This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Lahtinen, Tuomas J.; Hämäläinen, Raimo P.; Jenytn, Cosmo

On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis

Published in:
European Journal of Operational Research

DOI:
[10.1016/j.ejor.2019.09.004](https://doi.org/10.1016/j.ejor.2019.09.004)

Published: 01/04/2020

Document Version
Peer reviewed version

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Lahtinen, T. J., Hämäläinen, R. P., & Jenytn, C. (2020). On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis. *European Journal of Operational Research*, 282(1), 201-210. <https://doi.org/10.1016/j.ejor.2019.09.004>

This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Lahtinen, Tuomas J.; Hämäläinen, Raimo P.; Jenytn, Cosmo

On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis

Published in:
European Journal of Operational Research

DOI:
[10.1016/j.ejor.2019.09.004](https://doi.org/10.1016/j.ejor.2019.09.004)

E-pub ahead of print: 01/01/2019

Document Version
Peer reviewed version

Please cite the original version:
Lahtinen, T. J., Hämäläinen, R. P., & Jenytn, C. (2019). On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis. *European Journal of Operational Research*.
<https://doi.org/10.1016/j.ejor.2019.09.004>

On preference elicitation processes which mitigate the accumulation of biases in multi-criteria decision analysis

Tuomas J. Lahtinen*, Raimo P. Hämäläinen, Cosmo Jenytin

tuomas.j.lahtinen@aalto.fi, raimo.hamalainen@aalto.fi, cosmo.jenytin@aalto.fi

Systems Analysis Laboratory, Department of Mathematics and Systems Analysis,
Aalto University School of Sciences.

*Corresponding author

Abstract

In the practice of multi-criteria decision analysis, biased responses to the preference elicitation questions may impact the outcome of the process. In particular, there is a risk that the effects of biases accumulate in favor of a single alternative or a subset of alternatives. In this paper, we develop new bias mitigation techniques for multi-criteria decision analysis, which are based on the idea that the effects of biases can cancel out each other in the preference elicitation process. The benefits of the techniques include that the decision maker does not need to try to change her behavior to avoid biases, and there are no numerical adjustments of her judgements. The new techniques that we propose are: 1. Introducing a virtual reference alternative in the decision problem. 2. Introducing an auxiliary measuring stick attribute. 3. Rotating the reference point. 4. Restarting the decision process at an intermediate step with a reduced set of alternatives. We simulate computationally how these techniques help mitigate biases in the Even Swaps process when the decision maker exhibits the loss aversion bias, the measuring stick bias, and makes random response errors. The techniques can also be applied in weight elicitation using the SWING and trade-off methods to reduce the aforementioned biases.

1. Introduction

The mitigation of the effects of cognitive biases is an important issue in the practice of decision analysis. So far, behavioral research has focused on reducing biases in individual preference elicitation tasks (Montibeller and von Winterfeldt 2015). There are only a few studies which have analyzed debiasing over multiple steps in the decision analysis process (see, e.g. Anderson and Hobbs 2002, Jacobi and Hobbs 2007, Hämäläinen and Alaja 2008).

Taking the path perspective (Hämäläinen and Lahtinen 2016, Lahtinen and Hämäläinen 2016, Lahtinen et al. 2017) naturally suggests that in bias mitigation we should consider the whole path, i.e. the sequence of steps or tasks carried out, in the decision analysis process. The effects of biases can build up along the steps taken and create a large overall bias. For example, the preference elicitation questions may be posed such that the decision maker repeatedly gives too much weight for a certain attribute. Thus, the effects of biases build up in favor of the alternatives that score the best in this attribute, which may lead to the selection of an inferior alternative. However, it is also possible that the effects of biases cancel out each other (Kleinmuntz 1990, Anderson and Hobbs 2002). The likelihood of selecting an inferior alternative is greater when biases build up rather than cancel out each other.

Looking for bias minimizing paths can be a more attractive approach than trying to train the decision maker to reduce biases in the individual steps of the elicitation process. Then one would not need to interfere with the decision makers' natural ways of responding to the elicitation questions by debiasing

procedures or by forcing them to change their behavior. Training people to avoid biases in multi-criteria decision making is often not easy nor very successful (Hämäläinen and Alaja 2008). Furthermore, reducing biases by adjusting the numerical judgments obtained from experts or stakeholders (see, e.g. Bleichrodt et al. 2001, Anderson and Hobbs 2002, Jacobi and Hobbs 2007) can be problematic. People may not trust results that have been technically adjusted or corrected by the analyst.

This paper introduces four techniques for creating paths with reduced overall effects of biases in preference elicitation. The first technique is to introduce a virtual reference alternative in the decision problem. The second is to introduce an auxiliary measuring stick attribute to be used in the assessment of trade-offs between attributes. The third technique is to rotate the reference point used. Lastly, the fourth is to restart the decision making process at an intermediate step with a reduced set of alternatives.

The idea of the introduction of a virtual alternative, or an auxiliary measuring stick is to bring in a new element to the decision analysis process. This can impact the path followed in the preference elicitation process, and how the decision maker thinks about the problem. For example, it is well known in marketing that introducing an additional alternative can influence the results of a decision process (see, e.g. Huber et al. 1982, Farquhar and Pratkanis 1993).

The introduction of a virtual alternative may help in particular in mitigating reference point related effects, such as the loss aversion bias (Tversky and Kahneman 1991). This possibility has not received attention previously although virtual alternatives are, in fact, commonly used in standard preference elicitation procedures. A virtual alternative is a hypothetical alternative that is not included in the original set of decision alternatives. For example, in trade-off tasks, which are sometimes called two-attribute matching tasks, the decision maker adjusts a given virtual alternative to make it equally preferred to another virtual alternative. The characteristics of these virtual alternatives can affect the results obtained (see, e.g. Delqu 2003, Deparis et al. 2015). In SWING weighting (von Winterfeldt and Edwards 1986) the decision maker is instructed to imagine a virtual alternative, which serves as an initial reference point. Typically, this alternative has the worst possible consequence in each attribute.

The introduction of an auxiliary measuring stick attribute can help to mitigate the effect of the measuring stick bias, which is also called the scale compatibility bias (Tversky et al. 1988). In preference elicitation processes utilizing trade-off tasks, the effect of this bias depends on the measuring stick attributes used (Lahtinen and Hmlinen 2016). In a trade-off task, the measuring stick attribute refers to the attribute in which the decision maker gives her response. The bias refers to the tendency to give extra weight to the measuring stick attribute (Delqu 1993). One earlier approach suggested for the mitigation of this bias is the modified trade-off technique by Delqu (1997). Anderson and Hobbs (2002), in turn, suggested to

adjust the weights elicited from the decision maker by using estimated bias coefficients or to use an averaging procedure.

The idea of rotating the reference point is similar to using multiple anchor points in the estimation of consequences (see, e.g. Montibeller and von Winterfeldt 2015). This idea may also be helpful in mitigating reference point related effects. The intermediate restarting technique, in turn, can help to eliminate the impacts of biases that have built up during the earlier steps. For example, a preference elicitation process could be restarted after dominated alternatives are found and excluded from the decision problem. Although iteration between phases is common in multi-criteria decision analysis (see e.g. Belton and Stewart 2002), we are not aware of any cases where a preference elicitation process has been restarted in order to mitigate biases.

Behavioral experiments testing debiasing methods over a complete elicitation process can be extremely laborious as the subjects need to carry out multiple evaluations. A convenient alternative is to use computational simulations. This option is attractive, in particular, when assessing the overall effects of biases that occur in different steps of the process. The prerequisite for a computational analysis is that one has a model of the biases considered. Naturally, when drawing conclusions from such an analysis it is important to acknowledge that the models of biases are not perfect and approximately describe the behavior of actual decision makers. Models of biases are presented, for example, in Bleichrodt et al. (2001), Anderson and Hobbs (2002), Delquié (2003), Jacobi and Hobbs (2007), and Lahtinen and Hämäläinen (2016). In the decision analysis literature, computational analyses have previously been used, for example, to evaluate the impact of approximations of value functions (Stewart 1996), compare weighting methods in the presence of response errors (Jia et al. 1998), and analyze the value of information in portfolio decision analysis (Keisler 2004). In the area of multi-criteria optimization, computational analyses have been utilized, e.g., to assess the effects of cognitive biases and to improve interactive methods (see, e.g. Stewart 1999, Stewart 2005, Ojalehto et al. 2016).

In this paper, we computationally demonstrate how the aforementioned techniques perform. As an example, we use the Even Swaps process (Hammond et al. 1998, 1999), however the ideas are also applicable with other decision analysis methods. One of the reasons for considering Even Swaps is that the decision maker is explicitly given the possibility of following different paths in the process. In our computational analysis, the decision maker is assumed to exhibit the loss aversion bias, the measuring stick bias, and to make non-systematic response errors.

2. Even Swaps and biases

The Even Swaps process (Hammond et al. 1998, 1999) supports choosing an alternative from a set of alternatives that are described with multiple attributes (see, e.g. Table 1). In an even swap, an alternative is replaced with a preferentially equivalent virtual alternative, which differs from the original alternative in two attributes. The decision maker conducts even swaps to make attributes irrelevant and to find dominated alternatives. These can be eliminated from the decision problem. An attribute is irrelevant if all alternatives share the same level of performance on that attribute. The decision maker carries out swaps until only one alternative remains in the consequence table. Software support for the Even Swaps method is provided by the Smart-Swaps software (Mustajoki and Hämäläinen 2005, 2007). The Even Swaps method is well known but there is a surprisingly limited number of reported applications (Kajanus et al. 2001, Gregory and Wellman 2001, Wakshull 2002, Hurley and Andrews 2003, Luo and Cheng 2006, Gomes et al. 2012, and Altun et al. 2016). Dereli and Altun (2012) have developed an extension of the Even Swaps method to negotiation support.

Table 1: A simple illustrative consequence table related to the choice of an apartment.

Attributes	Apartment alternatives		
	A	B	C
Rent (euros per month)	700	900	800
Size (square meters)	30	40	35
Condition score (1 to 10)	4	6	10

In the process, the decision maker conducts a series of even swap tasks consisting of two steps. First, the decision maker determines a *fixed change* in one attribute describing an alternative. For example: “The size of apartment B is changed from 40 to 35 square meters.” Second, the decision maker needs to specify a *compensatory change* in another attribute called the *measuring stick attribute*. For example: “To compensate for the decrease in size of apartment B, how much does the rent have to decrease from 900 so that the resulting modified alternative is equally preferred to the original one?” Table 2 shows the modified apartment B, which is dominated by apartment C.

Table 2: An example of an even swap in alternative B. The decision maker states that a change in Rent from 900 to 810 is equally valuable to a change in Size from 40 to 35. The modified apartment B is dominated by apartment C and thus B can be removed from the analysis.

Attributes	Apartment alternatives		
	A	B	C
Rent (euros per month)	700	900 810	800

Size (square meters)	30	40 35	35
Condition score (1 to 10)	4	6	10

Lahtinen and Hämäläinen (2016) have shown that in the even swap task people give extra weight to both the measuring stick attribute, and to the attribute where the change is a loss. These behaviors reflect the measuring stick bias and the loss aversion bias that have also been found in two-attribute matching tasks (Delqu 1993, 1997, Anderson and Hobbs 2002, Bleichrodt and Pinto 2002, Deparis et al. 2015). In the Even Swaps process, the measuring stick bias works in favor of those alternatives that are the best in the measuring stick attribute used. Extra weight to the loss equates to the decision maker providing too large compensatory changes when they are gains, and too small compensatory changes when they are losses. Therefore, alternatives tend to become more attractive every time they are swapped. We assume that the increase in attractiveness depends positively on the size of the fixed change.

The attribute elimination procedure described in Hammond et al. (1998) is the basic version for the Even Swaps process. In this procedure, attributes are made irrelevant one by one until only one attribute remains. In a simple version of the attribute elimination method, the decision maker has one fixed *reference alternative* throughout the process, and she uses the same measuring stick attribute in all swaps. The decision maker carries out swaps to make the attribute specific consequences of all other alternatives equal to those of the reference alternative in all attributes besides the measuring stick attribute. One can combine the attribute elimination procedure with pairwise comparisons of alternatives. In each pairwise comparison, the decision maker fixes one of the two alternatives as the reference alternative and performs swaps in the other alternative until either alternative becomes dominated.

Due to the measuring stick bias, the attribute elimination method is likely to favor those alternatives which are the best in the measuring stick attribute. Due to the loss aversion bias, the method is likely to favor those alternatives whose consequences, in attributes other than the measuring stick attribute, differ the most from the consequences of the reference alternative.

3. Bias mitigation techniques and methods

This section presents four techniques and five methods for mitigating the overall effect of biases, and demonstrates these for the Even Swaps process. Techniques 1 and 3 are alternative techniques for mitigating the loss aversion bias. Technique 2 can help to mitigate the measuring stick bias. Technique 4 aids in preventing the accumulation of biases and response errors when alternatives are eliminated one by one. The five methods make use of different combinations of these techniques. In this paper, for the sake of conciseness, we limit ourselves to consider what we think are the five most interesting combinations of

the bias mitigation techniques for the Even Swaps process. More bias mitigation methods can be developed by combining the techniques in different ways and by using variations of them.

Technique 1: Introduce a virtual reference alternative.

To minimize the effect of the loss aversion bias, the magnitudes of fixed changes made in the alternatives should be as close to each other as possible. When following the attribute elimination procedure, one approach is to create a virtual reference alternative whose consequence in each attribute is approximately equal to the average of the consequences of the original alternatives. For example, in cases with two alternatives and continuous attributes this ensures that the fixed changes made in the alternatives are equal in magnitude. In the computational analysis, the effectiveness of this technique is also studied in cases with more than two alternatives. Table 3 provides an example where a virtual alternative is brought in as the reference alternative. If the decision maker aims to make, e.g. the attribute Condition score to be an irrelevant attribute, the fixed changes needed in alternatives A, B, and C are 3, 1, and 3, respectively. The sizes of the fixed changes would be more imbalanced if one of the original alternatives was used as the reference alternative.

Table 3: A consequence table with a virtual reference alternative.

Attributes	Apartment alternatives			
	A	B	C	Virtual
Rent (euros per month)	700	900	800	800
Size (square meters)	30	40	35	35
Condition score (1 to 10)	4	6	10	7

Technique 2: Introduce an auxiliary measuring stick attribute.

To minimize the measuring stick bias, one technique is to introduce an additional attribute in which the alternatives are similar and use this as the measuring stick attribute. Therefore, the extra weight that the measuring stick attribute gains does not work in favor of any of the alternatives. Often, such an attribute can be found among the attributes originally left out of the analysis as they did not differentiate the alternatives. For example, commute time might be such an attribute if one is choosing apartments located in the same area (Table 4).

Table 4: Commute time has been introduced in the table to be the auxiliary measuring stick attribute.

Attributes	Apartment alternatives			
	A	B	C	Virtual

	A	B	C
Rent (euros per month)	700	900	800
Size (square meters)	30	40	35
Condition score (1 to 10)	4	6	10
Commute time (minutes)	60	60	60

Technique 3: Rotate the reference point.

Rotating the reference alternative aids in minimizing the effect of the loss aversion bias. When the attribute elimination method is used, the reference alternative can be rotated each time an attribute is eliminated. In this way, approximately the same number of swaps is carried out in all alternatives and the effects of the loss aversion in the different swaps partially cancel out each other. For example, consider a decision maker selecting between alternatives A and B presented in Table 3. If alternative B is the reference alternative, then all swaps would be carried out in the alternative A and the effects of loss aversion would accumulate in favor of alternative A. Instead, the decision maker could first use B as the reference point and change A's condition score from 4 to 6, thus making A's and B's condition scores equal. Next, the decision maker could change B's size to be equal to that of A. At this point, the attribute Rent would indicate the dominated alternative.

Technique 4: Intermediate restarting of the process with a reduced set of alternatives.

Pruning the set of decision alternatives changes the decision task. This idea can be used in different methods. In Even Swaps it is possible to restart the process each time after an alternative is eliminated. This removes the effects of biases and response errors that have built-up during the earlier steps. This technique can be particularly useful when the attribute elimination procedure is used in conjunction with the pairwise comparisons of alternatives. For example, consider a decision maker selecting between alternatives A, B and C presented in Table 3. Assume that the decision maker carries out two swaps so that alternative A dominates alternative B. At this point, the decision maker can eliminate alternative B from the consequence table and restart the Even Swaps process with the original unmodified alternatives A and C.

One can use the above techniques in different combinations to generate bias mitigation methods like those described below for the mitigation of biases in the Even Swaps process.

The Basic method: The attribute elimination procedure with a fixed reference alternative.

Initialization: Choose a reference alternative and a measuring stick attribute among the set of original alternatives and attributes.

Carry out even swaps such that the consequences of all the alternatives become equal to the consequences of the reference alternative in all attributes besides the measuring stick attribute. Use the same measuring stick attribute and reference alternative in every swap.

Example: The decision maker carries out the following process with Table 1. She decides that apartment C is the reference alternative and that rent is the measuring stick attribute. She carries out a total of four swaps in which apartments A and B are modified such that their size and condition score become equal to those of the apartment C. The compensatory changes are made in rent. In the end, rent indicates the only non-dominated alternative.

Method A: Attribute elimination method with a virtual reference alternative (Technique 1).

Initialization: Introduce a virtual reference alternative such that for each attribute, the consequence of the virtual reference alternative is equal to the average of the consequences of the original alternatives.

Choose a measuring stick attribute among the set of original attributes.

Carry out even swaps as in the Basic method.

Example: The decision maker carries out the following process with Table 3, which includes a virtual reference alternative. She decides that rent is the measuring stick attribute. She carries out a total of six swaps in which apartments A, B and C are modified such that their size and condition score become equal to those of the virtual reference alternative. The compensatory changes are made in rent. In the end, rent indicates the only non-dominated alternative.

Method B: Attribute elimination method with an auxiliary measuring stick (Technique 2).

Initialization: Choose a reference alternative from the set of original alternatives. Introduce an auxiliary measuring stick attribute in which the consequences of all alternatives are equal.

Carry out even swaps as in the Basic method.

Example: The decision maker carries out the following process with Table 4, which includes commute time as an auxiliary measuring stick attribute. She decides that alternative C is the reference alternative. She carries out a total of six swaps in which apartments A, B and C are modified such that their rent, size and condition score become equal to those of apartment C. The compensatory changes are made in commute time. In the end, commute time indicates the only non-dominated alternative.

Method C: Attribute elimination method with a virtual reference alternative (Technique 1) and an auxiliary measuring stick (Technique 2).

Initialization: Introduce a virtual reference alternative and an auxiliary measuring stick.

Carry out even swaps as in the Basic method.

Example: The decision maker carries out the following process with a modified version of Table 3, to which she has additionally introduced commute time as the auxiliary measuring stick attribute. She carries out a total of nine swaps in which apartments A, B and C are modified such that their rent, size and condition score become equal to those of the virtual reference alternative. The compensatory changes are made in commute time. In the end, commute time indicates the only non-dominated alternative.

Method D: Pairwise comparisons of alternatives using an auxiliary measuring stick (Technique 2), alternating reference alternative (Technique 3), and using intermediate restarting of the process (Technique 4).

Initialization: Introduce an auxiliary measuring stick attribute.

Carry out even swaps on a pair of alternatives such that their consequences become equal in all attributes except for the measuring stick attribute. Change the reference alternative after each swap. Eliminate the dominated alternative. Restore the original consequences of the remaining alternative in the considered pair. Carry out pairwise comparisons until only one alternative remains.

Example: The decision maker begins by carrying out the following steps to compare apartments A and B described in Table 4. Following technique 3, the rent of apartment A is made equal to the rent of apartment B, the size of B is made equal to the size of A, the condition of A is made equal to the condition of B. Following technique 2, all compensatory changes are made in commute time. Assuming that apartment A becomes dominated, the decision maker should eliminate this alternative from the table, and next carry out a similar process with apartment B and C. Following technique 4, she should use the original unmodified consequence data for apartment B.

Method E: Pairwise comparisons of alternatives using an auxiliary measuring stick (Technique 2), virtual reference alternative (Technique 1), and intermediate restarting of the process (Technique 4).

Initialization: Introduce an auxiliary measuring stick attribute.

Select a pair of alternatives. Introduce a virtual reference alternative whose consequence in each attribute is the average of the consequences of the two alternatives under comparison. Carry out even swaps such that the consequences of both alternatives become equal to the consequence of the reference alternative in all attributes besides the measuring stick attribute. Eliminate the dominated alternative and the virtual

reference alternative. Restore the original consequences of the remaining alternative. Carry out pairwise comparisons until only one alternative remains.

Example: The decision maker begins by carrying out the following steps to compare apartments A and B described in Table 4. Following technique 1, she introduces a virtual reference alternative with rent of 800 euros, size of 35 square meters, condition score of 5, and commute time of 60 minutes. She carries out total of six swaps in which apartments A and B are modified such that their rent, size and condition score become equal to those of the virtual reference alternative. Following technique 2, the compensatory changes are made in commute time. Assuming that apartment A becomes dominated, the decision maker should eliminate this alternative and the virtual reference alternative from the table, and next carry out a similar process with apartment B and C. Following technique 4, she should use the original unmodified consequence data for apartment B.

In appendix A, we prove that Method E eliminates the effects of biases completely if the decision maker's behavior is assumed to follow the model described in Section 4 and there are no random response errors. Basically, introducing an auxiliary measuring stick eliminates the measuring stick bias. Creating a new virtual reference alternative for each pairwise comparison of alternatives evens out the effect of the loss aversion bias within these pairwise comparisons. Intermediate restarting prevents the effects of biases from accumulating.

The number of swaps required is similar for all the methods besides Method E, which requires about twice as many swaps as the other methods. The number of swaps required with the Basic method is at most $(N - 1)(K - 1)$, $N(K - 1)$ with Method A, $(N - 1)K$ with method B, NK with Method C, $(N - 1)K$ with Method D, and $2(N - 1)K$ with Method E. The number of alternatives is denoted with N and the number of attributes with K .

4. Computational analysis

The effectiveness of each method described in the previous section is studied computationally in different settings using a Monte Carlo type simulation. We generate 225 000 000 cases each of which defines a decision problem and a decision maker. The parameters related to the decision problem include the number of alternatives, the number of attributes, and the consequences of the alternatives. The parameters related to the decision maker include the weight profiles describing the preferences of the decision maker, the magnitudes of biases, and whether the decision maker is assumed to make random response errors. In each case, the Even Swaps process is simulated computationally assuming that the decision maker follows the Basic method and Methods A-E. In total, 1 350 000 000 computational simulations of the Even Swaps process are completed. The performance criterion used to evaluate a method is the

percentage of cases in which this method gives the “correct result”, i.e. the same result as a bias and error free process would give.

In the even swap task, the decision maker determines a compensatory change Δ_m in the measuring stick attribute m to compensate for a fixed change Δ_r in the attribute r . We assume that if the decision maker is unbiased, her swaps would follow an additive linear value function. However, due to the measuring stick bias the weight of the measuring stick attribute, w_m , is increased by a factor S . Due to the loss aversion bias, the weight of the attribute in which the change is a loss is increased by a factor L . In addition, the decision maker can make random response errors. These are modeled with a random variable e .

If the fixed change is a loss, the compensatory change given by the decision maker is:

$$\Delta_m = \frac{L w_r}{S w_m} \Delta_r e. (1)$$

If the fixed change is a gain, the compensatory change given by the decision maker is:

$$\Delta_m = \frac{w_r}{L S w_m} \Delta_r e. (2)$$

Figure 1 describes the generation of the different cases included in the computational analysis. First, the decision problems are chosen to include 2, 5, or 8 alternatives and 3, 5, or 8 attributes. Next, 5000 consequence tables are generated for each problem size. Only non-dominated alternatives are included in each table. To generate one table, the attribute specific consequences of the alternatives are drawn from a uniform distribution. The table is then normalized such that, in each attribute the value 0 is assigned to the alternative that has the worst performance and the value 1 to the alternative that has the best performance. If there is a dominated alternative in the table, the table is rejected and a new one is generated. Next, 100 weight profiles are generated for each number of attributes. In each profile, the weights sum up to 1 and all weights are greater than 0.05. These weight profiles are drawn from a uniform distribution over the space of feasible weights. Next, the following bias coefficients S : 1.0, 1.1, 1.2, 1.3, 1.4, and L : 1.0, 1.1, 1.2, 1.3, 1.4 are chosen to be considered. These values are assumed to describe typical magnitudes of biases and they cover the point estimates obtained in Lahtinen and Hämäläinen (2016). Lastly, the cases defined by abovementioned parameter values are analyzed with and without random response errors. The random variable e is assumed to follow a log-normal distribution with median 1.0 and standard deviation 0.10. These properties are assumed to describe typical random response errors.

Figure 1: Generation of cases

Figure 2 describes in principle the computational simulation of the Even Swaps process. First an initialization is performed according to the method used. When the Basic method or Method B is followed, one of the alternatives available is randomly selected to be the reference alternative. When the Basic method or Method A is followed, one of the attributes in the consequence table is randomly selected to be the measuring stick attribute. The auxiliary measuring stick attribute for Methods B, C, D and E is created such that its weight is the average of the weights of the original attributes. After the auxiliary attribute is introduced, the weights are rescaled to sum to 1. In the simulated Even Swaps process, the fixed changes are selected based on the method used. When Method D or E is followed, swaps are carried out repeatedly on a selected pair of alternatives until one of them becomes dominated. The pair of alternatives is selected randomly among the remaining alternatives. In addition, when Method E is applied, a new virtual reference alternative is included in the consequence table every time a new pair is selected. This step is included only in Method E and is not shown in Figure 2. The compensatory changes are calculated based on equations (1) and (2) and the parameters related to the case.

Figure 2: Computational simulation of the Even Swaps process

Note that the generation of cases in the computational analysis includes both selected parameter values, i.e. magnitudes of biases, the magnitude of response error, and the sizes of the decision problems, as well as randomly generated parameter values, i.e. consequences of the alternatives and the weight profiles. The idea is to cover a wide range of possible situations. The individual simulation runs, in which a selected method is used in a single case include randomness due to the random response errors that are included in half of the cases, and in the way the decision maker is assumed to use the method. This does not affect the conclusions made in Section 5, because the results are analyzed on an aggregate level. In principle, it would be possible to perform multiple iterations using the same method in a given case. This could be desirable if one would want to analyze the performance of the methods more extensively in a defined case.

5. Results

The results in this section are based on the outcomes of the 1 350 000 000 computational simulations of the Even Swaps process. In total, there are 225 000 000 unique cases, each of which is a combination of number of alternatives, number of attributes, the consequences of alternatives with respect to the attributes, weight profile, magnitudes of biases, and whether random response errors are included. In each case the Even Swaps process is computationally simulated using all the six methods.

The overall performances of the methods are shown in Table 5. Methods A, C, and D perform better than the Basic method by 6 to 7 percentage points. Method E has the highest performance score. It finds the same result as the bias-free process in 98 percent of the cases studied and finds the correct result in all cases when there are no random response errors. The Basic method and Method B have the worst performance. Still, there are cases in which these methods outperform Methods A, C, and D (Table 6). Thus, one might wish to also consider these methods when choosing the method to be used in a real-life situation.

Table 5: The overall performances of the methods.

Method	Percentage of correct results
Basic	87
A	93
B	86
C	94
D	93
E	98

Table 6: Pairwise comparison of methods in cases with no response error. The cells show the percentage of cases in which using the method on the row leads to the correct result and using the method on the column does not.

	Basic	A	B	C	D	E
Basic		2	3	2	3	0
A	7		9	1	4	0
B	2	3		1	3	0
C	9	3	9		4	0
D	9	4	10	3		0
E	11	6	12	4	6	

Figures 3 to 7 illustrate that Methods A, C, D, and E perform better than the Basic method across different types of cases. The performance ranking of Methods A, C, and D depends on the magnitudes of the biases (Figures 3 and 4), on the number of attributes (Figure 6), and on the number of alternatives (Figure 7). However, there are not large differences between the performances of these three methods in any of the figures.

Methods A and C perform equally well when the magnitude of the loss aversion bias is 1.3 or higher (Figure 4). Thus, this implies no added benefit of using an auxiliary measuring stick attribute in addition to a virtual reference alternative when the magnitude of loss aversion is high. Figure 8 emphasizes that Method B performs better than the Basic method in cases where the magnitude of loss aversion bias is low relative to the magnitude of the measuring stick bias. The nonlinear shape of the surface for the Basic method in Figure 8 suggests that there is an interaction between the effects of the loss aversion bias and the measuring stick bias when the Basic method is used.

Figure 3: Performances for different magnitudes of the measuring stick bias.

Figure 4: Performances for different magnitudes of the loss aversion bias.

Figure 5: Performances for different magnitudes of the response error.

Figure 6: Performances for cases with different number of attributes.

Figure 7: Performances for cases with different number of alternatives.

Figure 8: Illustration of the differences between the Basic method and Method B with different magnitudes of biases using quadratic surfaces fitted to the data.

Overall, these results demonstrate how different procedures for carrying out the Even Swaps process affect its outcomes when the decision maker exhibits biases in a systematic way. Furthermore, it is shown that the likelihood of finding the best alternative can be increased by following a bias mitigating procedure. The results presented here rely on the assumption that the decision maker's preferences can be approximated with an additive value function.

6. Bias mitigation in the SWING method

So far, behavioral issues related to the SWING procedure have received very limited attention although SWING is a widely used multi-criteria weighting method. For example, the measuring stick bias has not been studied in the SWING procedure. The new bias mitigation techniques create interesting opportunities related to the SWING method. In this section, the use of techniques 1 to 3 with SWING is briefly characterized. In addition, we suggest a procedure for using technique 4, i.e. intermediate restarting, with the SWING method. Previously, methods for reducing the impact of errors and biases in SWING weighting have been discussed by Mustajoki et al. (2005).

In the standard SWING method, the decision maker is instructed to imagine a virtual reference alternative with the consequences on the worst level for every attribute. The ‘swing’ related to an attribute is a hypothetical improvement, in which the reference alternative is improved from the worst to the best consequence level in the attribute. The attribute with the most valuable swing is typically selected as the measuring stick and given fixed 100 importance points. Next, the rest of the swings are given importance points that are between 0 and 100. The attribute weights can be calculated by normalizing the importance points to sum to 1.

Technique 1: Introduce a virtual reference alternative.

Due to the choice of the virtual reference alternative in the standard SWING procedure, the swings represent ‘gains’ related to improving attribute specific consequences from the worst to the best level. Following the idea of technique 1, testing the SWING procedure with other virtual reference alternatives would be interesting. For example, the virtual reference alternative could have the consequences on the best level in every attribute. In this case, the swings would represent losses related to deteriorating the attribute specific consequences from the best to the worst level. The framing of swings as losses instead of gains may have an impact on the decision maker’s behavior due to the loss aversion phenomenon. Furthermore, it could be possible to frame swings as gains in some attributes and as losses in others.

Technique 2: Introduce an auxiliary measuring stick attribute.

Selecting the most important attribute as the measuring stick may create a bias in the SWING process. For example, if the measuring stick attribute is given too much weight, the process favors the alternative that has the best consequence in this attribute. Evidence for such a phenomenon can be found in existing research but further studies are needed (Lahtinen and Hämäläinen 2016, p. 897). One way to mitigate this phenomenon is to introduce an auxiliary measuring stick attribute, in which the alternatives have equal or approximately equal consequences. In addition, one would need to define a range for this attribute to be considered in the swing weighting.

Technique 3: Rotate the reference point.

In SWING it is not necessary to use only one fixed reference alternative or a measuring stick attribute. For example, rotating the measuring stick attribute within the procedure could help mitigate the measuring stick bias. Furthermore, one could carry out the process multiple times using different reference alternatives. One reference alternative could be a virtual alternative with worst possible consequences in every attribute, and another one could be a virtual alternative with best possible consequences in every attribute. In the former case, swings are framed as gains and in the latter as losses. The results obtained with different reference alternatives could be averaged.

Technique 4: Intermediate restarting of the process with a reduced set of alternatives.

Following the intermediate restarting technique, the SWING process could be carried out in two stages. In stage one, the alternatives are scored using attribute weights elicited with the standard SWING procedure. In stage two, the lowest scoring alternatives are eliminated such that the attribute ranges are reduced. Then the SWING process is carried out with a reduced set of alternatives and reduced attribute ranges.

The benefit of this process is that the decision maker can revisit the evaluation of the top alternatives of the first stage using updated attribute ranges. In the second stage, the attribute ranges in certain attributes may become very small. The decision maker can focus his attention on the attributes in which the top alternatives of initial evaluation rounds differ from each other. Such a process may help to cope with the range insensitivity phenomenon, which causes the attribute ranges used in SWING to influence the ranking of the alternatives (see, e.g. Fischer 1995).

Consider, for example, the decision problem shown in Table 7. When alternatives A, B and C are included in the problem, the wide range in attribute 1 is determined by alternative C. Therefore, due to the range insensitivity phenomenon it is likely that the decision maker gives too little weight to attribute 1. Consequently, the decision process is likely biased in favor of alternative B, which has a mediocre score in attribute 1 and the highest score in attribute 2. However, if alternative C is eliminated and the SWING process is restarted with the alternatives A and B only, then the swing range for attribute 1 would be reduced. In this situation, repeating the SWING procedure could change the ranking between alternatives A and B compared to the situation in which alternative C was included in the problem.

Table 7: Consequence table for the SWING example.

Attributes	Alternatives			Attribute range for A, B, C	Attribute range for A, B
	A	B	C		

Attribute 1	10	5	0	0 to 10	5 to 10
Attribute 2	7	10	8	7 to 10	7 to 10

7. Discussion

The risk of biases in decision analysis processes has been long acknowledged (see, e.g. Howard 1980, p. 14, Keeney 1982, p. 819) with an increasing interest today (Montibeller and von Winterfeldt 2015). However, thus far, there has been a limited interest in analyzing the overall effects of biases. Biases matter in particular, when they accumulate in such a way that the rank order of alternatives is affected.

Things we can do:

- Acknowledge the fact that the overall effect of biases depends on how the preference elicitation process is structured, and on the set of alternatives under comparison.
- Pay attention to the interplay of biases. For example, Lahtinen and Hämäläinen (2016) describe how the starting point phenomenon can increase the effect of the splitting bias.
- Carry out computational modeling of the overall effects of biases in order to analyze the effectiveness of bias mitigation methods and to design new procedures.
- Develop new bias mitigation approaches based on the idea that biases can cancel out each other along the preference elicitation path.
- Look for the debiasing measures that have the greatest positive impact. For example, Keeney (1992) suggests that it is particularly important to mitigate biases already in the problem structuring phase.

This paper studies bias mitigation in a complete preference elicitation process. Yet, preference elicitation is only one phase in the overall decision analysis process. In practice, it is important to pay attention and manage behavioral phenomena in the entire process including the phases related to creating objectives and alternatives. In general, we believe that a systemic perspective is essential in the search for more effective ways to mitigate biases in decision analysis processes. The growing literature on behavioral operational research also suggests taking a systemic approach to all model based problem solving and decision support (Hämäläinen et al. 2013, Franco and Hämäläinen 2016).

8. Conclusions

The overall effect of biases can be mitigated by designing the preference elicitation process such that the effects of biases in different steps of the process counteract each other. Mitigating the effects of biases in

this way does not force the decision maker to change her behavior or to learn to avoid biases and there are no numerical adjustments of her judgements. The bias mitigation approaches presented in this paper complement the earlier research in which the focus has been on minimizing the effect of biases in individual steps of the preference elicitation process.

In practice, it is likely best to tailor the bias mitigation methods to be used based on the information available about the case at hand. This may include conducting computational analyses before implementing the real process. Relevant information to be considered include, e.g., the number of alternatives, the number of attributes, the consequences of the alternatives, as well as estimates of the magnitudes of the biases of the decision makers in the specific context at hand. The performance ranking of the new debiasing methods depends on the values of these parameters. Moreover, the results of this paper refer to the case when the preferences of the decision maker follow the additive linear value function and simple bias coefficients can be used to represent the effects of the biases.

A computational approach, such as the one used in this paper, for the evaluation of the performance of bias mitigation methods can significantly aid in the process of developing and testing new debiasing approaches. Previous research on the effects of biases in decision analysis has primarily used behavioral experiments as a research method. However, testing debiasing approaches by using behavioral experiments for a sequential process is not easy. In the future, we would like to see the experimental and the computational approaches used together.

The bias mitigation techniques demonstrated in this paper with the Even Swaps process provide a basis for designing bias mitigation approaches for other decision analysis processes as well. This paper shows how to apply the ideas in the SWING procedure. Moreover, in weight elicitation based on trade-off tasks, the idea of a virtual measuring stick attribute is directly applicable. In addition, the techniques could be useful in interactive multi-criteria optimization. A general conclusion based on the perspective taken here is that more attention should be paid to the overall effects of biases in decision analysis processes.

Appendix A.

In this appendix we prove that Method E completely eliminates the effects of the measuring stick bias and the loss aversion bias if the decision maker's behavior is assumed to follow the model described in Section 4 without random response errors.

When the decision maker follows Method E, she carries out a sequence of pairwise comparisons of alternatives, in which she compares two original unmodified decision alternatives. Therefore, it suffices to

show that Method E reveals the alternative with the greater unbiased overall value in each pairwise comparison of alternatives.

Thus, we consider two alternatives, $A = (A_1, \dots, A_n, X)$ and $B = (B_1, \dots, B_n, X)$. The elements of the vectors A and B with indices from 1 to N indicate the attribute specific consequences of the alternatives with respect to the N attributes considered in the original decision problem. Without loss of generality we assume that the alternatives' consequences differ in all these attributes. X is the consequence of the alternatives in the auxiliary measuring stick attribute. A is better than B in attributes A_{better} , B is better than A in attributes B_{better} . The attribute weights are $w_1, \dots, w_n, w_x$. The unbiased overall values of the alternatives without the attribute X are $V(A) = \sum_{i \in 1, \dots, N} w_i A_i$ and $V(B) = \sum_{i \in 1, \dots, N} w_i B_i$. The weights as well as the bias coefficients are greater than zero.

When Method E is applied, in the fixed changes of the swaps that are carried out, the consequences of both alternatives are changed to $\frac{A_i+B_i}{2}$ in the attributes 1 to N . In the compensatory changes, the consequences of the alternatives are changed in the auxiliary attribute. Let A_{aux} and B_{aux} denote A 's and B 's resulting values in the auxiliary attribute. Following the equations (1) and (2) presented in Section 4, as a result of the swaps carried out, A 's consequence in the auxiliary attribute becomes:

$$\begin{aligned} A_{aux} &= X + \frac{L}{S} \sum_{i \in A_{better}} \frac{w_i}{w_x} (A_i - \frac{A_i+B_i}{2}) - \frac{1}{LS} \sum_{i \in B_{better}} \frac{w_i}{w_x} (\frac{A_i+B_i}{2} - A_i) \\ &= X + \frac{1}{2 S w_x} [L \sum_{i \in A_{better}} w_i (A_i - B_i) - \frac{1}{L} \sum_{i \in B_{better}} w_i (B_i - A_i)]. \end{aligned}$$

B 's consequence in the auxiliary attribute becomes:

$$B_{aux} = X + \frac{1}{2 S w_x} [L \sum_{i \in B_{better}} w_i (B_i - A_i) - \frac{1}{L} \sum_{i \in A_{better}} w_i (A_i - B_i)].$$

The difference between the alternatives' consequences in the auxiliary attribute is:

$$\begin{aligned} A_{aux} - B_{aux} &= \frac{L}{2 S w_x} [\sum_{i \in A_{better}} w_i (A_i - B_i) - \sum_{i \in B_{better}} w_i (B_i - A_i)] + \frac{1}{2 L S w_x} [\sum_{i \in A_{better}} w_i (A_i - B_i) - \sum_{i \in B_{better}} w_i (B_i - A_i)] \\ &= (\frac{L}{2 S w_x} + \frac{1}{2 L S w_x}) [\sum_{i \in 1, \dots, N} w_i A_i - \sum_{i \in 1, \dots, N} w_i B_i] = \frac{1}{2 S w_x} (L + 1/L) [V(A) - V(B)]. \end{aligned}$$

We can see that the auxiliary attribute reveals the alternative with the greater overall value since, the difference between the alternatives' consequences in the auxiliary attribute, $A_{aux} - B_{aux}$, is positively proportional to the difference of the unbiased overall values of the alternatives, $V(A) - V(B)$. ■

References

- Altun, K., von Zedtwitz, M., and Dereli, T. (2016). Multi-issue negotiation in quality function deployment: Modified Even-Swaps in new product development. *Computers & industrial engineering*, 92, 31-49.
- Anderson, R.M., and Hobbs, B.F. (2002). Using a Bayesian approach to quantify scale compatibility bias. *Management Science*, 48(12), 1555-1568.
- Belton, V., and Stewart, T.J. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers.
- Bleichrodt, H., Pinto, J.L., and Wakker, P. (2001). Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Management Science*, 47(11), 1498-1514.
- Bleichrodt, H., and Pinto, J.L. (2002). Loss aversion and scale compatibility in two-attribute trade-offs. *Journal of Mathematical Psychology*, 46(3), 315-337.
- Delquié, P. (1993) Inconsistent trade-offs between attributes: New evidence in preference assessment biases. *Management Science*, 39(11), 1382-1395.
- Delquié, P. (1997) "Bi-matching": A new preference assessment method to reduce compatibility effects. *Management Science*, 43(5), 640-658
- Delquié, P. (2003). Optimal conflict in preference assessment. *Management Science*, 49(1), 102-115.
- Deparis, S., Mousseau, V., Öztürk, M., and Huron, C. (2015). The effect of bi-criteria conflict on matching-elicited preferences. *European Journal of Operational Research*, 242(3), 951-959.
- Dereli, T., and Altun, K. (2012). Modified Even-Swaps: A novel, clear, rational and an easy-to-use mechanism for multi-issue negotiation. *Computers & Industrial Engineering*, 63(4), 1013-1029.
- Farquhar, P. H., and Pratkanis, A. R. (1993). Decision structuring with phantom alternatives. *Management Science*, 39(10), 1214-1226.
- Fischer, G. W. (1995). Range sensitivity of attribute weights in multiattribute value models. *Organizational Behavior and Human Decision Processes*, 62(3), 252-266.
- Franco, L.A., and Hämmäläinen, R.P. (2016). Engaging with behavioural OR: On methods, actors, and praxis. In Kunc, M, Malpass, J, White, L (Eds.) *Behavioural operational research: Theory, methodology and practice*, 1-14.
- Gomes, L. F. A. M., Rangel, L. A. D., and Fernandes, P. P. 2012. Decision aiding in plastic surgery: a multicriteria analysis. *Pesquisa Operacional*, 32(2), 371-387.

- Gregory, R., and Wellman, K. 2001. Bringing stakeholder values into environmental policy choices: A community-based estuary case study. *Ecological Economics*, 39(1), 37-52.
- Hurley, W. J., and Andrews, W. S. 2003. Option analysis: using the method of even swaps. *Canadian Military Journal*, 4(3), 43-46.
- Hämäläinen, R. P., and Alaja, S. (2008). The threat of weighting biases in environmental decision analysis. *Ecological Economics*, 68(1), 556-569.
- Hämäläinen, R. P., Luoma, J., and Saarinen, E. (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. *European Journal of Operational Research*, 228(3), 623-634.
- Hämäläinen, R. P., and Lahtinen, T. J. (2016). Path dependence in Operational Research – How the modeling process can influence the results. *Operations Research Perspectives*, 3, 14-20.
- Hammond, J.S., Keeney, R.L., and Raiffa, H. (1998). Even swaps: A rational method for making trade-offs. *Harvard Business Review*, 76, 137-150.
- Hammond, J., Keeney, R.L., and Raiffa, H. (1999). *Smart choices: A practical guide to making better decisions*. Harvard Business School Press.
- Howard, R. A. (1980). An assessment of decision analysis. *Operations Research*, 28(1), 4-27.
- Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1), 90-98.
- Jacobi, S. K., and Hobbs, B. F. (2007). Quantifying and mitigating the splitting bias and other value tree-induced weighting biases. *Decision Analysis*, 4(4), 194-210.
- Jia, J., Fischer, G. W., and Dyer, J. S. (1998). Attribute weighting methods and decision quality in the presence of response error: a simulation study. *Journal of Behavioral Decision Making*, 11(2), 85-105.
- Kajanus, M., Ahola, J., Kurttila, M., and Pesonen, M. (2001). Application of even swaps for strategy selection in a rural enterprise. *Management Decision*, 39(5), 394-402.
- Keeney, R. L. (1982). Decision analysis: an overview. *Operations Research*, 30(5), 803-838.
- Keeney, R.L. (1992). *Value-focused Thinking: a Path to Creative Decision Making*. Harvard University Press.
- Keisler, J. (2004). Value of information in portfolio decision analysis. *Decision Analysis*, 1(3), 177-189.

- Kleinmuntz, D. N. (1990). Decomposition and the control of error in decision analytic models. In Hogarth, R.M. (Ed.) *Insights in decision making: A tribute to Hillel J. Einhorn*, University of Chicago Press, 107-126.
- Lahtinen, T. J., and Hämäläinen, R. P. (2016). Path dependence and biases in the even swaps decision analysis method. *European Journal of Operational Research*, 249(3), 890-898.
- Lahtinen, T. J., Guillaume, J. H., and Hämäläinen, R. P. (2017). Why pay attention to paths in the practice of environmental modelling? *Environmental Modelling & Software*, 92:74-81.
- Luo C.-M., and Cheng, B.-W. (2006). Applying Even-Swap Method to Structurally Enhance the Process of Intuition Decision-Making, *Systemic Practice and Action Research* 19(1), 45-59
- Montibeller, G., and von Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Analysis*, 35(7), 1230-1251.
- Mustajoki, J., and Hämäläinen, R. P. (2005). A preference programming approach to make the even swaps method even easier. *Decision Analysis*, 2(2), 110-123.
- Mustajoki, J., Hämäläinen, R. P., & Salo, A. (2005). Decision support by interval SMART/SWING—incorporating imprecision in the SMART and SWING methods. *Decision Sciences*, 36(2), 317-339.
- Mustajoki, J., and Hämäläinen, R.P. (2007). Smart-Swaps—A decision support system for multicriteria decision analysis with the even swaps method. *Decision Support Systems*, 44(1), 313-325.
- Ojalehto, V., Podkopaev, D., and Miettinen, K. (2016). Towards automatic testing of reference point based interactive methods. In: Handl J. et al. (Eds.) *Parallel Problem Solving from Nature – PPSN XIV*, Springer, 483–492.
- Stewart, T. J. (1996). Robustness of additive value function methods in MCDM. *Journal of Multi-Criteria Decision Analysis* 5(4), 301-309.
- Stewart, T.J. (1999). Evaluation and refinement of aspiration-based methods in MCDM. *European Journal of Operational Research*, 113(3), 643-652.
- Stewart, T. J. (2005). Goal programming and cognitive biases in decision-making. *Journal of the Operational Research Society*, 56(10), 1166-1175.
- Tversky, A., Sattath, S., and Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95(3), 371-384.

Tversky, A., and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039-1061.

von Winterfeldt, D., and Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge University Press.

Wakshull, M.N. (2002). Application of “Even Swaps” to Normalize Qualitative and Quantitative Risk Valuations, *Proceedings of the Project Management Institute Annual Seminars & Symposium*, San Antonio, Texas.