Revised version, submitted 5.6.2015

Path Dependence in Operational Research – An Illustration with the Even Swaps Decision Analysis Method

Raimo P. Hämäläinen* and Tuomas J. Lahtinen
Systems Analysis Laboratory, Department of Mathematics and Systems Analysis,
Aalto University School of Science, P.O. Box 11100, 00076 Aalto, Finland
raimo.hamalainen@aalto.fi, tuomas.j.lahtinen@aalto.fi

Abstract

There are usually alternative ways of using models to solve a problem in Operational Research. Different modeling processes can lead to different outcomes. This path dependence can originate from the system created by the modeling team, learning, procedure, behavior, motivation, uncertainty or external causes, and interplay of these factors. To demonstrate the phenomenon, we show experimentally that path dependence can occur in the Even Swaps decision analysis method. This provides an example of path dependence originating from procedural and behavioral origins. We explain the results by the accumulation of the effects of two known biases; loss aversion and scale compatibility. Strategies to reduce the accumulation of the effects of these biases are suggested.

Keywords: Behavioral Operational Research, OR Practice, Decision Analysis, Path Dependence, Lock-in, Scale Compatibility

1. Introduction

Path dependence is a concept discussed in economics (David 1985, Arthur 1989, Martin and Sunley 2006), policy studies (Webster 2008, Levin et al. 2012), sociology (Sterman and Wittenberg 1999, Mahoney 2000, Gartland 2005), political science (Pierson 2000), and organizational decision making (Sydow et al. 2009). It is usually related to the lock-in phenomenon which refers to the emergence of strong anchor points from which it is not easy to move forward. The most famous example is how the QWERTY layout has become the worldwide standard for keyboards (David 1985). More generally, the meaning assigned to path dependence is that 'history matters', i.e., the current state of the world depends on the path taken. In this paper we want to bring path dependence into the focus also in Operational Research (OR) in general. We see that the topic is of both theoretical and practical interest in model based problem solving and decision making. Earlier research on path dependence in other disciplines has focused on exposing and describing it. In OR we also want to find ways to avoid it or to mitigate the risks related to it. Path dependence is directly related to the emerging area of Behavioral Operational Research (Hämäläinen et al. 2013).

* Corresponding author. Tel.: +358 500677942

The possibility that two valid but different OR modeling processes lead to different results and policy recommendations has been noted already early (Landry et al. 1983). Also the literature on best practices in OR (see, e.g. Morris 1967, Pidd 1999, Walker 2009) does implicitly acknowledge the possibility of path dependence since alternative practices are seen to be possible. Little (1970) and Ormerod (2008) have suggested to adjust the OR intervention and the model adaptively as the process evolves and the project team's competence develops. The literature on the ethics of OR also provides examples that the OR process matters (Rauschmayer et al. 2009, Ormerod 2013). Although these issues have been observed early, the subsequent research is still rather limited and scattered.

We use the term path dependence to refer to all the phenomena where different paths lead to different outcomes. By a path we mean the sequence of steps that are taken in the modelling or problem solving process. The steps can include, for example, the framing and structuring of the problem, the choice of model, the order in which different parts of the model are specified and solved, and the way in which data or preferences are collected.

In many contexts we would naturally want to minimize the possibility and effects of path dependence. This is the case in particular in prescriptive modeling when the goal is, e.g. to optimize or to reach efficiency. Also in important policy problems one should, at least, be aware of the possibility of path dependence and its origins and of the possible range of its consequences. Yet, there are situations where the main benefits expected from an OR project are related to learning and to the creation of a shared understanding of the problem as a whole. Then path dependence might not be a serious concern. In fact, reaching different conclusions along different paths could improve learning.

We believe that the term path dependence is useful as an integrative perspective and will help to give structure to the future studies. In this paper, we first discuss the phenomenon and its occurrence in OR in general and identify seven types of origins for path dependence: systemic, learning, procedure, behavior, motivation as well as uncertainty and external origins.

To demonstrate the phenomenon we show experimentally how path dependence can emerge in decision analysis from procedural and behavioral origins. The experiment considers the Even Swaps method (Hammond et al. 1998, 1999) which is well known, simple and uses clearly defined paths. In this method a path consists of the sequence of even swaps that the decision maker (DM) carries out to eliminate alternatives and attributes one by one until the 'best' alternative is found.

Multiple strategies exist for carrying out the Even Swaps process, each leading the decision maker to a different path. We show the existence of path dependence by experiments where Even Swaps is used with the Smart-Swaps software (Hämäläinen et al. 2004, Mustajoki and Hämäläinen 2005, 2007). Different paths are shown to lead to different choices. This is explained by the accumulation of the effects of successive biased even swaps. The biases in the swaps are shown to be due to scale compatibility and loss aversion effects. Estimates of the magnitudes of these biases in the even swap tasks are also provided. We suggest ways to reduce the risk of path dependence in the Even Swaps method.

2. Path dependence

We describe seven drivers and origins of path dependence: systemic, learning, procedure, behavior and motivation, as well as uncertainty and external origins. These drivers can interact and occur together. We hope that thinking of these perspectives and the related examples, one can better understand the phenomenon and its possible emergence in OR.

2.1. Systemic origins

Systemic origins of path dependence relate to the system under study as well as to the system formed by the people and the organizations involved in the problem solving process. In the practice of OR the organizational system set up to solve the problem can cause path dependence, for example, due to lock-in and irreversibilities. A problem solving team and a stakeholder group can get locked-in to one approach and one software as they become more and more involved in using it. This leads to a problematic situation when there can be new better approaches available but the problem solving team keeps on using the old one. It is not easy to give up old habits (Barnes et al. 2004). Also the groupthink phenomenon (Janis 1982) in OR-communities of practice can lead to lock-in. The members of the problem solving team can convince each other of the correctness of the approach designed by the team because critical thinking might challenge the team's authority. Cultural effects can also be seen to be systemic. For example, in developing countries the problem solving culture can be different from western countries and this can impact the OR process (Lee and Olson 1980).

Irreversibility exists if it is impossible to undo the steps taken and start the process again once one path is started. Irreversibility can be due to organizational reasons (Sydow et al. 2009). Consider, for example, a problem that has initially been split into sub-problems that are separately dealt with in different groups. Later it can turn out that another partitioning would likely lead to a better solution. Revisiting the initial partitioning may, however, be impossible due to personnel, budget or time constraints. Irreversibility is an important factor to be taken into account when dealing with the modeling of wide ranging policy issues such as climate policies (Webster 2008).

Systemic origins of path dependence can also be ‘mathematical’. The dynamics of nonlinear systems can create path dependence due to increasing returns, bifurcation points as well as feedback loops. It is also well known that complex nonlinear systems can be very sensitive to the initial conditions.

Increasing returns is identified as the cause of path dependence in the seminal paper on technological development by Brian Arthur (1989). The paper notes that systems with increasing returns can include multiple equilibria to which the process can end up. Path dependence due to increasing returns can also exist in the development of regional economies (Martin and Sunley 2006) and in organizational decision making (Sydow et al. 2009). The sources of increasing returns and subsequent path dependence discussed in the above mentioned papers include learning, coordination, and complementarity effects. These phenomena can exist in OR problem solving too.

Bifurcation points are typical, e.g., in fishery models (Clark 1990). In fishery management problems it can be that overfishing has caused the collapse of a fishery and then it can be impossible to restore it in the short run by regular fishery management policies. Thus optimizing the policy is dependent on the history. The modelling of feedback loops is the focus in systems dynamics (see, e.g., Sterman 2000) and the results are typically sensitive to behavioral assumptions. Sterman and Wittenberg (1999) demonstrate that feedback loops can drive path dependence in the development of science. In their model, higher confidence in a scientific paradigm increases the rate at which the paradigm is used to solve puzzles and vice versa. The same argument could also apply to OR problem solving.

2.2. Learning

During the modeling process the OR specialists as well as the problem owners and stakeholders learn and their understanding increases about the problems which are being modelled. The interests of the modeling team can be directed to different aspects and perspectives as they learn different characteristics of the problem (see, e.g. Lane 1993). The fact that learning takes places in the modeling process has been recognized especially in systems dynamics (see, e.g. Sterman 1994, Lane 1992) and problem structuring (Checkland 2000, Ackerman 2012) as well as in participatory decision analysis (McDaniels et al. 2004, Salo and Hämäläinen 2010). Therefore, it is likely that learning often also has an impact on the outcome.

In participatory processes the time of formal engagement with the problem owners and stakeholders can be important. If the start and the introduction of the OR facilitator is delayed, the stakeholders can have committed to a heuristic problem solving process. Thus their problem framing and expectations of the results may have already been fixed. Then it can be difficult to launch an open structured problem solving process and unlearn the early ideas.

Keeney (1994) suggests that people should spend more initial effort to consider the values related to a problem. He argues that we commonly first think about alternatives, and not values, and this reduces our creative thinking. For example, decision makers may spend too much time on thinking about incremental changes in the status quo solution. Following the advice of Keeney, in OR problem solving it might be beneficial to start the process by carefully studying the goals and objectives of the decision makers and stakeholders. Only then should one choose the actual problem solving approach to be used. Experimental research suggests that the use of value-focused thinking influences the objectives and alternatives which are considered (see, Leon 1999, Bond et al. 2008, Selart et al. 2011).

2.3. Procedure

Procedural origins of path dependence relate to the properties and structure of the problem solving procedure used.

Procedural path dependence can be due to the technical properties of the mathematical methods used. For example, it is well known that stepsize effects exist in the solution algorithms used in OR. Use of different stepsizes can lead to different solutions. We can end up to a local or the global optimum depending on the iteration parameters used in numerical optimization. The solution that is found can also depend on the initial starting point. Technical path dependence has been shown to exist in the construction of regression models in statistical analysis where the forward selection and backward elimination methods for variable selection can produce different models (see, e.g. Derksen and Keskelman 1992).

Effects related to the order of steps taken can occur in sequential decision processes and lead to path dependence even without any behavioral causes. For example, when multiple decision makers are involved in strategic decision making the order of choices often has an impact on the outcome. A well-known effect is the so called first mover advantage in games which has been discussed in different economic settings and management decisions (see, e.g., Lieberman and Montgomery 1988, Varian 1992).

In multi-method processes (see, e.g. Hämäläinen et al. 2001, Franco 2011) it is possible that the order in which the methods are used can affect the outcome. In climate modeling the choice of the initial perspective can be important. For example, whether the modeling process is started with a socioeconomic or an atmospheric perspective can have an effect on which issues will be given the most attention to.

In large modeling problems it can be impractical or difficult to build an overall aggregate model. Rather, the problem needs to be decomposed into sub-problems which are solved separately. The decomposition method and the order in which different subsystems are modelled can affect the solution. Such problems can be found, e.g. in healthcare and airline industries (Jun et al. 1999, Barnhart et al. 2003).

2.4. Behavior and motivation

Behavioral drivers of path dependence relate to the possible effects of different behavioral phenomena and cognitive biases (see, e.g. Hämäläinen et al. 2013, Montibeller and Winterfeldt 2015). These can have different overall effects on different paths. Thus the result can depend on the path followed. The experiment on path dependence in the Even Swaps method presented in this paper demonstrates one such case. In multi-criteria decision analysis it is known that different steps related to, for example, problem structuring and alternative generation can influence the outcome, see the discussion in Section 3.

Anchoring (Tversky and Kahneman 1974) is one behavioral phenomenon which can influence the outcome of the OR process. Anchoring effects have been observed in decision support systems (George et al. 2000), in multi-criteria optimization (Buchanan and Corner 1997), elicitation of preferences (see, e.g., Lenert 1998, van Exel et al. 2006), negotiation (Kristensen and Gärling 1997), as well as in valuation, probability estimation and forecasting (for a review, see Furnham and Boo 2011). So the choice of initial steps to be taken can direct the OR process to different paths due to anchoring.

Behavioral reasons and biases can also lead to lock-in type situations in modeling. The status quo bias (Samuelson and Zeckhauser 1988) refers to the tendency to prefer the current solution or approach over the possible new ones. The sunk cost effect (Arkes and Blumer 1985) refers to the phenomenon where people want to keep on committing resources to a project to which they have previously invested. This happens regardless of whether the earlier investments have been successful or not. For example, an organization can have initially adopted a certain modelling based tool, such as a spreadsheet model, to support its operations. Over time this model can have grown excessively and become unwieldy and nontransparent. Still the organization can keep on using the old model. The reason can be the sunk costs and effort put into the old approach.

Motivational origins of path dependence are related to situations where people's goals affect the problem solving process. This risk is high when the problem is messy and controversial with alternative modeling approaches being possible. An unethical modeler may choose an approach which leads to results that are desired by her (Rauschmayer et al. 2009). The confirmation bias (Nickerson 1998) can lead the modeler to seek for approaches that support his prior beliefs about the right solution to the problem. Strategic behavior is likely to be found in group processes. The stakeholders in participatory modelling projects can try to influence the outcome by strategic behavior, for example, by intentionally emphasizing some features of the problem (Hämäläinen 2015). In bargaining it can be favorable to strategically select the initial offer or even misrepresent one's preferences to set the process on a favorable path (see, e.g. Raiffa 1982).

2.5. Uncertainty and changes in the external environment

Uncertainty can exist in the model assumptions as well as in the external environment. If the same modeling process is repeated, it can lead to different outcomes due to changes in the external environment. The basic assumptions of the model are not always clear and fixed. Different estimates of the parameters in a model naturally lead to different results. Large structural uncertainties are faced, for example, in climate models (see, e.g. Webster et al. 2003) which include many important subsystems, such as, socioeconomic, weather, solar, oceanic and industrial systems. In the comprehensive aggregate model there can remain uncertainties related to the interaction of the different subsystems.

One suggested solution to identify and mitigate the effects of structural uncertainty is the use of multi-modeling and averaging out the errors in different model based predictions (Knutti et al. 2010). However, the question of how to weight the outputs from different models creates new behavioral challenges in multi-modelling.

Scenario analysis is traditionally used to account for future uncertainties in policy modeling (see, e.g. Schoemaker 1991). We can also try to find robust adaptive strategies which can be adjusted based on how the future unfolds (Rosenhead 1990). However, developing and adopting an adaptive policy can be difficult in practice (Lempert and Groves 2010).

Changes in the external environment can relate, for example, to the market situation. In many political and economic decisions the timing of the start of the decision making process can be very crucial. Delaying the start can mean that values of the model parameters have changed which again can make some outcomes unreachable. However, sometimes when uncertainties exist it can be beneficial to postpone the early decisions to wait for more accurate information to become available (Arrow and Fisher 1974). Model based maintenance strategies (see, e.g. Scarf 1997) provide an example where wearing is an external driver of the process.

2.6. What to do in practice?

Awareness of path dependence is the first step. Based on the arguments provided above we propose the following check list. Following the checklist should help to detect path dependence, to decide whether it is problematic and to find out whether it could or should be avoided.

1. What is the main goal of the modeling process – learning or prescriptive modelling?
    - Path dependence is likely to be more problematic for prescriptive modelling.
2. Is path dependence a real risk and do we want to avoid it?
    - Discuss the possibility of path dependence and its origins with the stakeholders and the problem solving team.
    - If path dependence is unavoidable due to systemic reasons or due to unforeseeable changes in the modeling environment, acknowledge this fact when communicating the results.
3. Consider measures and risks related to the system created by the problem solving process.
    - Is the problem solving team balanced and are we aware of the risk of groupthink?
    - Consider alternative approaches when starting the problem solving process.
    - Consider adopting a value focused framework before choosing which approach is followed in the problem solving. Choosing the approach first may restrict one's creative thinking.
    - Is it possible that we have a methodological lock-in to use our favorite model or the one we are familiar with?
4. Consider procedural, behavioral and motivational factors.
    - What behavioral biases can be related to the problem solving steps and what are their impacts?
    - Consider if the risks of biases can be reduced with the redesign of the problem solving procedure.
    - Can modelers or stakeholders in the problem solving team have hidden motivational or strategic agendas?

5. Consider technical factors related to the problem under study.
    - Can nonlinearities and feedback loops in the system studied create path dependence?
    - Are there irreversibilities that need to be taken into account?
    - What are the sources of uncertainty related to the system?
    - Does the problem decomposition create a risk?
6. Is it possible to use multiple models?
    - Multi-modelling can reveal path dependencies.
    - Consider if an independent parallel problem solving process is needed.
    - Would a peer review of the process be useful?
7. Consider the possibility of an adaptive modelling approach.
    - Pay attention to the process and plan for intermediate checkpoints where the approach and available data is re-evaluated.
    - Consider if it is possible and necessary to restart the problem solving process with new data and new assumptions.

3. Biases and path dependence in Decision Analysis

Before going to the demonstration of path dependence in the Even Swaps method we consider path dependence and its possible causes in decision analysis (DA) more generally.

All decision analysis (DA) processes consist of a sequence of steps and typically different paths can be followed. Biases are likely to be an important driver for path dependence in DA because it directly involves and works with subjective data elicited from people. The sequential character of the process can lead to accumulation of effects of biases in successive statements in preference assessment.

The existence of biases and other behavioral effects in DA is well known (see, e.g., Weber et al. 1988, Weber and Borcherding 1993, Delquié 1993, 1997, Fischer 1995, Pöyhönen et al. 2001, Pöyhönen and Hämäläinen 2001, Anderson and Hobbs 2002, Carlson and Bond 2006, Jacobi and Hobbs 2007, Hämäläinen and Alaja 2008). These effects can be related to, e.g., problem framing, preference elicitation tasks, and to how information is presented. Montibeller and Winterfeldt (2015) provide a recent review of biases in decision and risk analysis. They also discuss the possibilities and needs for debiasing in decision support.

The concept of constructed preferences discussed in psychological literature (Slovic 1995, Lichtenstein and Slovic 2006) relates directly to path dependence in decision making, as noted by

Payne et al. (1999). The idea is that people do not have stable underlying preferences. Instead they construct preferences during the decision making process, and the process has an impact on the preferences that are formed. This can naturally lead to path dependence because different paths can lead to different preferences.

The possibility that the decision maker (DM) is anchored to the initial point was mentioned already by French (1984) in the context of multi-criteria optimization. In the experiment by Korhonen et al. (1990) path dependence was suggested to be caused by effects related to prospect theory.

In this paper we focus on two well-known phenomena, the scale compatibility (Tversky et al. 1988, Slovic et al. 1990) and the loss aversion biases (Tversky and Kahneman 1991). We show how they can be related to path dependence in DA. The scale compatibility bias refers to that DMs tend to give extra weight to the measuring stick used in a trade-off task (Delquié 1993, 1997, Anderson and Hobbs 2002, Bleichrodt and Pinto 2002). Measuring stick is the attribute in which the DM gives the response in the task. This bias can lead to path dependency as several different measuring sticks can be employed in the DA process. Loss aversion can create path dependency when many reference points can be used.

4. Scale compatibility and loss aversion as causes of path dependence in the Even Swaps method

4.1. The Even Swaps method and the measuring stick attribute

The Even Swaps method (Hammond et al. 1998, 1999) helps the DM to identify the 'best' alternative out of a set of multiattribute alternatives. The DM carries out a sequence of even swaps in which she changes an alternative in two attributes such that the modified alternative is preferentially equivalent to the original one. The goal is to make swaps so that alternatives become dominated and can be eliminated or so that attributes become irrelevant. The process continues until only one alternative remains. The conducted sequence of swaps forms the path of the process. The method allows the DM to choose the path freely. Ideally, she would end up with the same alternative on each path.

The DM carries out the even swap in two steps. First she selects a change in one attribute of the alternative. This we call a *reference change*. Then she gives a compensating *response change* in another attribute which we call the *measuring stick attribute*.

A straightforward strategy for carrying out the Even Swaps process, suggested by Hammond et al. (1998), is to use even swaps to repeatedly make attributes irrelevant until only one remains. At this point the most preferred alternative can be readily identified. We call this the *attribute elimination strategy*. The *pricing out* method by Keeney and Raiffa (1976) is an attribute elimination strategy in which all attributes but the monetary one are made irrelevant and money is used as the measuring stick in every swap.

4.2. Scale compatibility

It is known that people tend to give extra weight to the response attribute, i.e. the measuring stick, in two-attribute matching tasks (Delquié 1993, 1997, Anderson and Hobbs 2002, Bleichrodt and Pinto 2002). This is referred to as the scale compatibility bias. The task of determining the response change of an even swap is equivalent to giving a response in a two-attribute matching task. Therefore one can expect that the scale compatibility bias is found in a similar manner in even swaps as in matching tasks. The bias would cause the measuring stick attribute to get extra weight in the even swap. This would cause the result of an Even Swaps process to depend on the measuring stick attributes used.

When a single measuring stick attribute is used throughout the Even Swaps process, the DM repeatedly carries out even swaps in which this same attribute receives extra weight. This way the effects of the scale compatibility bias can accumulate. This leads us to the following hypothesis:
Hypothesis 1: An Even Swaps process where only one measuring stick is used favors the alternatives that are good in this measuring stick attribute.

4.3. Loss aversion

Loss aversion refers to people's tendency to give extra weight to losses compared to corresponding gains (Tversky and Kahneman 1991). Bleichrodt and Pinto (2002) show that people are loss averse in two-attribute matching tasks. Asking for the response change in an even swap task is equivalent to a two-attribute matching task. Therefore one can expect that the loss aversion bias also exists in the even swap tasks.

In the even swap task an alternative is changed in two attributes. One of these changes made in the alternative is a gain and one is a loss. A loss averse DM will give extra weight to the loss. This results in the situation where this alternative becomes more attractive in each swap. If the reference change of the even swap is a loss then the compensatory response change is a gain. In this case, the

DM overstates the response change because she gives extra weight to the reference change. If the reference change of the even swap is a gain then the compensatory response change is a loss. In this case, the DM understates the response change because she gives extra weight to it. In either case the even swap increases the attractiveness of this alternative.

When the same alternative is repeatedly swapped, then loss aversion can make this alternative better and better. This way the effects of the loss aversion bias can accumulate in favor of this alternative. This leads us to the following hypothesis:

Hypothesis 2: The Even Swaps process favors the alternative in which the most swaps are conducted.

4.4. Modeling scale compatibility and loss aversion

We present a simple approach to modeling the scale compatibility and loss aversion biases. This model is based on the Anderson and Hobbs (2002) model to estimate the magnitude of scale compatibility. We include a new loss aversion parameter in the model and assume that the value function for each attribute is linear.

The following notation is used. The reference change of an even swap in attribute k is $x_k \rightarrow x'_k$ and the response change in the measuring stick attribute m is $x_m \rightarrow x'_m$. The magnitude of the corresponding trade-off ratio is denoted by

$$r_{mk} = \left|\frac{x'_m - x_m}{x_k - x'_k}\right|. \tag{1}$$

The weights of attributes m and k are denoted by w_m and w_k. The coefficients describing the increase in weight due to biases are S and L for scale compatibility and loss aversion respectively. For unbiased DM they would equal to one. Using these notations the trade-off ratio is given in the following way.

If the reference change $x_k \rightarrow x'_k$ in attribute k is a loss:

$$r_{mk} = \frac{w_k \cdot L}{w_m \cdot S} \cdot e. \tag{2}$$

If the response change $x_m \rightarrow x'_m$ in the measuring stick attribute m is a loss:

$$r_{mk} = \frac{w_k}{w_m \cdot S \cdot L} \cdot e. \tag{3}$$

The coefficient e represents a random error which is assumed log-normally distributed with one as the median as in Anderson and Hobbs (2002).

To obtain estimates for the bias coefficients we take logarithms of (2) and (3) and get:

$$\ln(r_{mk}) = \ln(w_k) - \ln(w_m) - \ln(S) \pm \ln(L) + \ln(e), \tag{4}$$

where the sign of $\ln(L)$ depends on whether the loss is in the attribute m or in the attribute k. This model can be estimated with ordinary least squares regression.

4.5. Illustration of path dependence in Even Swaps

To help the reader understand how the loss aversion and scale compatibility biases can lead to path dependence we provide the following simple illustrative example. The initial consequences table is given in Table 1 on the left. In this illustration we assume the DM's preferences to follow a linear additive value function with equal attribute weights. For illustrative purposes, overall values of the alternatives are shown below each consequences table. These overall values would not be available in a real case because we would not know the DM's preferences.

When the DM conducts even swaps she exhibits scale compatibility and loss aversion such that her swaps follow equations (2) and (3) with bias coefficients $S = 1.3$ and $L = 1.1$. These magnitudes for bias coefficients are selected because they are close to the average magnitudes observed in our experiments. Theoretically, when S or L differs from 1 we can always construct a similar illustrative example where the outcome depends on the path followed.

When there are no biases, alternative B gives the highest value for the DM. Alternative A would contribute 75 per cent of the value of B. However, when biases are assumed the DM can end up with either alternative. That is, there exists path dependence. We show this by considering the following two paths:

Path 1: Attribute X is used as the measuring stick and all swaps are carried out in A.

Path 2: Attribute X is used as the measuring stick and all swaps are carried out in B.

On both paths the DM has to conduct two swaps in order to find a non-dominated alternative. The scale compatibility bias works in favor of A on both paths because A is better than B in attribute X which is used as the measuring stick on both paths. On each path, the loss aversion bias works in favor of the alternative in which all swaps carried out. When the DM goes along path 1 she ends up with alternative A (Table 1). When the DM goes along path 2 she ends up with alternative B (Table 2).

Table 1: Consequences tables on path 1.

Initial consequences table.

	Alternatives	
Attributes	A	B
X	2	0
Y	1	0
Z	0	4
Overall value	3	4

Consequences after swap 1. Attribute Z becomes irrelevant.

	Alternatives	
Attributes	A	B
X	-0.8	0
Y	1	0
~~Z~~	~~4~~	~~4~~
Overall value	4.2	4

Consequences after swap 2. Alternative B becomes dominated.

	Alternatives	
Attributes	A	B
X	0.05	0
~~Y~~	~~0~~	~~0~~
~~Z~~	~~4~~	~~4~~
Overall value	4.05	4

Swaps on path 1 explained:

Swap 1. Reference change: Alternative A improved in the attribute Z from 0 to 4.

Response change: Alternative A worsened in the measuring stick attribute X from 2 to α.

The response change is a loss so α is based on equation 3:

$$\left|\frac{\alpha-2}{0-4}\right| = \frac{1}{1\cdot 1.3\cdot 1.1} \Leftrightarrow \alpha = -4\cdot\frac{1}{1\cdot 1.3\cdot 1.1}+2 \Leftrightarrow \alpha \approx -0.80.$$

Swap 2. Reference change: Alternative A worsened in the attribute Y from 1 to 0.

Response change: Alternative A improved in the measuring stick attribute X from -0.8 to β.

The reference change is a loss so β is based on equation 2:

$$\left|\frac{\beta-(-0.80)}{1-0}\right| = \frac{1\cdot 1.1}{1\cdot 1.3} \Leftrightarrow \beta = 1\cdot\frac{1\cdot 1.1}{1\cdot 1.3} - 0.80 \Leftrightarrow \beta \approx 0.05.$$

Table 2: Consequences tables on path 2.

Initial consequences table

	Alternatives	
Attributes	A	B
X	2	0
Y	1	0
Z	0	4
Overall value	3	4

Consequences after swap 1. Attribute Z becomes irrelevant.

	Alternatives	
Attributes	A	B
X	2	3.4
Y	1	0
~~Z~~	~~0~~	~~0~~
Overall value	3	3.4

Consequences after swap 2. Alternative A becomes dominated.

	Alternatives	
Attributes	A	B
X	2	2.7
~~Y~~	~~1~~	~~1~~
~~Z~~	~~0~~	~~0~~
Overall value	3	3.7

5. Experiment

The subjects were Finnish speaking, mostly second year engineering students (N=148) from Aalto University. The subjects used the Even Swaps method with the Smart-Swaps software (Hämäläinen et al. 2004, Mustajoki and Hämäläinen 2005, 2007) in decision tasks related to selecting a rental apartment or a summer job.

Each session consisted of two steps. First the subjects were given a 15 minute tutorial on the method and they practiced using it for 10 minutes. Then the subjects were given a sheet that contained the instructions for carrying out the experiment. The subjects proceeded at their own pace. Completing the tasks took from half an hour to one and a half hours. Incomplete responses and those which do not follow the instructions are excluded from analysis.

5.1. The decision tasks

Four different decision tasks are used in the experiment. The consequences tables for each of them are given in Tables 3-6. The subjects were told to think that the alternatives are equally good with respect to any other attribute that is not shown in the table. The 1 to 3 scales used for the attributes condition, atmosphere and flexibility are from the worst to the best. These scales were described to the subjects in more detail on the instruction sheets of the experiment.

Table 3: Consequences table for Task 1 (job)

	Alternatives			
Attributes	A	B	C	D
Salary	2600€	1850€	2800€	2100€
Daily working hours	7.5 h	9 h	8.5 h	7 h
Job atmosphere	2	3	1	2
Commuting time	60 min	45 min	30 min	35 min
Flexibility	1	3	1	2

Table 4: Consequences table for Task 2 (apartment)

	Alternatives			
Attributes	A	B	C	D
Size	25m^2	27m^2	20m^2	32m^2
Commuting time	40 min	5 min	15 min	25 min
Rent	300€	450€	350€	500€
Condition	3	1	2	3

Table 5: Consequences table for Task 3 (job)

	Alternatives			
Attributes	A	B	C	D
Salary	2600€	1850€	2800€	2100€
Daily working hours	7.5 h	9 h	8.5 h	8 h
Job atmosphere	2	3	1	2
Commuting time	60 min	45 min	30 min	35 min
Flexibility	1	3	1	2

Table 6: Consequences table for Task 4 (apartment)

	Alternatives	
Attributes	B	D
Size	27m^2	32m^2
Commuting time	5 min	25 min
Rent	450€	500€
Condition	1	3

5.2. Paths

Each subject carried out one apartment related task and one job related task. Both tasks were carried out two or three times. Each time the subject followed a different path, i.e., conducted a different sequence of swaps. The different paths resulted from different instructions given (Table 7).

The Pricing (PRI) and Hours (HRS) paths are based on the use of the attribute elimination strategy. The dominance (DOM) and irrelevance (IRR) paths are used as references for these paths. The DOM and IRR paths are based on the use of the Even Swap proposal features of the Smart-Swaps software. The software allows the user to ask for proposals which help her to complete the Even Swaps process with as few swaps as possible (Mustajoki and Hämäläinen 2005).

- The proposals for dominance aims at domination of alternatives with as few swaps as possible.
- The proposals for irrelevance aims at making attributes irrelevant with as few swaps as possible.

The proposal specifies the alternative that would be modified, the reference change and the measuring stick attribute. The user gives the response change based on her preferences. Figure 1 illustrates a proposal for dominance.

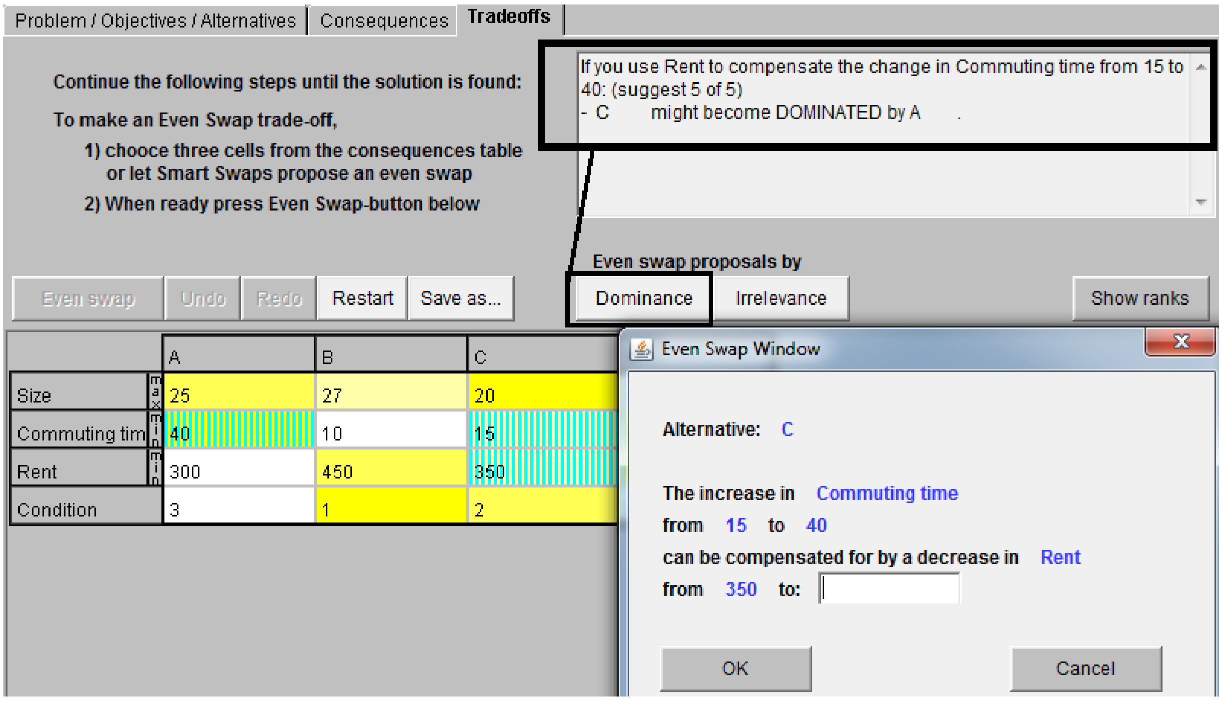

Figure 1: Interface for the Smart-Swaps software (Hämäläinen et al. 2004). Even swap proposal by dominance is highlighted with the bolded frame.

Table 7: Paths related to different instructions.

Path	Subjects instructed to
Pricing (PRI)	use an attribute elimination strategy to make all attributes but the monetary one irrelevant. Money used as the measuring stick in every swap.
Hours (HRS)	use an attribute elimination strategy to make all attributes but daily hours irrelevant. Daily hours used as the measuring stick in every swap.
Dominance (DOM)	follow the suggestions provided by the feature 'even swap proposals by *dominance*' of the Smart-Swaps software.
Irrelevance (IRR)	follow the suggestions provided by the feature 'even swap proposals by *irrelevance*' of the Smart-Swaps software.
Swaps in one alternative (Swaps in B/D)	carry out all swaps in the same alternative, B or D.*

*=This instruction was used only in Task 4 which includes only two alternatives.

The subjects were divided into two groups with different tasks and instructions (Table 8). Tasks 1 and 2 were used with the first group, tasks 3 and 4 were used with the second group. The subjects repeatedly carried out a job related task followed by an apartment related task. The order in which the subjects followed the paths was altered across subjects.

Table 8: The experiments

Task	N	Paths		
Task 1 (job)	98	PRI	DOM	IRR
Task 2 (apartment)		PRI	DOM	IRR
Task 3 (job)	50	HRS	DOM	-
Task 4 (apartment)		Swaps in B	Swaps in D	-

6. Results

6.1. Path dependence

Path dependence is studied by comparing the subjects' choices across the paths. Path dependence exists if different paths lead to different final alternatives. We can only compare the results between paths but cannot determine which the 'true' most desirable alternative is for each subject.

For the analysis the alternatives in each task are grouped into two sets (Table 9). In Task 1 and Task 2 the alternatives are grouped based on their performance in the money attribute which is used as the measuring stick on the PRI path. In Task 3 the alternatives are grouped based on their performance in daily working hours which is used as the measuring stick on the HRS path.

Table 9: Sets of alternatives

Task	Set 1	Set 2
Task 1 (job)	'high-salary jobs' A and C	'low-salary jobs' B and D
Task 2 (apartment)	'low-rent apartments' A and C	'high-rent apartments' B and D
Task 3 (job)	'low-hours jobs' A and D	'high-hours jobs' B and C
Task 4 (apartment)	alternative B	alternative D

Hypothesis 1 predicts that the percentage of subjects ending up with any of the alternatives in Set 1 is higher on the PRI path than on the IRR and DOM paths. Hypothesis 1 also predicts that the percentage of subjects ending up with any of the alternatives in Set 1 is higher on the HRS path than on DOM path. The reason is that only one measuring stick attribute is used on PRI and HRS paths and the alternatives in Set 1 are better in this attribute.

Hypothesis 2 predicts that the percentage of subjects ending up with alternative B is higher on the Swaps in B path than on the Swaps in D path. The reason is that on the Swaps in B path all swaps are conducted in alternative B.

Table 10 (column 5) gives the percentage of subjects ending up with an alternative in Set 1 on each path. This illustrates the effect sizes. Additionally, we tested whether these results change if only subjects with data from all paths are included for each task. We found the same pattern of results. None of our conclusions would change if that data was used.

The statistical test results are presented in Table 11. Statistical McNemar's test (1947) is used because we have binomial paired data. The test is based on comparing the results of the *same* subjects between a pair of paths. Only subjects with data from both paths are included in each comparison. The null distribution of the test statistic k is binomial(K, 0.5) where K is the number of subjects who end up with alternatives in different sets on each of the two paths. The statistic *k* is the number of subjects who end up with any of the alternatives in the Set 1 on the PRI, HRS or Swaps in B path *and* end up with any of the alternatives in the Set 2 in the other path.

Table 10: Number and percentage of subjects who ended up with an alternative in Set 1.

Task	Path	N	Number of subjects who ended up with	Percentage of subjects who ended up with	Percentage of swaps with money as measuring stick
Task 1 (job)			high-salary job	high-salary job	
	PRI	67	42	63	100
	IRR	98	56	57	55
	DOM	98	28	29	19
Task 2 (apartment)			low-rent apartment	low-rent apartment	
	PRI	45	36	80	100
	IRR	97	61	63	34
	DOM	96	51	53	31
Task 3 (job)			low-hours jobs	low-hours jobs	
	HRS	45	34	76	N/A†
	DOM	45	32	71	N/A
			apartment B	apartment B	
Task 4 (apartment)	Swaps in B	38	19	50	N/A
	Swaps in D	38	8	21	N/A

† N/A=Not available

Table 11: Results of statistical tests.

Task	Paths	Prediction by hypothesis 1 or 2	N	K	k	p-value
Task 1 (job)	PRI, IRR	Hypothesis 1: PRI favors high-salary alternatives	67	18	11	0.24
	PRI, DOM	Hypothesis 1: PRI favors high-salary alternatives	67	34	28	0.0004***
Task 2 (apartment)	PRI, IRR	Hypothesis 1: PRI favors low-rent alternatives	45	18	13	0.05*
	PRI, DOM	Hypothesis 1: PRI favors low-rent alternatives	45	20	18	0.0002***
Task 3 (job)	HRS, DOM	Hypothesis 1: HRS favors low-hours alternatives	45	16	9	0.4
Task 4 (apartment)	Swaps in B, Swaps in D	Hypothesis 2: Swaps in B favors alternative B	38	15	13	0.004**

*=Statistical significance level p <0.05, **=p<0.01, ***=p<0.001.

The difference between two paths is statistically significant with p<0.001 in two cases, p<0.01 in one case and p<0.05 in one case (Table 11). In the remaining two cases the difference between the paths is small but in the direction predicted by the hypothesis.

In addition, Table 10 shows that the subjects used money as the measuring stick more frequently on the IRR than on the DOM path. This possibly explains why the IRR path has been more favorable than the DOM path for the alternatives that are good in the monetary attribute. We also compare the choices by subjects who conducted Task 1 (job) on the PRI path and the choices by subjects who conducted Task 3 (job) on the HRS path. Two statistically significant differences are found. Job C is chosen more often on the PRI path (40%) than on the HRS path (11%) (Z-statistic: 3.35, p-value: 0.0004). Job D is chosen more often on the HRS path (53%) than on the PRI path (36%) (Z-statistic:-1.84, p-value: 0.03). These results are in line with hypothesis 1 because Job C has better salary and Job D has lower working hours.

We conclude that these results show the existence of path dependence. Hypothesis 1 is clearly supported when a monetary attribute is used as the measuring stick. Hypothesis 2 is supported.

6.2. Magnitudes of loss aversion and scale compatibility

Altogether the subjects carried out several thousand even swap tasks during the experiment. We form eight data sets out of this data and estimate the model (4) separately for each of them. The data set *Apartment task – All attributes* includes all swaps conducted in a task related to choice of an apartment. The data set *Job task – All attributes* includes all swaps conducted in a task related to choice of a job. The rest of the data sets are formed such that each contains all swaps where the same pair of attributes was traded-off against each other. Only those pairs of attributes are studied in which the measurement scale for both attributes is continuous. Table 12 shows the estimates of the bias coefficients.

Table 12: Estimates of magnitudes of biases for different data sets.

Data set	N	R^2	ln(L)	p-value	ln(S)	p-value	Loss aversion coefficient L	Scale compatibility coefficient S
Apartment task								
All attributes	2155	0.95	0.1	3.6E-13	0.19	2.7E-31	1.11	1.21
Rent, Commuting time	474	0.89	0.078	0.12	0.36	3.9E-10	1.08	1.43
Size, Commuting time	419	0.52	0.12	0.0048	0.16	1.6E-5	1.13	1.17
Size, Rent	554	0.97	0.058	0.0043	0.11	8.1E-7	1.06	1.12
Job task								
All attributes	3096	0.98	0.12	3.2E-23	0.29	1.1E-90	1.13	1.34
Salary, Commuting time	576	0.95	0.074	0.12	0.33	1.1E-11	1.08	1.39
Working hours, Commuting time	305	0.95	0.15	0.0030	0.3	5.1E-9	1.16	1.35
Working hours, Salary	592	0.99	0.053	0.10	0.21	6.5E-12	1.05	1.23

We make the following three observations.

Observation 1. The scale compatibility and loss aversion biases are found in even swaps similarly as in matching tasks. The bias coefficients S and L are greater than one in all data sets. This confirms our initial assumption that these biases exist in Even Swaps.

Observation 2. The estimates of magnitude of scale compatibility obtained for data sets "Apartment task – All attributes" and "Job task – All attributes" are 1.21 and 1.34 respectively. These are close to the average of the subject specific scale compatibility estimates by Anderson and Hobbs (2002) which is 1.32 when their subject 4 is excluded as an outlier. This suggests the interesting conclusion that the average magnitude of scale compatibility can be a general bias effect which is of the same magnitude in different contexts. The experiment by Anderson and Hobbs (2002) was in the context of fisheries management and the subjects were managers instead of students. Anderson and Hobbs (2002) report Bayesian posterior probability distributions of bias for each subject. The means of these distributions are (1.5, 2.07, 1.01, 10.9, 1.25, 1.45, 0.65) (R. Anderson, personal communication, September 10, 2013).

Observation 3. The estimates obtained for S range from 1.12 to 1.43 and the estimates obtained for L range from 1.05 to 1.16. A possible explanation is that the magnitude of the bias depends on which attributes are traded-off against each other. This would be contrary to the assumption by Anderson and Hobbs (2002) that the bias coefficient *do not* depend on which attributes are traded-off against each other. This matter could be investigated in a further study.

We conducted residual analysis that suggests the assumption of homogeneity of variance and the assumption of normality to be reasonable. Multicollinearity is not found in the data sets.

6.3. Summary of results

When the subjects go through the Even Swaps process and use money as the measuring stick in the even swap tasks, i.e. they give responses in money, they end up favoring those alternatives which are good in the monetary attribute. When two alternatives are compared such that the same alternative is modified in every swap, the subjects favor the modified alternative. These results can be explained with the accumulated effect of successive statements biased by scale compatibility and loss aversion. These biases are also found in the trade-off data provided by the experiments.

7. Reducing the accumulation of bias effects in Even Swaps

In applications of decision analysis one should be aware of the risk of path dependence. For example, one might think that the straightforward pricing out method, which corresponds to the pricing path in Even Swaps, is always safe to use. This method can, however, favor the alternatives which are good in the monetary attribute. Below we suggest a strategy for carrying out the Even Swaps process that leads to a path that mitigates the cumulative effect of the biases.

To reduce the impact of the loss aversion bias, the DM can carry out even swaps evenly in all of the alternatives. This way the effects of the bias do not accumulate in favor of a single alternative but are distributed evenly across the alternatives. Effect of the scale compatibility bias can be reduced by using a measuring stick in which the consequences of the alternative differ the least. Scale compatibility gives extra weight to the measuring stick attribute. This extra weight does not matter if the alternatives differ only a little in this attribute.

These ideas can be illustrated using the example of section 4.5. The attribute Y should be selected as the measuring stick because the alternatives' consequences differ the least in this attribute. One swap should be conducted in each of the alternatives to distribute the effect of loss aversion evenly. Following these suggestions would lead to a path shown in Table 13. Along this path the DM would end up with alternative B which theoretically gives her the highest value.

Table 13: Illustration of path where cumulative effect of bias is reduced.

Initial consequences table.

	Alternatives	
Attributes	A	B
X	2	0
Y	1	0
Z	0	4
Overall value	3	4

Consequences after step 1. Attribute X becomes irrelevant.

	Alternatives	
Attributes	A	B
~~X~~	~~0~~	~~0~~
Y	2.7	0
Z	0	4
Overall value	2.7	4

Consequences after step 2. Alternative A becomes dominated.

	Alternatives	
Attributes	~~A~~	B
~~X~~	~~0~~	~~0~~
Y	~~2.7~~	3.4
~~Z~~	~~0~~	~~0~~
Overall value	2.7	3.4

Another idea to reduce the accumulated effect of biases is to restart the Even Swaps process with the original consequences of the remaining alternatives at certain points during the process. This would remove any error or bias that has accumulated in the remaining alternatives. One possibility would be to restart the process with the remaining alternatives every time when an alternative is eliminated. The following example illustrates this idea: The decision task includes three alternatives A, B and C. The DM conducts four swaps in B which result in the elimination of C. Here loss aversion has accumulated and made the modified B more attractive than the original B because all swaps were conducted in B. Thus the DM is biased towards B if she continues the Even Swaps process with A and modified B. The accumulated bias could be removed by restarting the Even Swaps process with alternative A and the original, unmodified, alternative B.

This idea can be applied in the Smart-Swaps software by using the backtracking feature which allows the DM to cancel previously made swaps. Every time an alternative is dominated the DM can backtrack to the start of the process and then eliminate the dominated alternative and continue with the remaining ones.

One can utilize all of the above mentioned ideas together with the following strategy for carrying out the Even Swaps process:

1. Identify a pair of alternatives which are close to each other in some attribute that can be used as the measuring stick attribute.
2. Carry out swaps to make the consequences of these two alternatives equal in all other attributes. Alternate between conducting swaps in the two alternatives.
3. Eliminate the dominated alternative and restart the Even Swaps process with the reduced set of alternatives.
4. Continue until only one alternative remains.

Using the strategy described above will take at most (N-1)(K-1) swaps where N is the number of alternatives and K is the number of attributes. One needs N-1 pairwise comparisons to eliminate all but one alternative. In each pairwise comparison at most K-1 swaps must be conducted to make all but one attribute irrelevant. In the attribute elimination method, e.g. the pricing out method, the upper limit for the number of swaps is the same. The strategy presented here averages out the effect of biases as suggested by Kleinmuntz (1990).

It should be noted, however, that averaging responses to different preference elicitation questions does not always lead to the most accurate judgment. For example, a decision maker might give the most thoughtful responses when using money as the measuring stick if she is used to evaluating benefits in terms of money. In such a case using money as the measuring stick might lead to more accurate judgments than using several measuring sticks and averaging the responses. Therefore, averaging and debiasing procedures (see, e.g., Anderson and Hobbs 2002 or Jacobi and Hobbs 2007) should be employed with careful consideration.

One could also think that sensitivity analysis can help to identify and avoid path dependence in Even Swaps. However, performing sensitivity analysis in a ‘traditional way’ is difficult in the Even Swaps method. This is because of the sequential nature of the swapping process. For example, changing the response change in the first swap by 10% might cause the subsequent swaps to be unfeasible. Thus the DM might have to revise all swaps accordingly. Instead, as a form of sensitivity analysis, one could carry out the whole Even Swaps process along multiple paths and compare the results.

8. Conclusions

The possibility that two valid modeling processes can lead to different results was recognized early (Landry et al. 1983). However, this issue has later received very little attention. In this paper we point out the relevance of this issue in particular in policy related OR processes and in decision analysis. We suggest that it is useful to consider the related phenomena under the integrative term, path dependence. We discuss seven, possibly interplaying, origins or drivers of path dependence in OR: system, learning, procedure, behavior and motivation, as well as uncertainty and external origins. Thinking of these perspectives the practitioner should be better able to identify path dependence and find ways to analyze whether there is possibility and need to avoid it.

We claim that acknowledging the possibility of path dependence challenges us to critically evaluate and improve our OR practices. Increased awareness is the natural first step. The practitioner can, for example, consider the check list provided in this paper. The adaptive approach is one natural way to mitigate the effects of path dependence. In this approach the process is revised at checkpoints, where intermediate results are obtained, learning has occurred, and possibly new data has become available. We suggest that good practices in OR should always include the idea of adaptive modelling.

In decision analysis path dependence is likely to emerge from behavioral origins because DA directly involves and works with subjective data elicited from people. Our experiment demonstrates how path dependence emerges in the Even Swaps method. The suggested causes for our observations are the scale compatibility and loss aversion biases. The issue of path dependence is important to consider in normative decision support where the aim is to give one correct outcome to be implemented. One option is to analyze the same problem following different paths. Debiasing methods can also be considered. This type of remedy for path dependence is demonstrated by the bias averaging strategy for the Even Swaps method developed in this paper.

Today we should consider the OR process and the people engaged in the system created in the problem solving as a whole. The essential question is to understand the human impact on the whole OR process. This naturally leads us to consider the path followed in the process. Future research should consider the human drivers of path dependence in more detail in different situations and in different modeling processes. These drivers can be related to the OR facilitator and to the problem owners as well as to the stakeholders involved.

References

Ackermann, F., 2012. Problem structuring methods 'in the Dock': Arguing the case for soft OR. European Journal of Operational Research 219 (3), 652-658.

Anderson, R.M., Hobbs, B.F., 2002. Using a Bayesian Approach to Quantify Scale Compatibility Bias. Management Science 48 (12), 1555-1568.

Araña, J.E., León, C.J., 2008. Do emotions matter? Coherent preferences under anchoring and emotional effects. Ecological Economics 66 (4), 700-711.

Arkes, H.R., Blumer, C., 1985. The psychology of sunk cost. Organizational behavior and human decision processes 35 (1), 124-140.

Arrow, K.J., Fisher, A.C., 1974. Environmental preservation, uncertainty, and irreversibility. The Quarterly Journal of Economics 9 (1), 312-319.

Arthur, W.B., 1989. Competing technologies, increasing returns, and lock-in by historical events. The Economic Journal 99 (394), 116-131.

Barnhart, C., Belobaba, P., Odoni, A. R., 2003. Applications of operations research in the air transport industry. Transportation science 37 (4), 368-391.

Barnes, W., Gartland, M., Stack, M., 2004. Old Habits Die Hard: Path Dependency and Behavioral Lock-in. Journal of Economic Issues 38 (2), 371-377.

Bleichrodt, H., Pinto, J.L., 2002. Loss Aversion and Scale Compatibility in Two-Attribute Trade-Offs. Journal of Mathematical Psychology 46 (3), 315-337.

Bond, S.D., Carlson, K.A., Keeney, R.L., 2008. Generating Objectives: Can Decision Makers Articulate What They Want? Management Science 54 (1), 56-70.

Buchanan, J.T., Corner, J.L., 1997. The Effects of Anchoring in Interactive MCDM Solution Methods. Computers & Operations Research 24 (10), 907-918.

Carlson, K.A., Bond, S.D., 2006. Improving Preference Assessment: Limiting the Effect of Context through Pre-exposure to Attribute Levels. Management Science 52 (3), 410-421.

Checkland, P., 2000. Soft systems methodology: a thirty year retrospective. Systems Research and Behavioral Science 17, S11-S58.

Clark, C.W, Mangel, M., 1979. Aggregation and fishery dynamics: a theoretical study of schooling and the purse seine tuna fisheries. Fishery Bulletin 77 (2), 317-337.

Clark, C.W. 1990. Mathematical Bioeconomics: The Optimal Management of Renewable Resources. Wiley Interscience, New York.

David, P.A., 1985. Clio and the Economics of QWERTY. The American Economic Review 75 (2), 332-337.

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.

Derksen, S., Keselman, H.J., 1992. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. British Journal of Mathematical and Statistical Psychology 45 (2), 265-282.

van Exel, N.J.A., Brouwer, W.B.F., van der Berg, B., Koomanschap, M.A., 2006. With a little help from an anchor: Discussion and evidence of anchoring effects in contingent valuation. The Journal of Socio-Economics 35 (5), 836-853.

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., Lord, E., 2011. Understanding multi-methodology: evaluating the perceived impact of mixing methods for group budgetary decisions. Omega 39 (3), 362-372.

French, S., 1984. Interactive Multi-Objective Programming: Its Aims, Applications and Demands. The Journal of the Operational Research Society 35 (9), 827-834.

Furnham, A., Boo, H.C., 2011. A literature review of the anchoring effect. The Journal of Socio-Economics 40 (1), 35-42.

Gartland, M.P., 2005. Interdisciplinary views of sub-optimal outcomes: Path dependence in the social and management sciences. The Journal of Socio-Economics 34 (5), 686-702.

George, J.F., Duffy, K., Ahuja, M., 2000. Countering the anchoring and adjustment bias with decision support systems. Decision Support Systems 29 (2), 195-206.

Hammond, J.S., Keeney, R.L., Raiffa, H., 1998. Even swaps: A rational method for making trade-offs. Harvard Business Review 76 (2), 137-149.

Hammond, J.S., Keeney, R.L., Raiffa, H., 1999. Smart Choices: A Practical Guide to Making Better Decisions. Harward Business School Press, Boston, MA.

Hämäläinen, R., Kettunen, E., Marttunen, M., Ehtamo, H., 2001. Evaluating a framework for multi-stakeholder decision support in water resources management. Group Decision and Negotiation 10 (4), 331-353.

Hämäläinen, R.P., Mustajoki, J., Alanaatu, P., Karttunen, V., Arstila, A., Nissinen, J., 2004. Smart-Swaps – Smart Choices with Even Swaps. Computer software, Systems Analysis Laboratory, Aalto University, http://smart-swaps.aalto.fi/ (url checked 05.06.2015).

Hämäläinen, R.P., Alaja, S., 2008. The threat of weighting biases in environmental decision analysis. Ecological Economics 68 (1), 556-569.

Hämäläinen, R.P., Luoma, J., 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., 2015. Behavioral modelling research – the missing perspective. Manuscript.

Jacobi, S.K., Hobbs, B.F., 2007. Quantifying and mitigating the splitting bias and other value tree-induced weighting biases. Decision Analysis 4 (4), 194-210.

Janis, I.L., 1982. Groupthink: Psychological Studies of Policy Decisions and Fiascoes. Wadsworth, USA.

Jun, J.B., Jacobson, S.H., Swisher, J.R., 1999. Application of Discrete-Event Simulation in Health Care Clinics: A Survey. The Journal of the Operational Research Society 50 (2), 109-123.

Keeney, R.L., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. John Wiley & Sons, New York.

Keeney, R.L., 1994. Creativity in decision making with value-focused thinking. Sloan Management Review 35, 33-41.

Kleinmuntz, D.K., 1990. Decomposition and control of error in decision-analytic model. In Robin M. Hogarth (ed.) Insights in Decision Making: A Tribute to Hillel J. Einhorn, Unviersity of Chicago Press, Chicago, 107-126.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G. A., 2010. Challenges in combining projections from multiple climate models. Journal of Climate 23 (10), 2739-2758.

Korhonen, P., Moskowitz, H., Wallenius, J., 1990. Choice Behavior in Interactive Multiple-criteria Decision Making. Annals of Operations Research 23(1), 161-179.

Kristensen, H., Gärling, T., 1997. The Effects of Anchor Points and Reference Points on Negotiation Process and Outcome. Organizational Behavior and Human Decision Processes 71 (1), 85-94.

Landry, M., Malouin, J-L., Oral, M. 1983. Model validation in operations research. European Journal of Operational Research 14 (3), 207-220.

Lane, D.C., 1992. Modelling as learning: A consultancy methodology for enhancing learning in management teams. European Journal of Operational Research 59 (1), 64-84.

Lane, D.C., 1993. The road not taken: observing a process of issue selection and model conceptualization. System Dynamics Review 9 (3), 239-264.

Lee, S.M., Olson, M.D., 1980. The challenge of operations research in the developing country. The Journal of Technology Transfer 4 (2), 37-45.

Lempert, R.J., Groves, D.G., 2010. Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west. Technological Forecasting & Social Change 77 (6), 960-974.

Lenert, L.A., Cher, D.J., Goldstein, M.K., Bergen, M.R., Garber, A., 1998. The Effect of Search Procedures on Utility Elicitations. Medical Decision Making 18 (1), 76-83.

Leon, O.G., 1999. Value-focused thinking versus alternative-focused thinking: Effects on generation of objectives. Organizational Behavior and Human Decision Processes 80 (3), 213-227.

Levin, K., Cashore, B., Bernstein, S., Auld, G., 2012. Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change. Policy Sciences 45 (2), 123-152.

Lichtenstein, S., Slovic, P. (eds.) 2006. The construction of preferences. Cambridge University Press, New York.

Lieberman, M.B., Montgomery, D.B., 1988. First-mover advantages. Strategic Management Journal 9 (1), 41-58.

Little, J.D.C., 1970. Models and Managers: The Concept of Decision Calculus. Management Science 16 (8), B466-B485. Reprinted in Management Science 50 (12 Supplement), 1841-1853.

Mahoney, J., 2000. Path dependence in historical sociology. Theory and Society 29 (4), 507-548.

Martin, R., Sunley, P., 2006. Path dependence and regional economic evolution. Journal of Economic Geography 6 (4), 395-437.

McDaniels, T.L., Gregory, R., 2004. Learning as an objective within a structured risk management decision process. Environmental Science & Technology 38 (7), 1921-1926.

McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12 (2), 153-157.

Montibeller, G., von Winterfeldt, D., 2015. Cognitive and Motivational Biases in Decision and Risk Analysis. Risk Analysis. DOI: 10.1111/risa.12360

Morris, W. T., 1967. On the Art of Modeling. Management Science 13 (12), B707-B717.

Mustajoki, J., 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., 2007. Smart-Swaps - A decision support system for multicriteria decision analysis with the even swaps method. Decision Support Systems 44 (1), 313-325.

Nickerson, R.S., 1998. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology 2 (2), 175-220.

Ormerod, R. J., 1997. The Design of Organisational Intervention: Choosing the Approach. Omega 25 (4): 415-435.

Ormerod, R. J., 2008. The transformation competence perspective. Journal of the Operational Research Society 59 (11), 1435-1448.s

Ormerod, R. J., Ulrich, W., 2013. Operational research and ethics: A literature review. European Journal of Operational Research 228 (2), 291-307.

Payne, J.W., Bettman, J.R., Schkade, D.A., 1999. Measuring Constructed Preferences: Towards a Building Code. Journal of Risk and Uncertainty 19 (1-3), 243-270.

Pidd, M., 1999. Just Modeling Through: A Rough Guide to Modeling. Interfaces 29 (2), 118-132.

Pierson, P., 2000. Increasing returns, path dependence, and the study of politics. American Political Science Review 94 (2), 251-267.

Pöyhönen, M., Vrolijk, H., Hämäläinen, R.P., 2001. Behavioral and procedural consequences of structural variation in value trees. European Journal of Operational Research 134 (1), 216-227.

Pöyhönen, M., Hämäläinen, R.P., 2001. On the convergerence of multiattribute weighting methods. European Journal of Operational Research 129 (3), 569-585.

Raiffa, H., 1982. The art and science of negotiation. Harvard University Press.

Rauschmayer, F., Kavathatzopoulos, I., Kunsch, P.L., Menestrel, M. Le., 2009. Why good pratice of OR is not enough – Ethical challenges for the OR practitioner. Omega 37 (6), 1089-1099.

Rosenhead, J., 1990. Rational Analysis: Keeping Your Options Open. In J. Rosenhead, J. Mingers (eds.) Rational Analysis for a Problematic World: Problem Structuring Methods for Complexity, Uncertainty and Conflict, John Wiley & Sons, Chichester.

Salo, A., Hämäläinen, R.P., 2010. Multicriteria Decision Analysis in Group Decision Processes. Handbook of Multicriteria Decision Analysis, Springer, New York, 167-187.

Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. Journal of Risk and Uncertainty 1 (1), 7-59.

Selart, M., Johansen, S.T., 2011. Understanding the Role of Value-Focused Thinking in Idea Management. Creativity and Innovation Management 20 (3), 196-206.

Scarf, P.A., 1997. On the application of mathematical models in maintenance. European Journal of Operational Research 99 (3), 493-506.

Schoemaker, P.J.H., 1991. When and How to Use Scenario Planning: A Heuristic Approach with Illustration. Journal of Forecasting, 10 (6), 549-564.

Slovic, P., Griffin, D., Tversky, A., 1990. Compatibility effects in judgment and choice. In Robin M. Hogarth (ed.) Insights in Decision Making: A Tribute to Hillel J. Einhorn, University of Chicago Press, Chicago, 5-27.

Slovic, P., 1995. The construction of preference. American Psychologist 50 (5), 364-371.

Sterman, J.D., 1994. Learning in and about complex systems. System Dynamics Review 10 (2-3), 291-330.

Sterman, J.D., Wittenberg, J., 1999. Path Dependence, Competition, and Succession in the Dynamics of Scientific Revolution. Organization Science 10 (3), 322-341.

Sterman, J.D., 2000. Business dynamics: systems thinking and modeling for a complex world. Boston: McGraw-Hill.

Sydow, J., Schreyögg, G., Koch, J., 2009. Organizational Path Dependence: Opening The Black Box. Academy of Management Review 34 (4), 689-709.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science 185 (4157), 1124-1131.

Tversky, A., Sattath, S., Slovic, P., 1988. Contingent Weighting in Judgment and Choice. Psychological Review 94 (3), 371-384.

Tversky, A., Kahneman, D., 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. Quarterly Journal of Economics 106 (4), 1039-1061.

Varian, H.R., 1992. Microenonomic Analysis. Third edition. Norton: New York.

Walker, W.E., 2009. Does the best practice of rational-style model-based policy analysis already include ethical considerations? Omega 37 (6), 1051-1062.

Weber, M., Eisenführ, F., von Winterfeldt, D., 1988. The effects of splitting attributes on weights in multiattribute utility measurement. Management Science 34 (4), 431-445.

Weber, M., Borcherding, K., 1993. Behavioural influence on weight judgments in multiattribute decision making. European Journal of Operational Research 67 (1), 1-12.

Webster, M., Forest, C., Reilly, J., Babiker, M., Kicklighter, D., Mayer, M., Prinn, R., Sarofim, M., Sokolov, A., Stone, P., Wang, C., 2003. Uncertainty analysis of climate change and policy response. Climatic Change 61 (3), 295-320.

Webster, M., 2008. Incorporating Path Dependency into Decision-Analytic Methods: An Application to Global Climate-Change Policy. Decision Analysis 5 (2), 60-75.