research agenda, which is based on our review of the existing literature on biases and evidence about their effects, as well as on the practices for debiasing.

Further exploration of motivational biases. Motivational biases are very important in MCDA, ranging from issues related to obvious conflicts of interest to subtle influences of professional association or preferences for outcomes of an analysis. Motivational biases can lead to excluding alternatives that are possible competitors of a preferred alternative, adding objectives that favor a preferred alternative or eliminating objectives that favor others, and, perhaps most importantly, manipulating weights to favor an alternative. While there exists some literature on motivational biases, it is not directly connected to the judgment tasks involved in MCDA. Therefore much more research is needed to better understand the effect of motivational biases in MCDA and how to reduce these biases.

Testing best practices to reduce cognitive biases. MCDA practitioners employ many “best practices” in debiasing, but few of those have been tested experimentally. Thus a high-priority item on our research agenda is to identify these best practices, and to test them in controlled experiments. Examples are the use of counterfactuals to reduce anchoring and probing and prompting strategies to reduce omission biases.

Testing best practices to reduce motivational biases. This is a virtually unexplored field. MCDA practitioners use some “tricks” to reduce motivational biases (counterfactuals, hypothetical bets, scoring rules), but with the exception of scoring rules, none of these have been tested. There is a huge opportunity for experimental researchers to explore current best practices and to test their effectiveness in reducing motivational biases.
8. References


[29] Gabrielli, W.F. and von Winterfeldt, D. Are importance weights sensitive to the range of alternatives in multiattribute utility measurement? Los Angeles, CA, Research Report 78-6, Social Science Research Institute, University of Southern California, 1978.


