Model-structuring
in
public decision-aiding

by

Carlos A. Bana e Costa[1] and Euro Beinat[2]

[1] *Centre for Management Studies of Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal*
and
Department of Operational Research, London School of Economics, Houghton Street, London WC2A 2AE, U.K. (c.bana@lse.ac.uk)

[2] *SPINlab - Spatial Information Laboratory, Vrije Universiteit Amsterdam, 1081HV, The Netherlands*
and
Salzburg University, Kapitelgasse 4-6, A-5020 Salzburg, Austria
(euro@beinat.net)

Working Paper LSEOR 05.79 *ISBN 0 7530 1701 6*

First published in Great Britain in 2005
by the Department of Operational Research
London School of Economics and Political Science

Typeset, printed and bound by:

The London School of Economics and Political Science
Houghton Street
London WC2A 2A

Working Paper LSEOR 05.79 *ISBN 0 7530 1701 6*

Model-structuring in public decision-aiding

Carlos A. ***Bana e Costa***[1]*, Euro Beinat*[2]

[1] *Centre for Management Studies of Instituto Superior Técnico, Technical University of Lisbon,*
Av. Rovisco Pais, 1049-001 Lisbon, Portugal and
Department of Operational Research, London School of Economics and Political Science,
Houghton Street, London WC2A 2AE, U.K. (c.bana@lse.ac.uk)

[2] *SPINlab - Spatial Information Laboratory, Vrije Universiteit Amsterdam, 1081HV, The Netherlands and Salzburg University, Kapitelgasse 4-6, A-5020 Salzburg, Austria*
(euro@beinat.net)

Abstract. This paper addresses structuring of multicriteria decision aid models. Structuring provides the actors involved in a decision context with a common language for debating and arguing about their preferences and facilitates the identification of decision opportunities and the construction of new alternatives. Structuring is decomposed into (1) problem-definition, (2) model-structuring and (3) impact assessment and analysis. The concept of policy unit is used to generalise the concept of decision actor and to introduce group, individual and mixed structuring frameworks. The notion of concern is central to structuring activities and is extensively discussed. Descriptors are then defined and classified in terms of natural, proxy and constructed descriptors. Several methods that can be used to select or construct descriptors are presented. The paper also provides extensive examples of good practice and operational tips for the structuring phase.

Keywords: *structuring, multicriteria decision aid, concern, descriptor, best practice.*

1 Introduction

The subject of this paper is the structuring of a multicriteria decision aid model. It emphasises the complexity involved in approaching ill-structure public decision-making problems involving multiple points of view. We argue that structuring should provide the actors involved in a decision context with a common language for debating and arguing about their preferences and for learning from each one perspective. It should facilitate and stimulate the identification of decision opportunities and the construction of new alternatives. Technically, the structuring activities should result on the definition of a sound operational basis for evaluating the pros and cons of options and appraise their impacts against the actors' multiple points of view.

We decompose structuring in activities linked with (1) problem-definition, (2) model-structuring and (3) impact assessment and analysis, that is, the estimation of the consequences of implementing each one of the options considered.

In problem-definition the decision context is first characterised, the boundaries and scope of the analysis are established, the actors involved and their primary motivations and objectives for the analysis are identified, and the potential options that will constitute the point of application of the analysis are typified. The driving question for interactively structuring the problem is "what is the problem?" (cf. Rosenhead, 1996). At the end of this activity the analyst should be in "a position to suggest and describe to the decision-maker the benefits that an analytic strategy will have. The resolution of the analytic strategy is included in problem-structuring" (Buede, 1987).

Many times, the problem is not well defined and therefore it cannot be properly decided the type of decision analysis model that is more adequate to be created. Should one concentrate on modelling the complexity steaming from multiple objectives or on modelling uncertainty, or both dimensions? Should one create a model to compare options or to evaluate the intrinsic value of each one of them, or both things together (*cf.* Bana e Costa, 1996)? This last question introduces a more general and important question in problem-definition: What type of "problematic" should the analyst adopt?[1] For many people, decision aid consists purely of aiding a decision-maker to choose the best of several options. Roy (1985) first pointed out that, in reality, there are other possible problematics in decision-aiding, different to that of simple choice, such as ranking and assignment problem situations. Inspired from Roy's thoughts, Bana e Costa (1996) proposed to distinguish between the two main basic problematics, comparative evaluation (including choice of a best option, choice of a small number of options, successive choice, and ranking problems) and the problematic of intrinsic evaluation (including assignment problems and screening).

All the above questions are important to the definition of the problem at hand. This paper assumes that (1) problem-definition has previously been completed and (2) it gave enough information to decide that the adequate modelling strategy to be followed is the development of a multicriteria analysis in which (3) the explicit modelling of uncertainty issues is not a primary goal.

---

[1] "We use the word problematic to describe theanalyst's conception of the way he envisions the aid he will supply in the problem at hand (...)" (Roy, 1996, p. 57).

The paper is structured as follows. First we introduce the concept of policy unit, which generalise the concept of decision actor and helps describing conflicts in typical public policy situations. We then introduce the main structuring frameworks, group, individual and mixed frameworks, which provide the operational contexts for structuring. The notion of concern is then introduced, and the features of a family of key concerns are extensively discussed. Descriptors are then defined and classified in terms of natural, proxy and constructed descriptors. The paper provides an extensive analysis of methods that can be used to select or construct descriptors, and an extensive discussion on good practice and operational tips for this essential structuring phase. The paper then elaborates on impact analysis and introduces the concept of intrinsic evaluation.

2 Policy Units

To exemplify the type of decision problems in which the contents of this paper can be particularly useful, we refer to the case-study of selecting an investment policy for road infrastructures in the Lisbon Metropolitan Region (LMR) for the period of 1999-2004 (see Bana e Costa, 2001), from which we extracted the section related to problem-definition:

> "LMR (a regional policy unit) comprises the eighteen municipalities (local policy units). The population living in LMR is reaching the 3 millions inhabitants, 25% of which in Lisbon. {…} The transport networks, public and private, are insufficient to cope with the demand, and traffic congestion is a daily event in LMR. The decision framework in LMR requires a complex political negotiation system. One reason for this is the fact that there are no administrative regions in Portugal. Particularly in what concerns LMR, *Junta Metropolitana de Lisboa* (JML) – an executive body composed by all the eighteen municipal Mayors – has only the power conceded by the municipalities. Nonetheless, it is an important forum for debate and negotiation about inter-municipal interventions that require the agreement of the respective Mayors. Moreover, a common position of the municipalities increases their power of negotiation with the Central Government. JML intervenes in five main areas: land use planning, management and development, environment, transports and transport infrastructures, housing and social facilities, and European funds and investments. Thirty four roads projects were proposed by JML for financial support from the EC in the period 1999-2004, in order to improve the quality of LMR's road network. However, their total cost (84 billion PTE) was twice as high as the expected available budget (45 billion PTE), which implied that choices were necessary. In any case, six "fixed" projects (FPs) – that complete the National Road Network grid in congested areas and use up to 28.5% of the budget – should be implemented. The projects have different impacts at the regional and local levels. Consequently, a project considered important to the global benefit of LMR can bring no benefit for a specific municipality, and vice-versa. This may result into two types of conflicts: on the one hand, between the "best" for the region and the "best" for some municipalities (vertical conflicts); on the other hand, between municipalities (horizontal conflicts). So, a "best compromise" policy option should be discovered not only by comparing the regional attractiveness of the packages but also analyzing the degree of conflict that they can create at the local level." (Bana e Costa, 2001, pp. 111-112)

This case helps us to introduce the notion of *policy-unit*. A policy unit is a methodological construct that allows the evolution to be performed at various policy levels and/or spatial scales. It generalises the concept of *decision maker* common in the decision analysis literature. A policy unit is defined as an individual or group which has, or represents, a coherent perspective for the purpose of the evaluation. It is usually characterised by a specific set of policy concerns, partially or totally different from those of other policy units, and is interested in a specific set of impacts of the policies.

Policy units can be institutional decision makers, interest groups, administrative or political bodies representing areas (e.g., nations, regions, provinces, counties, boroughs, etc.) and the like. While policy units often reflect the institutional arrangements of policy and public decision making, for the purpose of this paper they are characterised by a coherent view on concerns and impacts of the policy, independent of affiliation or position in the decision hierarchy.

Policy units are a very useful construct in public decision making that involves the use of land and/or spatially distributed resources. For instances, the multiplicity of actors involved directly and indirectly in any public policy decision creates a situation in which conflicts are likely to arise, and these conflicts have often a clear spatial dimension: they arise between different interest groups in any one location, for example between users and non-users of a proposed new transport facility. They also arise between similar groups in different locations, for example between users in different locations who will benefit to differing degrees by the final design of a project or between those groups suffering noise from a project whose final benefit will be concentrated elsewhere. The attractiveness of a policy depends on the spatial level at which it is evaluated (the scale dimension) and the location where it is evaluated (the location dimension). For instance, a policy proposal that makes sense regionally may be unattractive when national interests and concerns are taken into account (or vice versa). The concept of policy unit allows a clear structuring of these different concerns and provides a way to analysis difference and conflicts.

The spatial component of a policy unit often originates from the hierarchical structure of geo-political areas. In the simplest case, the policy units are hierarchically organised and correspond to administrative areas which retain the decision making power for selecting, approving and implementing a policy. In general, units are organised in a network which combines hierarchical relationships and other types of coalition and combinations. The set of units is also dynamic and may change due to the formation of coalitions or to the exclusion or inclusion of policy units as a result, for instance, of modifications to the policy measures. In Figure 1 we outline ways in which policy and decision-making at different spatial levels interact in the case of large transport infrastructures in Europe (TENs, see European Commission, 2004).

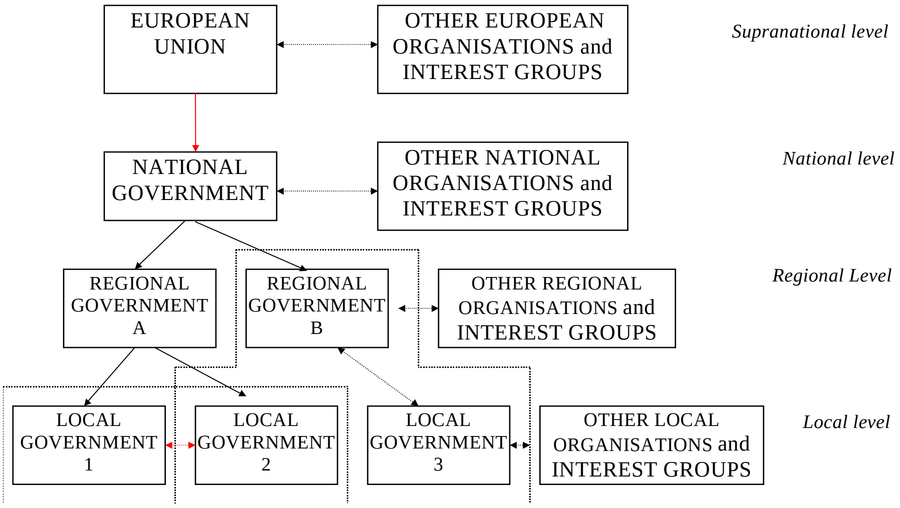


Figure 1. Structure of policy units.

The structure outlined in Figure 1 illustrates a system where there are policy bodies which can take decisions, at four different levels, supra national (EU), national, regional and local. These policy units can include both the elected governmental bodies at each level and, parallel to these, sets of other representative bodies which are either involved directly in the decision making process, or have influence over decision taken. These can include important planning and environmental advisory bodies and the operators of transport services. Each of these policy levels will take decisions regarding transport policy or projects for its own area of responsibility. In some cases bodies at different levels are limited in their powers by bodies at other levels, e.g. local government typically operates within a statutory framework determined by national government, and most operational transport policy decisions are vested at national rather than EU level by the principle of subsidiary.

As said earlier, this paper assumes that the definition of the problem has already been performed, and therefore, if appropriate, the policy-units have already been identified.

3 Structuring frameworks

Model-structuring is an interactive and learning process seeking to build a more-or-less formal representation which integrates the objective components of the problem and the subjective objectives of the actors, in such a way that the value-systems of the actors are made explicit. This representation can be constructed either for all policy units simultaneously or for each unit independently. The former, called *collective* or *group structuring* (GS), is a cooperative group structuring exercise, in which the policy units interact with the aim of reaching a common model for structuring the analysis: in particular, all units share the same key concerns (see Section 5). In the latter, called *individual structuring* (IS), each unit expresses its concerns independently of the other units. In both cases impact assessment leads to an estimate of the impacts of each option on each unit. The evaluation model can also be the same for all units or be unit-specific. In *group evaluation* (GE) all units use the same value judgements to evaluate impacts. In *individual evaluation* (IE) each unit has its own individual evaluation model which is partially or totally different from that of the others. This gives rise to three realistic operational frameworks:

1. *Group framework: group structuring and evaluation.* A model of values is constructed collectively and shared by all the units. Disagreements between units on concerns, value judgements etc. must be solved through the model building process. Failing to solve these disagreements would lead to an interruption of the process. Once they are solved, under such a group framework conflicts between units can only emerge when the impacts are not the same for all units.

2. *Individual framework: individual structuring and evaluation.* Each unit performs an independent structuring and an independent evaluation. The models employed for this purpose might be partially or totally different from those used by other units. The conflicts can be caused by any variable used for the assessment and in the construction of the units' models of values.

3. *Mixed framework: group structuring and individual evaluation.* The units structure the problem in a common way, but employ different value judgements for the evaluation of an option's impacts. Disagreements as regards concerns are solved during the structuring phase. Beyond that point, units employ different evaluation schemes. Conflicts can emerge because the units are affected by different impacts, or because they base their evaluation on different value judgements.

When several policy units are involved in policy analysis and the concerns of each one of them are structured in an individual model, the structuring exercise is essentially repeated for all units involved.

Because structuring is a mixture of art and science, it is easy to understand why significant differences, as well as many common ideas, arise in the approaches proposed in the multicriteria literature – see, for example, Keeney and Raiffa (1976), Saaty (1980), von Winterfeldt and Edwards (1986), Keeney (1992), Roy (1996). Essentially there are two basic model-structuring strategies. The most common one consists of starting by defining a set of options and then analysing the characteristics of these options to find out those that are relevant from the viewpoint of the actors (they can be called “active characteristics” in opposition to those “passive” attributes that have no impact in the actors’ value systems). These characteristics can then be used as criteria to evaluate the options. On the other hand, one can start by finding out the key objectives the actors want to achieve and then create options designed to achieve them. Keeney (1992) calls the former “alternative-focused thinking” and the latter “value-focused thinking”. He supports this latter approach: “The primitive notion for a decision problem should be values and not alternatives. If, in fact, we begin with values, we might not even think of situations as decision problems, but rather as decision opportunities” (Keeney, 1988, p. 466).

In the framework of policy analysis, we consider that “value-focused thinking” is the appropriate strategy. Indeed, a policy should be viewed as a means to achieve objectives, and therefore it is difficult to conceive that policy options can be defined prior to the decision of what they are wanted to achieve. In this sense, to choose a policy among several options is a decision opportunity rather than a decision problem.

4 The notion of concern

Once the problem has been defined and therefore the principal objectives and the type of options to achieve them has been identified, model structuring usually starts with a broad discussion of the issues of concern for evaluating the attractiveness of options. They can emerge during the discussion either as stated goals and objectives or as active characteristics and expected consequences of options. All of them are aspects or factors relevant for the analysis, which we will refer to simply as “concerns”, provided that their value meaning is well defined and perceived by everybody to avoid ambiguity and misunderstanding. This notion of concern corresponds to the initial notion of “point of view” of Bana e Costa (1990) and is consistent with the concept of “value dimension” of von Winterfeldt and Edwards (1986). To some extent it also corresponds to the common broad definition of “objective” by Keeney and Raiffa (1976) and Keeney (1992). However, even if an objective stated by an actor is obviously a concern for that actor, many types of concerns do not emerge or are not actually stated as objectives. All concerns contribute to make explicit the actors’ value systems, but not all the concerns correspond directly with objectives.

Actors’ objectives and options’ characteristics have different but complementary roles in the model-structuring process. For example, one of the main arguments against the Betuweroute railway project (see Beinat and van Drunen, 1998) was that its costs are higher than the benefits it will produce. However, “costs” is not a concern in all policy units. While the national authorities (who pay for the project) want to minimise costs, the municipalities involved in the routing decision may well disregard costs. Thus, *to minimise costs* is a stated objective for the national authorities, while costs is not a concern for all the municipalities as it is not an active characteristic of the project that impacts the value systems of all of them.

The role of the facilitator during the interactive, learning process of model-structuring consists in stimulating the reflection of the actors to progressively:

- Make all sorts of concerns emerge, clarify their meaning, and to analyse why and for what they are relevant, in order to

- Achieve a shared definition of a coherent family of key-concerns in the perspective of the policy unit(s), and to
- develop descriptors to make the key-concerns operational for evaluating the attractiveness of options.

5 Defining a family of key concerns

The literature in the field of multicriteria value analysis often refers to two ways to structure concerns: the bottom-up and the top-down approaches (cf. Buede, 1987, von Winterfeldt and Edwards, 1986). In practice a mix of the two is actually more adequate. An objective can emerge without reference to any explicit characteristic of an option. On the other hand, the analysis of characteristics, attributes or consequences of options often contributes to reveal hidden objectives or other types of implicit concerns. This provides a better understanding of the problem and of the actors' values, and improves communication between actors at different levels.

When facing ill-defined and complex situations, involving several actors and with many issues to be addressed simultaneously, structuring tools, such as 'post-it' sessions and cognitive and oval mapping techniques can be used to facilitate group-work and to help structure the issues (Eden and Ackermann, 1998 and 2001; Eden, 2004[2]). They serve to disentangle concerns and means-ends or cause-effect relationships from the cloud of aspects which usually emerge at the early stages of the analysis. Aspects are re-described, clarified, eliminated, decomposed, linked and/or grouped, to avoid ambiguity and eliminate redundancies. This learning process can be facilitated by questions used as devices to generate ideas and stimulate reflection and debate, until a final set of key concerns is identified (for a recent application see Bana e Costa et al., 2005b).

To illustrate such a process, take the example of the Lisbon road network case (cf. Bana e Costa 2001). Several experts were invited to a workshop to discuss what concerns should be considered in designing and evaluating alternative investment options in the regional road network. Following Phillips (1990), participants were "chosen to represent the differing perspectives on the problem. The modelling process helps to create a new perspective out of the separate views. If any key perspective is not represented, the recommendations of the group may be rejected by an influential individual who feels that the group did not consider certain crucial factors."

To stimulate the discussion and make new concerns emerge, the experts were confronted with questions like: "Why or for what is this concern important?", or "how can this objective be achieved?" (see Keeney, 1992). The structuring process developed in an interactive decision conferencing socio-technical framework (see Phillips and Bana e Costa, 2005) during which several chains and coalitions of concerns were identified and helped to highlight the *key* concerns shared by the group.

---

[2] For a detailed description of these and other structuring tools see (Rosenhead and Mingers, 2001). See also (Watson and Buede, 1987) and (Belton and Stewart, 2002).

Key Concerns

A *key-concern* is an individual concern, or a cluster of concerns, in terms of which the actors agree to a separate evaluation of the impacts of the options. The organised collection of impacts is the impact table. It provides an overview of the impacts of the options against the key-concerns identified in the structuring process. To organise an impact table inherently requires two necessary conditions:

1) it should be possible to describe, at least qualitatively, the plausible impacts in terms of the key-concerns within the specific problem context;
2) it should be possible to estimate, with more or less certainty, the impact of each option against the key-concern

The first condition ensures that it is possible to associate a key-concern with a descriptor of impacts. This is why broadly defined concerns that are often present in evaluations cannot be taken as key-concerns. Examples are: economic growth; social effects; environmental impacts. Similarly, this applies to generic strategic objectives such as: “to protect and enhance the natural and built environment”; “to contribute to an efficient economy”; “to support sustainable economic growth in appropriate locations” (*cf.* DETR, 2000). The second condition implies, by definition, that a key-concern must be measurable and operational in the context in which options are to be evaluated (*cf.* Keeney, 1992).

Going back to the example above, several key-concerns were identified in the Lisbon case. Reflecting on how the broad concern of *improving the quality of the regional road network* could be achieved, the concerns *improving the accessibility* and *the connectivity of the network* emerged (example of top-down procedure). To achieve these objectives, new roads need to be constructed (this makes clear that options are means to achieve ends). When thinking about the characteristics of these road projects, a new “area of concern” arose: *environmental effects* (example of bottom-up procedure). This has been further decomposed in a few more specific concerns, such as *noise pollution*, *air pollution*, *watershed* and *land use effects* (top-down). Each of these more specific concerns is a value dimension, component or facet of the same broad concern: and so on, in a mixture of top-down and bottom-up procedures.

If the evaluation of options is to be based on a comparison of their impacts on each key-concern, then a key concern must be an independent evaluation axis. This is possible if the options can be ranked with respect to that concern irrespective of their impacts in any other aspects (ordinal independence). This is often an acceptable working hypothesis. However, if the goal of the evaluation is not only the ranking of options, that is to analyse if one option is more attractive than another, but also to measure if the difference in attractiveness is small or large, then a key-concern must respect the stricter property of cardinal independence. This is also necessary for constructing an additive evaluation model, and it is one of the reasons why, very often, several concerns have to be clustered to form one single key-concern (see examples in Bana e Costa et *al.*, 1998 and 2005b).

To clarify this, consider the case of environmental effects (see Beinat, 1997). Collins and Glysson (1980) maintain that mutual ordinal independence – which is one type of “judgmental independence” (see von Winterfeldt and Edwards, 1986) – can easily be established for most environmental measures: we always prefer less pollution to more pollution, less noise to more noise, less degradation to more degradation, and so on, regardless of other aspects of the problem. This may not be the case for difference independence (the “cardinal” independence).

Let us consider a hypothetical choice problem between alternative road projects which cause different levels of air pollution effects. The key-concern “health effects due to toxic emissions” can be decomposed into the concentrations of several chemicals.

The question arises if any of these can be taken as an independent evaluation axis. The evaluation independence condition seems rather natural. But how about difference independence? Suppose for simplicity only two substances are involved. The answer relies on the way the two chemicals interact. Three kinds of interaction are usually distinguished (Figure 2): *additivity* (the combined effect is the sum of effects), *antagonism* (the combined effect is less than the sum) and *synergism* (the combined effect is more than the sum). Some substances exhibit particularly dangerous behaviour only in the presence of some other substances. In the case of “photochemical smog”, for instance, waste gases produced by automobiles, heating systems and industries interact with oxygen under the influence of sunlight. The result is the production of highly irritating substances, such as ozone, which affects the respiratory system. Several other cases of synergism are known (see, among others, Kraak *et al.*,1993; Moriarty, 1990; Ariëns *et al.*, 1976). However it is also possible that the simultaneous presence of different substances will result in a lower global effect. For instance, Ariëns *et al.* (1976) cite the use of thiosulfate against cyanide intoxication: the chemical reaction results in the formation of the substance rhodanite which has little toxicity.

Another example can be found in Moriarty (1990; see also Beinat 1997). It shows that in the case of the adverse effects of Hydrochloric acid (HCl) and sulphur dioxide (SO_2) on the photosynthesis of plants, the individual effects of HCl depends on the concentration of SO_2, and vice versa.

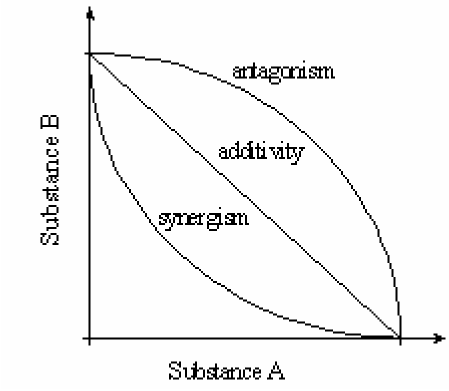


Figure 2. Iso-effect curves for substances A and B with different effect interaction.

Yet another example is the attractiveness of transport projects in terms of noise pollution on the basis of the concerns “peak noise level during the night” and “noise level during the whole day”. These concerns are (mutually) ordinally independent: whatever the average noise, low peaks during the night are always preferred to high peaks (and vice versa). Cardinal independence, on the contrary, is not obeyed. The difference of attractiveness between two night peak levels may not remain the same independent of the average noise level during the whole day. People can be more tolerant to an increase of night peak levels when the average noise level during the whole day is low. If they are subject to the same increase of night peak levels when already suffering a high average noise pollution during the whole day, the effect is magnified.

The peaks and the average both contribute to a decrease of environmental quality, but their combination is likely to produce a more than additive deterioration. In other words, it is hard to say if a given increase in the noise average produces a large or small deterioration of environmental quality without knowing the peaks during the night. These two concerns cannot be used as separate key-concerns. A good candidate would be a combined noise concern, but the combination should not be additive!

It is important to stress that the independence conditions discussed above refer to judgements (or preferences), and not to environmental (that is, physical or statistical) relations between concerns. Ignoring this distinction is a rather common pitfall of the analysis. For instance the costs and safety of a road infrastructure can be taken as two (judgmentally independent) key-concerns even if they are statistically correlated, in the sense that higher safety usually requires higher investments. As said by von Winterfeldt and Edwards (1986, p.43), "one can ignore environmental correlations {…}, unless they indicate redundancy."

Judgemental independence guarantees that the set of key-concerns is decomposable, so that the options can be evaluated in each key-concern at a time. In addition, a set of key concerns should be consensual, exhaustive, non-redundant and as concise as possible[3]. These properties give to a set of key-concerns the status of a *(coherent) family* (Roy, 1996).

Concision, or minimality of the model, is important to avoid the common temptation of including "everything" in the model, and not only what is essential for the purpose of the analysis in the eyes of the policy units involved. As pointed out by Hoobs and Meier (2000), "information 'pollution' (the generation of so much data concerning the performance of alternatives on numerous criteria that the information can not be digested by stakeholders)" is a potential weakness of multicriteria analysis applications if structuring is not carefully and properly conducted. Redundancy often leads to double counting, the consequence of which is the overvaluation of some impacts. On the other hand the model should be exhaustive, but this only applies to the key-concerns. This may, to a certain extent, act against concision, but it avoids the risk of leaving out any key-concern for a certain policy unit – which contradicts the desire for consensus. Finally, it should be stressed that in many circumstances it is difficult to appraise the impacts of options on certain key-concerns. However this does not justify their exclusion from the model. It is always better to construct a complete model even if some impacts are only roughly appraised, simply because an incomplete model can lead to deceptive conclusions.

Table 1 shows the family of seven key-concerns used in the Lisbon case (see Bana e Costa, 2001). They are grouped in areas of concern. This family was assessed as decomposable and identified as consensual – that is, collectively (regional level) and individually (local level) shared by the municipalities. These key-concerns reflect fundamental values to be taken into account when comparing alternative options. Notice that construction costs are not included as a key-concern, which may appear in contradiction with the exhaustiveness requirement, but in this case the options under evaluation are packages of road projects, defined in such a way that all have similar total costs, close to the fixed budget available. However, this does not prevent an unfair distribution of investments between policy units (municipalities), which would contradict consensus. Such an equity concern was later considered in the conflict analysis.

---

[3] See (Keeney, 1992, section 3.5) and (Belton and Stewart, 2002, section 3.4.2) for complementary discussions of desirable properties of a family of key-concerns.

Table 1. Tabular representation of a family of key concerns.

Areas of Concern	Key-concerns
Transport	Accessibility Connectivity
Environment	Noise Pollution Air Pollution Urban Land Use Non-Urban Land Use
Urban development	Urban Development Potential

The core result of this structuring process is the family of key-concerns. Instead of a tabular representation as in Table 1, it is common in complex problems to organise the areas of concern and key-concerns in a tree form, often referred to as a *value tree*.

Value trees offer a useful visual overview of the structure of the concerns in several levels of increasing specification (this is why they are often alternatively called “hierarchies” – as, for example, in (Brownlow and Watson, 1987) and (Saaty, 1980) – although this designation induces the existence of some form of “subordination” to parent nodes, which may not be present at all). Some key-concerns can appear immediately after the top (“overall”) node; other key-concerns are peers of parent nodes corresponding to general areas of concern. Depending on the evaluation context, it may be useful to organise the tree in separating “cost” concerns from “benefit” concerns”. Moreover, when the purpose is to emphasise only the key-concerns, each one of these will always appear at the end of a branch starting at the overall node, eventually passing throughout several intermediate nodes (see example in Figure 3). In other cases, the tree can be structured with the purpose of giving a detailed overview of the value issues, in which case the key-concerns can appear at any level, that is, not necessary at terminating nodes.

On the basis of a family of key-concerns it is already possible to proceed to the evaluation of the attractiveness of the options for each concern. This is often based on direct rating or scoring of options as in the SMART approach (*cf.* von Winterfeldt and Edwards, 1986; Edwards and Barron, 1994). While this approach may be valid in many cases, the transparency of the result is limited, due to the number of implicit steps, assumptions and evaluations that are implied by the direct scoring of an option against a concern.

It is useful to stress that multicriteria decision aids can be used in two basic contexts: In the first, one single policy unit seeks to make a decision that does not seriously impact on, or require justification to, other policy units. In this case, the decision does not need substantial justification, and direct scoring of options can be sufficient. This contrasts with the context in which the decision is to be taken by one unit on behalf of several units not directly involved in the evaluation process, or by several units collectively. In such cases, the rationale for the evaluation of options must be clearly documented, and proper consideration of each interest or point of view must be demonstrated. This requirement necessitates the use of rather more complete models, “even where these may be less efficient, and/or may impose structures (of preferences for example) which may not strictly be justifiable empirically” (Bana e Costa *et al.*, 1997). Even if defining descriptors of plausible impacts in the key-concerns is often a hard task, it contributes decisively to well informed judgements and justified and transparent evaluation.

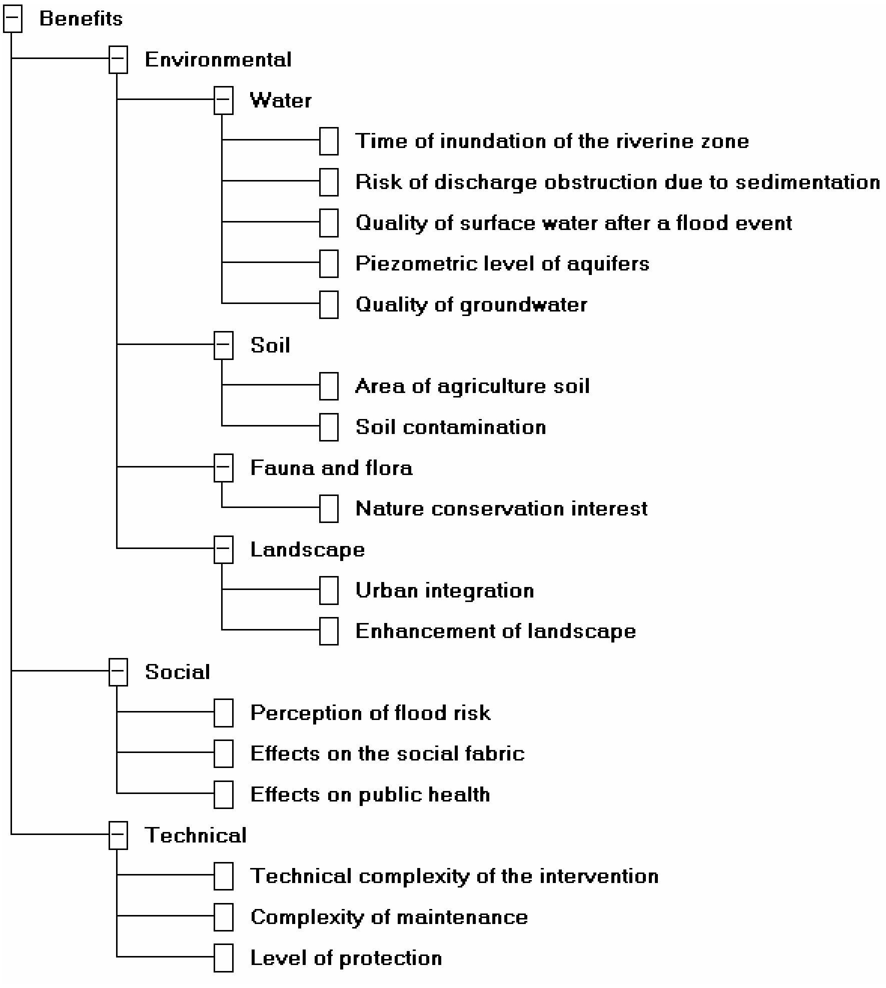


Figure 3. Example of a value tree of key-concerns (Bana et al., 2004).

6 Descriptors

A descriptor of impacts is an ordered set of plausible impact levels associated with a key concern. It is intended to:

1. Operationalise the appraisal of impacts (performances or consequences) of options in a key-concern, that is to measure (quantitatively or qualitatively) the degree to which the key-concern is satisfied.
2. Describe, as much as possible objectively, the impacts of options with respect to that key-concern. The more objectively the impacts are appraised, the better understood (less ambiguous) and therefore the better accepted (less controversial) will be the evaluation model.
3. Better frame the evaluation model, by restricting (whenever appropriate) the range of impact levels to a plausibility domain (from a most attractive or desirable level to a least attractive level). This impact-range can be defined by screening out impacts or options that are non-admissible or out-of-context. Depending on the type of evaluation framework, the impact-ranges of the descriptors can be defined for each unit individually or for all of them together.

4. Verify the ordinal independence of the corresponding key-concern. If dependence is detected in this phase, a feedback is necessary to re-structure the family of key-concerns, so that “all other things being equal” (ceteris paribus) comparisons of options may be made individually against each key-concern.

The contextual operational requirement will permit a key-concern to be described in such a way that clarifies why it is important for at least one unit, and, when several units are involved, each one will better understand the other ones’ values. This is also enhanced by the ordinal requirement imposed on a descriptor, making it an (ordinal) evaluation criterion – this being defined as a model allowing the comparison of options according to a particular point of view, along the lines of the definitions given in (Bouyssou, 1990) and (Bana e Costa *et al.* (1997). In fact, imposing the condition that the levels of the descriptor should be rank-ordered, in terms of their relative attractiveness, gives to the descriptor the structure of an ordinal preference scale[4], that at least makes it possible to appraise which of two options, if either, is more attractive than the other once their impacts have been assessed. Note that this preference is limited to the key-concern under consideration; hence it is called “partial” attractiveness, to distinguish it from “overall” attractiveness or preferability (i.e. at the level of the entire family of key-concerns).

The direction of attractiveness is often inherent to the descriptor, as, for example, when measuring the (gain in) accessibility from a new road by the (reduction in the) average travel time in a region. Conflicts between policy units can however arise when, for some of the units, more impact is better than less and, on the contrary, less is better for others. For example, the number of stations on a high speed rail line such as the Channel Tunnel Rail Link (Vickerman and Norman, 1999) might be desired to be minimised from a high level (European) perspective in order to maximise speeds between major cities, but maximised from a lower level (local) perspective in order to increase the perceived local benefits. In an individual structuring framework, one may argue that these conflicts could be avoided at this phase of structuring, simply by taking the attribute “number of stations” and postponing the question of the direction of preference to the conflict analysis phase. However, this will not contribute to shedding light on the problem.

Associating a descriptor to a key-concern is often a matter of choosing which of several existing attributes of the options is the most adequate in the specific problem-context.[5] For example, should accessibility impacts be measured on an absolute scale (time in minutes, or distance in km, or …?) or on a relative scale (deviation from a reference threshold, or a percentage, or …?)? It is worthwhile to caution here against the tendency to include in the model, not only one, but several attributes that are, in fact, alternative descriptors of the same key-concern; this is a common pitfall of analysis which introduces redundancy in the model, that may imply overweighting a key-concern in terms of its relevance for evaluating options. It is often also the case that the discussion of alternative descriptors actually reveals hidden

---

[4] See (Roy, 1996, § 8.1.2.1). The scale can be continuous or discrete.

[5] We take here an attribute in the common language sense of a characteristic of policy options, not in the classical and widely adopted sense given to it in multicriteria analysis since (Keeney and Raiffa, 1976) which uses the term attribute to designate what we call a descriptor (see also Keeney, 1992). Other designations appearing in the literature are performance measure, measure of effectiveness, criterion, evaluation measure (scale) (see Kirkwood, 1997).

concerns, which may imply a need to restructure the key-concern or even the family of key-concerns. Structuring is, indeed, a recursive learning process (Figure 4).

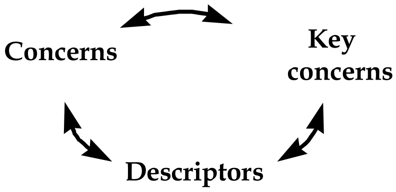


Figure 4. The learning cycle for the definition of concerns, key concerns and descriptors.

Depending on the context, alternative attributes may be more or less directly related with the key-concern. It is then useful to distinguish between direct and indirect descriptors:

1. The levels of a <u>direct, or natural, descriptor</u> directly reflect effects; such as the number of people affected by respiratory diseases.
2. The levels of an <u>indirect, or proxy, descriptor</u> indicate causes more than effects; such as the degrees of concentration of air pollutants that cause respiratory diseases.

It is not always certain that a natural descriptor is necessarily more adequate than a proxy one, as it is discussed in section 7.[6]

On the other hand, often neither a direct nor an indirect attribute exists which is appropriate to be used as the descriptor of the key-concern. This occurs when the key-concern has an intrinsically subjective nature (image, for instance), or for reasons of lack of information, or because the key-concern is a cluster of several interrelated elementary concerns or value dimensions that are *judgementally dependent*. Consider, for example, the key-concern “improving accessibility”. In the Lisbon Region Case, it has been described by the (direct) descriptor “average travel time decrease (%)”, assessed at either the municipal or regional levels. However, in other contexts, this attribute could be considered inadequate to describe the impacts in terms of improvement of accessibility, for different reasons, namely: 1) the technical conditions necessary to measure the average travel time so accurately do not exist, or 2) improving accessibility is actually a cluster of several concerns, like the geographical dispersion of the benefits, the contribution to removing existing bottlenecks, the degree to which the achievement of planning goals is favoured, etc.

In such cases, a <u>constructed descriptor</u> has to be developed within the specific context. The levels of a constructed descriptor can be qualitative or quantitative (or mixed) and several types can be conceived (like, for example, verbal descriptions of expected consequences, reference impact-profiles, visual representations, indices, etc.), as discussed in section 8.

As the choice or the construction of a descriptor for a key-concern inherently requires (qualitative, or ordinal) value judgments, conflicts of viewpoints may arise in group-structuring, because different policy units may tend to privilege different descriptors, even if all of them agree in the key-concern. Unavailability of data in some of the units can further complicate the task of defining one common descriptor.

---

[6] Some authors consider that a “natural scale is one that is in general use with a common interpretation by everyone.” (Kirkwood, 1997, p. 24), and therefore that they can be either direct or indirect.

7 Natural and proxy descriptors

A decision problem almost always allows for more than one representation and one choice of descriptors (Brownlow and Watson, 1987), and none of the choices is strictly right or wrong. In many applications the choice of descriptors may be related to the availability of appropriate information for describing the effects of policies (cf. also Haasstrup, 1994). Figure 5 shows a cause-effect chain used for describing the environmental consequences of transport policies which lead to environmental emissions. The box “environmental effects” in Figure 5 represents a natural descriptor, in the sense that it describes the consequences of a decision in terms that are more directly linked to the real interests of decision makers. Although this is the relevant aspect for decision-making, a proper quantification and description of environmental effects may be difficult. Other descriptors can also be used (for instance, the doses, the environmental concentration or even the emissions). They are proxies, which, at least in principle, are factually linked to the natural descriptors.

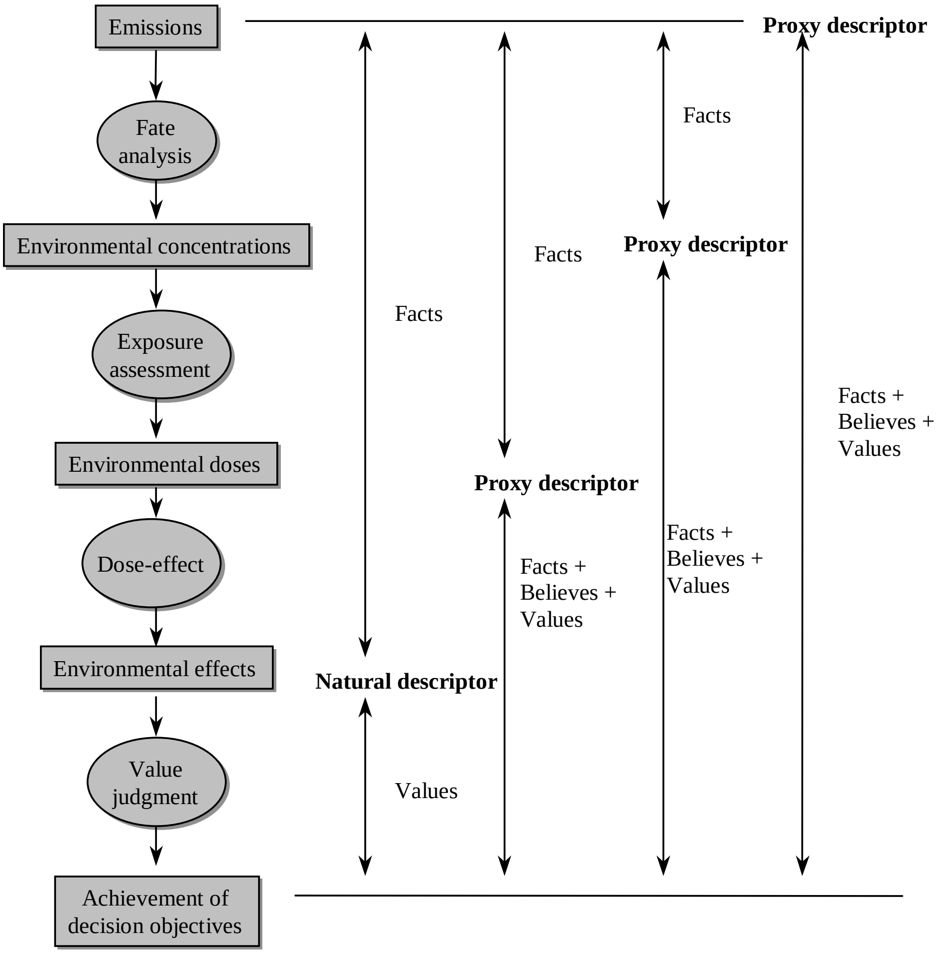


Figure 5. Alternative positions of natural and proxy descriptors in a simplified cause-effect chain for environmental pollution (adapted from Beinat 1997).

The measurement of the impacts or expected consequences of policies can be significantly simplified or even made straightforward with proxy descriptors. Yet significant complications usually appear at a later stage. To compare the achievements of different policies, it is necessary to determine the relation between the natural and the proxy descriptors. The use of a proxy descriptor also implies that the evaluation entangles facts and values. Figure 5 shows various options for the position of the descriptor in the cause-effect chain. By choosing a natural descriptor, the distinction between facts and values is clear, at least in principle. Factual information and value judgement have a distinct role and so have the people in charge of the assessment and of the evaluation. However, this distinction between facts and values may be impossible due to insufficient information on the natural descriptor. Proxy descriptors, being positioned at an earlier stage of the cause-effect chain, are less demanding. In the extreme case in Figure 5 the proxy descriptor measures the initial cause of environmental effects: the emissions. Describing the policies in terms of emissions is the easiest option among all the possible choices of proxy and natural descriptors. However, it is also the least clear in terms of the actual focus of the decision.

Evaluating policies with proxy descriptors complicates the matter and may introduce biases and distortions in the evaluation:

- It may affect the outcome of the analysis. As an example, Batterman *et al.* (1988) show how different environmental impact indicators highlight different areas for sulphur reduction in Europe and therefore different conclusions.
- It may determine substantial weight biases. For instance, Fischer *et al.* (1987) demonstrate that using the proxy “levels of dust emission” instead of the natural descriptor “person days per year of asthma due to dust exposure” leads to substantial and systematic overweighting of the proxy descriptor in the decision context.
- It may result in the generation of overlapping descriptors. For example, Beinat (1994) (see also Beinat *et al.*, 1994a) shows a case in which several proxy descriptors are necessary in order to describe the economic effects of infrastructure projects. They result in double counting and implicit overestimation of the importance of economic effects.

Advantages and disadvantages of the use of proxy descriptors can be linked to the position of the descriptor in the cause-effect chain. This often corresponds to the “distance” between the source of the impact and its consequences. Descriptors close to the cause (for example, emissions or environmental concentrations) are relatively easy to measure, owing to the accessibility of the medium to be measured (for instance, water, air, soil) and to the knowledge of the phenomenon determining the emissions (for instance, a combustion process). In comparison with descriptors positioned at later stages, the descriptor scores are more reliable, relatively accurate and less controversial. In addition, the close relation between the data used for the decision and the data used for designing technical solutions, for instance emission abatement devices, implies that the outcomes of the decision may be directly used to design better solutions. Finally, the responsibility for the cause of impact can be traced back with more accuracy. In spite of these advantages, assessing decision alternatives on this basis may be difficult. Very often, the relationship between the variable measured and the objectives of the decision is unclear and may hinder complicated processing of information. By evaluating alternatives on this basis, factual data and value judgement are mixed, making the distinction between technical and a political responsibility difficult.

Descriptors close to the consequences and the ultimate concern of the evaluation (for instance, effects on human health or on ecosystems) have different features. They allow a clearer interpretation of the costs and benefits of alternative options since they address the decision concerns directly. The distinction between factual information and value judgement is clearer. The political (value) judgement and the responsibility for the decision are addressed explicitly, increasing the transparency of the decision. Nevertheless, the use of descriptors close to the consequences of the decision implies the qualification (and possibly the quantification) of all the intermediate steps, starting from the initial cause. This sequence of steps may be understood only by a limited number of specialists. The lack of knowledge and/or information for this process may also imply that the descriptor impact-levels are affected by a high degree of imprecision and that the reliability of the information for the decision depends heavily on the quality of the assumptions made throughout this process. This may severely hamper the evaluation and comparison of decision options.

In practical applications, proxy descriptors close to the source of the problem are frequently used, implying that the advantages associated with the measurability, controllability and neutrality of the descriptor impact-levels are weighted heavily in applications[7].

8 Constructed descriptors

As mentioned earlier, frequently no natural or proxy attributes exist which are adequate to describe, in a specific decision context, the impacts in terms of a certain key-concern. A descriptor has therefore to be developed. The designation "constructed" descriptor is used to indicate precisely that the structuring task is not limited to the choice among alternative existing attributes. Several types of constructed descriptors can be developed. They can be characterised in terms of the number of value dimensions involved (one-dimension and multidimensional descriptors) and in terms of the way their impact levels are described, pictorially (pictorial descriptor), verbally (qualitative scale), or numerically (indices and formulas). Many combinations of these properties can be envisaged in practice; we will elaborate on:

- One-dimension qualitative scales: finite sets of verbal descriptions of one-dimension plausible impacts.
- Pictorial descriptors: finite set of reference visual representations of impacts.
- Multidimensional descriptors: sets of multidimensional impact levels describing plausible policy scenarios. It is useful to distinguish between *multidimensional (discrete) scales* and *indices.* The former are constructed by individual descriptions of each of their multidimensional impact levels, the latter are defined comprehensively by a formula that combines several dimensions analytically. Table 2 illustrates the three cases.

---

[7] The use of proxy descriptors is very common. For instance, Briassoulis (1995) listed and classified 210 environmental criteria used for industrial facility sitting. About 80% of these criteria appear to be proxies selected for their measurability and clarity.

Table 2. Examples of constructed descriptors.

Type	Context	Key concern	Descriptor
One-dimension qualitative scale	Evaluation of power plant sites (Keeney, 1992)	Public attitudes	*Support*: No groups are opposed to the facility and at least one group has organised support for the facility. *Neutrality*: All groups are indifferent or uninterested. *Controversy*: One or more groups have organised opposition, although no groups have action-oriented opposition. Other groups may either be neutral or support the facility. *Action-oriented opposition*: Exactly one group has action-oriented opposition. The other groups have organised support, indifference or organized opposition. *Strong action-oriented opposition*: Two or more groups have action-oriented opposition.
Pictorial descriptor	Evaluation of consequences of river works in Chile (see Nardini, 1998, p. 212-215)	The visual effects for a recreational area near a waterfall.	
Two-dimensions scale formed by all the plausible combinations of two intertwined aspects	Construction of the new railway line to the port of Lisbon (adapted from Bana e Costa *et al.*, 2001).	Effects on the railway service to the port during the construction of the new line.	*Very good*: It is possible to split up the construction of the new line in all its track-sections, while keeping in operation the old line. *Good*: It is possible to split up the construction in two track-sections only, while keeping in operation the old line. *Neutral:* It is impossible to split up the construction, while keeping in operation the old line. *Bad*: It is impossible to keep in operation the old line during the construction of the new line.

One-dimension qualitative scales

The simplest case of an one-dimension qualitative descriptor is when it is sufficient to develop a *dichotomous descriptor* distinguishing only two levels of impact, such as “yes” and “no” in contexts in which the matter of concern is just to know if consequences of implementing each option are expected or not.

The dichotomous descriptor {“support, or in favour” (most attractive level); “opposition, or against” (least attractive level)} can be used to take into account the behaviour of a certain policy unit towards the implementation of an option: for instance, the consistency of alternative highway layouts with local development strategies.

In general, however, in a group-structuring framework involving several policy units, intermediate levels between the two extremes, for instance, several reference degrees of support or opposition, must be described. In the Lisbon Region road case (Bana e Costa, 2001), the “potential for conflict of a policy option at the local level” was initially included in the family of key-concerns, operationalised by a qualitative scale where four reference levels were verbally described as follows: *consensual support* – all of the mayors of the 18 municipalities agree with the policy; *major support* – a majority of mayors agree with the policy and no mayor is strongly opposed to it; *controversy* – a minority of mayors is against the policy; and, *strong controversy* – a majority of mayors is against the policy. The construction of this descriptor was inspired by the five-level scale of plausible “public attitudes” shown in the second row of Table 2. This type of descriptor can be very useful in anticipating conflicts in group-structuring frameworks.

It is not uncommon to find studies of environmental impact assessment which use qualitative scales like {“very good”, “good”, “acceptable”, “bad”, “very bad”} without any objective description attached to the levels, or even {“no impact”, “minor impact”, “major impact”} with no clear description of what differentiates a major from a minor impact. Such ambiguous definitions of impact levels lack objectivity and therefore contribute neither to transparency in decision-making nor to the analysis of the causes of eventual disagreements about the attractiveness of options. Ambiguity should be avoided as much as possible in the construction of descriptors, although this can be difficult in cases where information is scarce. Eventually, the facilitator can ask the actors for examples of what is a good or a bad option, in terms of the key-concern to be described, or what would distinguish between a major and a minor impact, etc., and then start a discussion on the various responses in order to find shared verbal, or at least pictorial, descriptions of the different levels of impact.

Another common example of an ambiguous constructed descriptor is to take intervals defined upon a continuous scale as its impact levels. For instance, consider again the “accessibility” key-concern of the Lisbon case (see Bana e Costa, 2001) to which was associated the continuous descriptor “average travel time decrease (as a percentage of the status quo average travel time)”, ranging from a least plausible impact of -5% to a best plausible impact of +20%. A mistake would be to use, instead of this descriptor, a set of interval levels such as, for example, “bad: from –5% to 0%”, “acceptable: from 0% to 5%”, “good: from 5% to 10%”, “very good: from 10% to 15%”, “excellent: from 15% to 20%.”

Pictorial descriptors

When verbal descriptions are difficult to make, or seen as difficult to understand, visual representations of impacts can be used, like sketches, drawings, pictures (or selected parts of them – see an example in Krischer, 1976), videos, computer simulations, etc. Their main purpose is to convey a realistic contextual representation of the expected impacts in such as way that they can be assessed and compared. This is a necessary step to ensure that the actors involved in the evaluation perceive the meaning of an impact descriptor and can relate it to their value systems. Wenstøp and Seip (2001) elaborate on this association and link the need in decision making for well-founded values to the ability of descriptors to elicit emotions. Following Damásio (1994), the authors link emotions to neuro-physiological phenomena that appear to be important markers of how well founded are beliefs and values. To trigger an “emotion”, however, it is often necessary to elaborate the descriptor form so that it can be understood, put into context, and assimilated. In cases where numerical or textual descriptors are insufficient, pictures, animations or other multimedia forms, can reduce the barrier to the understanding of the descriptor meaning, and facilitate, if not allow at all, the elicitation of values.

An example of a pictorial descriptor is shown in the third row of Table 2. In this case the waterfall represents the single most attractive landscape element for a recreation area. The water works planned upstream have the effect of modifying the water flow, thus reducing the flow intensity and indirectly the visual quality of the falls. The pictorial descriptor has been used to simulate various levels of impact that correspond to the possible water flows and different project options. Through computer simulation, the waterfall appearance has been created for nine reference scenarios, ranging from the status-quo (preferred scenario) down to the worst case scenario corresponding to a radical change of the water flow and the virtual loss of the waterfall. The impacts of each of the planned alternatives could then be associated with one of the nine scenarios. This made it possible to compare the visual impacts of various alternatives in a meaningful way.

A similar approach is described by Wenstøp and Carlsen (1998) for the evaluation of hydropower development projects. To facilitate the assessment of several projects, they employed videos to represent the effects of various development scenarios. “Video recordings of the existing landscape with its rivers and waterfalls were taken from a helicopter. An artist had edited the video to show what it would look like after development. The video was played to the panels before the weighting process, and it was the impression of the moderators concluding the valuation session that this helped considerably in creating emotions around the trade-off decisions” (Wenstøp and Seip, 2001).

Video simulation was extensively used to compare the available options to protect Venice from high water (see Collegio di Esperti Internazionali, 1998) The three major options available[8] have some impacts that are easier to comprehend if visually represented. The landscape impacts of the mobile gates, for instance, were simulated though computer graphics from a variety of perspectives and in a variety of operative scenarios to facilitate the assessment of the change in the landscape from the status quo (see Collegio di Esperti Internazionali, 1998). In a similar way, simulated movies that show the tidal propagation in the lagoon under the effect of the protection measures helped convey the different water protection capabilities. This facilitated the comparison of the levels reached in the city at tidal peak, the duration of the inundation, the areas affected by high water, etc. conveying a feeling and a tangible perception of the meaning of implementing one solution instead of another.

Note that visual representations can be used to complement both one-dimension and multidimensional scales, or vice-versa. Moreover, several existing and well-known reference options can also be used to give reality to descriptions of several levels of a key concern. Other examples and a good discussion of the use of pictorial descriptors can be found in (Keeney, 1992, p. 109).

Multidimensional scales

To introduce the discussion on multidimensional scales, consider the key-concern “effects on the railway service to the port during the construction of the new line” described in the last row of Table 2. The (bi-dimensional) levels of the constructed descriptor are all the plausible combinations of the impacts in two aspects clustered in the key-concern: the possibility of splitting up the construction of the new line into track-sections, and the possibility of keeping the old railway line in operation during the construction of the new line. The former was described by the number of track-sections into which the construction could be split: 3, 2, or 0 track-sections. The latter was described dichotomously: “Yes: it is possible to keep in operation the old line during the construction” or “No: it is not possible to keep in operation

---

[8] The major options are: the closure of the lagoon with mobile gates to stop high tides; the introduction of lagoon morphological modifications to slow down tidal propagation: the raising of the city pavements to reduce tidal damage. A variety of combinations of these these options are also considered, leading to a vast portfolio of possible implementation measures.

the old line during the construction”. Why the two aspects were not considered each one as an individual key-concern? Because, regarding the effects on the railway service to the port during the construction, it was judged that splitting the construction into sections is technically irrelevant if at the same time it is not possible to keep the old line in operation. Thus the least attractive level of the descriptor was simply defined as “it is impossible to keep in operation the old line (during the construction of the new one).” Or, put another way, as shown in Table 3, three among all the possible combinations of the levels of the two aspects were judged as indifferent (equally attractive) in terms of their effects on the railway service during the construction.

Table 3. Development of a bi-dimensional scale using a two-entries table

		Is it possible to keep in operation the old line during the construction?	
		YES	NO
Number of track-sections into which the construction can be split:	3 sections	(YES; 3 sections) ***very good effect***	(NO; 3 sections) *indifferent to*
	2 sections	(YES; 2 sections) ***good effect***	(NO; 2 sections) *indifferent to*
	0 sections	(YES; 0 sections) ***neutral effect***	(NO; 0 sections) ***bad affect***

This clearly reveals that, although the concern with keeping the old line in operation is independent of the concern with splitting up the construction (because “yes” is always better than “no” regardless of whether it is possible or not to split up the construction), the second concern depends on the first one (because splitting into 3 sections is better than 2, which is better than none when it is possible to keep the old line in operation, but 3, 2, and 0 sections are indifferent if it is not possible to keep the old line in operation). It is worthwhile to note that an average sum of different scores assigned separately to the levels of each of the two interdependent concerns would not reproduce the overall ranking. Note also the lexicographical judgemental structure of this ranking.

A two-entries table like Table 3 greatly facilitates the construction of a bi-dimensional scale. Alternative developments are profiles-tables and tree designs (see Table 4 and Figure 6, respectively). Each cell in the two-entries table, or each path in the tree, or each line in the profiles-table, corresponds to one particular combination of levels (one and only one level of each dimension per profile). Contrary to a two-concerns table, tree designs or profiles-tables can be used to facilitate the construction of multi-dimensional scales combining more than two aspects. The advantage of developing a multi-dimensional scale with the support of a profiles-table is given by the easiness of eliminating rows corresponding to unfeasible profiles, merging rows corresponding to indifferent profiles, and reorder rows to order the impact levels by their relative attractiveness.

Table 4. Tabular development of a bi-dimensional qualitative scale

Effects on the railway service to the port during the construction of the new line						
	Is it possible to keep in operation the old line during the construction?		*The construction can be split into how many sections?*			Levels
	YES	*NO*	*3 sections*	*2 sections*	*0 sections*	
Profile 1	YES		3 sections			Very good effect
Profile 2	YES			2 sections		Good effect
Profile 3	YES				0 sections	Neutral effect
Profile 4		NO	3 sections			Bad effect
Profile 5		NO		2 sections		
Profile 6		NO			0 sections	

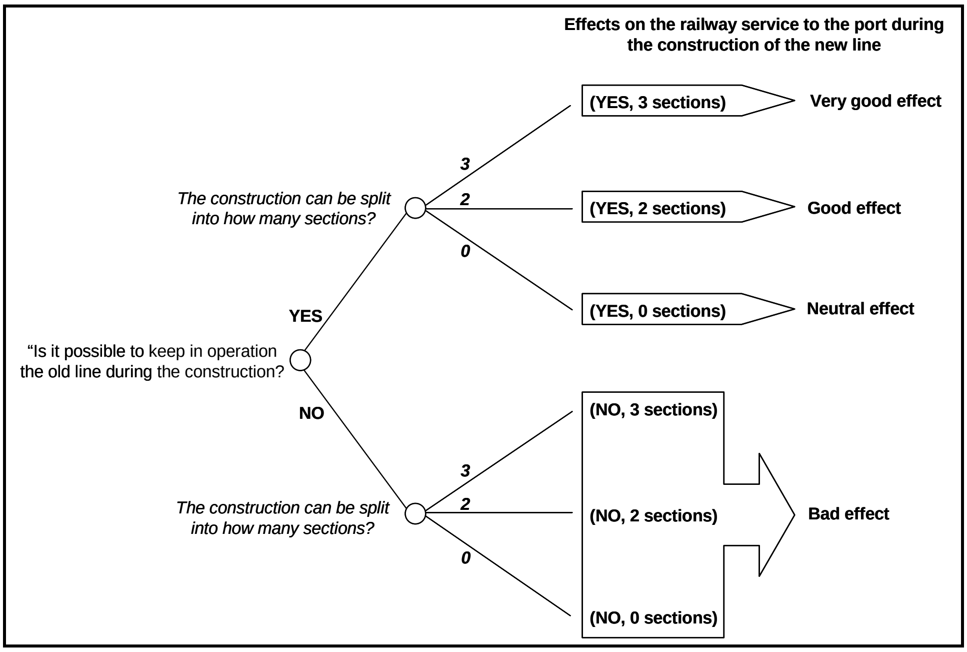


Figure 6. Design of a profiles-tree.

Table 5 resumes the basic steps of a systematic procedure to construct a multidimensional descriptor for a key-concern that clusters several intertwined dimensions.

When at least one of the dimensions is described by a quantitative continuous scale, step 1 can only be developed after breaking the continuous scale into a finite number of discrete levels. The impact range can be split up into a few intervals and the limits of these intervals (not the intervals) taken as reference levels which span the range and can then be combined with the levels of the other dimensions. To illustrate this, consider again the case of the new railway link to the port of Lisbon and the key-concern analysed above. At an early stage, instead of the number of track-sections, the experts thought initially in terms of "the average length of the track-sections into which the construction can be split", ranging from 200 metres to 600 metres (which was the full extension of the longest layout of the new line, corresponding to no possible split). Bisecting the range from 200 to 600 meters would give rise to three reference levels of average length – 200 meters, 400 meters, and 600 meters – that could then be combined with being or not being possible to keep the old line in operation. Actually, this was not done because it was later found out that the direction of preference for length, the less the better, could only be established for a given number of track-sections into which the construction could be split. Finally, note that it is not a necessary condition to take equally spaced consecutive reference levels; in some contexts, it is more adequate to take well-known references, as, for example, when thresholds of concentration of several pollutants are combined (*cf.* Beinat, 1997).

Table 5. A procedure to develop a multidimensional constructed scale.

Basic steps	What to do	Comments
Step 1	Define a discrete set of impact-levels in terms of each of the componential dimensions.	This may require making the quantitative dimensions discrete.
Step 2	Establish all possible combinations of the levels of the several dimensions.	Developing all possible multidimensional combinational can be called a "factorial design" (cf. Barron and Person, 1979); this can be facilitated using tables or trees.
Step 3	Eliminate the infeasible (not plausible) combinations.	
Step 4	Compare the desirability of the feasible combinations and group those ones that are judged as indifferent in terms of the key-concern; each group of profiles form a same plausible impact-level of the descriptor (if convenient, give a label to each level); rank the plausible impact-levels by decreasing relative attractiveness in terms of the key-concern.	This requires holistic comparisons of multidimensional profiles which can involve considerable judgemental effort; this can be facilitated using a pairwise comparison procedure. Clearly identify the most and least attractive multidimensional levels.
Step 5	Make a textual description of each plausible impact-level, as detailed as appropriate and as objective as possible.	If appropriate, use pictorial representations or mention to existing options to complement and to give reality to each description.

In practice, the number of facets of a key-concern and/or the number of their levels are often such that the full analysis of all possible impact-profiles is impracticable, even for a few dimensions with a few levels: for 4 dimensions each one described by 4 levels, 4^4=256 profiles would have to be analysed! One can then opt for developing the multi-dimensional descriptor on the basis of only a few reference impact-profiles, noticeably different in terms of their relative attractiveness, and each one formed by particular feasible combination(s) of impacts in the several aspects. There are several possible ways to reduce the number of combinations: to consider only a small number of reference impacts in the several aspects, and then to form all the combinations of these, and (or) to apply a procedure to select only a few combinations from among many possible ones.

In the process of structuring the aspects that should be considered in the comparison of alternative sites for the new international airport of Lisbon (Bana e Costa and Corrêa, 1999), one of the key-concerns identified was the "consonance with local development strategies". The concern is meant to shed light on the degrees to which:

- the siting of the new airport would be in agreement or disagreement with the economic development profile established in the Municipal Master Plan;
- the siting would be compatible or incompatible with the land uses foreseen by the Land Use and Master Plans;

- the choice would have a positive or negative effect on the local demographic trends;
- the industrial and urban infrastructures that exist at the site are sufficient to support the increase in population and activities that the new airport will bring.

Two reference impact-levels were defined for each of these four dimensions, as described in Table 6 (step 1 of the procedure in Table 5), giving rise to the total of sixteen combinations for which profiles are shown in Table 7 (step 2).

Table 6. Facets of the key-concern with the "consonance with local development strategies" and description of their reference impact-levels

Dimensions	Reference impact-levels
Agreement with the municipal development profile(s)	*Agreement:* The sitting of the new airport will contribute to the concretisation, or at least will not contradict, the objectives established in the Municipal Master Plan(s) with respect to functional and economic specialisation. *Disagreement:* Otherwise.
Land use compatibility	*Compatible:* The sitting of new airport will not conflict with the local land-uses foreseen by the Land Use Plan(s) and the necessary change of the land use of the site foreseen by the Municipal Master Plan(s) is admissible. *Incompatible:* Otherwise.
Effects on demographic trends	*Positive:* the increase of population induced by the new airport will contribute to the consolidation of the positive demographic trend that exists in the municipality, and eventually also to speed up population increase, in the case this is an objective of the Municipal Master Plan(s); or it will permit to invert the observed trend to territorial desertification and population aging, by contributing to filling up urban spaces or those to which the Municipal Master Plan(s) assigns urban use. *Negative:* The changing of the demographic trend that the new airport will provoke will clearly invert the desirable population trend defined in the Municipal Master Plan(s), or, it will eventually contribute to worsen the over occupation of the territory.
Infrastructural potential to support urban and industrial growth	*Existing*: the urban structure has enough capacity to absorb the socio-economic growth that will be induced by the new airport – the infrastructures' capacity thresholds defined by the Municipal Master Plan(s) will not be exceeded; there exists a enough developed urban network and available industrial areas nearby the site. *Non-existing:* Otherwise.

Six profiles (shaded in Table 7) were considered infeasible and, consequently, they were eliminated (step 3). The remaining ten profiles were pairwise compared (step 4), two pairs of which were considered to be equally attractive and therefore each pair was taken as different profiles of a same multidimensional impact-level (step 4); the derived ranking enabled the definition of an eight-levels constructed scale, labelled L8 (most attractive) to L1 (least attractive), as described in Table 8 (step 5).

Table 7. Tabular development of a constructed scale from reference impact-levels on the four dimensions of the key-concern "consonance with local development strategies" (adapted from Bana e Costa and Corrêa, 1999)

Consonance with local development strategies																	
Dimensions	***Step 1)*** Reference impact-levels	***Step 2)*** Multidimensional profiles															
Municipal development profile(s)	*Agree*	***A***	***A***	***A***	***A***	***A***	***A***	***A***	***A***								
	Disagree									***D***	***D***	***D***	***D***	***D***	***D***	***D***	***D***
Land use compatibility	*Compatible*	***C***	***C***	***C***	***C***					***C***	***C***	***C***	***C***				
	Incompatible					***I***	***I***	***I***	***I***					***I***	***I***	***I***	***I***
Effects on demographic trends	*Positive*	+	+			+	+			+	+			+	+		
	Negative			–	–			–	–			–	–			–	–
Infrastructural potential	*Existing*	***E***		***E***		***E***		***E***		***E***		***E***		***E***		***E***	
	Non-existing		***N***		***N***		***N***		***N***		***N***		***N***		***N***		***N***
Step 4) Multidimensional impact-levels		L8	*Step 3)*			L7	*Step 3)*			L6	L5	L3	L2	L5	L4	L2	L1

Facilitating the ranking of multidimensional profiles

The ten multidimensional impact-profiles that were not eliminated in step 3 (see Table 7) were pairwise compared in step 4 in order to derive a ranking in order of decreasing attractiveness in terms of the key-concern. The major limitation associated with the assessment of full profiles is the possibility of information overload (*cf.* Miller, 1956). The assessor has to compare multidimensional profiles mentally, without any type of support. This requires an ability to consider multiple impacts simultaneously, to make mental trade-offs and to compare profiles on the basis of all their pros and cons. Due to the complexity of the task, the assessor may be constrained to use *ad hoc* judgement strategies, such as focusing on the most prominent attributes, or simple decision rules such lexicographic ranking or other non-compensatory aids which simplify the judgement task (Kerstholt, 1992).

The comparison can be facilitated by specific tools, such as those for ordinal ranking included in the M-MACBETH software (www.m-macbeth.com). For each pair of profiles x and y, the evaluators is asked to answer to the following question procedure:

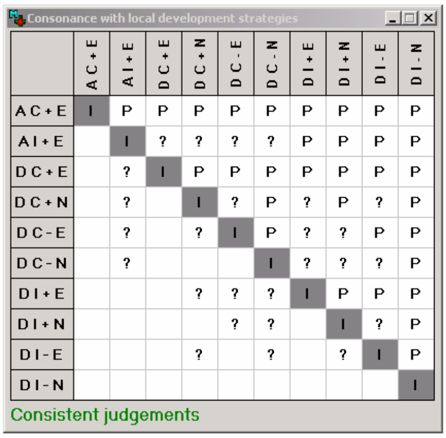
Is one of the profiles (x or y) more attractive than the other? YES NO
If yes, which profile (x or y) is more attractive than the other? x y

The answers obtained allow the definition of two binary relations on the set of profiles (*cf.* Bana e Costa *et al.*, 2005a): The relation P (strict preference) defined by xPy if and only if the evaluators judged x more attractive than y, and the relation I (indifference) defined by xIy if and only if the evaluator did not judge either of the two options more attractive than the other (i.e. if and only if the evaluator answered NO to the first question). The advantage of using the M-MACBETH software is in that, each time a pairwise comparison is made, it tests if the judgements formulated so far are consistent with the existence of a ranking. In case of inconsistency, a cycle exists, which is shown by the software, and it points out alternative suggestions to bypass the problem.

Table 8. Multidimensional constructed descriptor of "consonance with local development strategies" (adapted from Bana e Costa and Corrêa, 1999).

Impact levels	Description
L8	The siting of the new airport is in agreement with the municipal development profile(s), has positive effects in terms of demographic trends, is compatible with the land uses foreseen in municipal plans, and there is sufficient infrastructural potential to support urban and industrial growth induced by the new airport.
L7	The siting of the new airport is in agreement with the municipal development profile(s), has positive effects in terms of demographic trends, but it is incompatible with the land uses foreseen by municipal plans, although there is sufficient infrastructural potential to support urban and industrial growth induced by the new airport.
L6	The siting of the new airport is in disagreement with (contradicts) the municipal development profile(s), but it has positive effects in terms of demographic trends, is compatible with the land uses foreseen in municipal plans, and there is sufficient infrastructural potential to support urban and industrial growth induced by the new airport.
L5	The siting of the new airport is in disagreement with both the municipal development profile(s) and the land uses foreseen by municipal plans, but it has positive effects in terms of demographic trends and there is sufficient infrastructural potential to support urban and industrial growth induced by the new airport; or The siting of the new airport is in disagreement with the municipal development profile(s) and there is insufficient infrastructural potential to support urban and industrial growth induced by the new airport, although it has positive effects in terms of demographic trends and is compatible with the land uses foreseen in municipal plans.
L4	The siting of the new airport is in disagreement with the municipal development profile(s), there is insufficient infrastructural potential to support urban and industrial growth induced by the new airport, and it is incompatible with the land uses foreseen in municipal plans, although it has positive effects in terms of demographic trends.
L3	The siting of the new airport is in disagreement with the municipal development profile(s), it has negative effects in terms of demographic trends, although it is incompatible with the land uses foreseen in municipal plans and there is sufficient infrastructural potential to support urban and industrial growth induced by the new airport.
L2	The sitting of the new airport is in disagreement with the municipal development profile(s), it has negative effects in terms of demographic trends, and it is incompatible with the land uses foreseen in municipal plans, but there exists enough infrastructural potential to support urban and industrial growth induced by the new airport. or The siting of the new airport is in disagreement with the municipal development profile(s), it has negative effects in terms of demographic trends, and there is insufficient infrastructural potential to support urban and industrial growth induced by the new airport, although it is compatible with the land uses foreseen in municipal plans.
L1	The siting of the new airport is in disagreement with the municipal development profile(s), has negative effects in terms of demographic trends, is incompatible with the land uses foreseen in municipal plans, and there is insufficient infrastructural potential to support urban and industrial growth induced by the new airport.

Let us explain in detail how this pairwise comparison process evolved in the case being described. This would clearly require a significant amount of time, given that 45 judgements would be needed. In practice this was not so. The easiest way to start is by performing a dominance analysis (see Figure 7). All pairs of profiles x and y such that x (strictly) dominates y were identified (x strictly dominates y if and only if y is not more attractive than x in any dimension and x is more attractive than y in at least one dimension). All dominance relations were confirmed by the evaluator, which means that, in this case, each pair of dimensions are mutually ordinal preferential independent (which is not necessary always true). A "P" was then introduced in the respective cells of the pairwise comparison table of the software, totalling 32 P's, as shown in Figure 7. As expected, the software points out that the dominance judgments are consistent with the existence of a ranking, although they are not enough to derive it. Therefore, 13 comparisons still remained to be done. These holistic judgements can require significant cognitive effort, involving mental trade-offs, and this is precisely the reason why inconsistency between different judgements can naturally occur. To facilitate the task, the profiles were reordered by decreasing order of the number of "P's" in the respective rows and columns (see Figure 8. a).



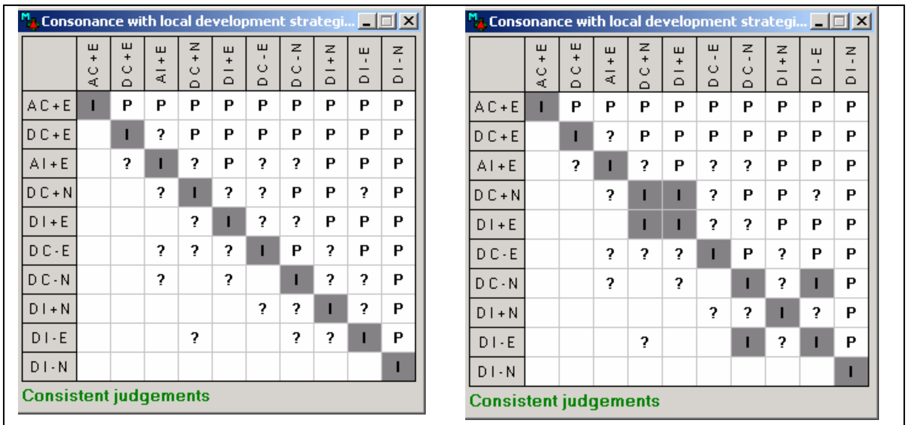
	AC+E	AI+E	DC+E	DC+N	DC-E	DC-N	DI+E	DI+N	DI-E	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
AI+E		I	?	?	?	?	P	P	P	P
DC+E		?	I	P	P	P	P	P	P	P
DC+N		?		I	?	P	?	P	?	P
DC-E		?		?	I	P	?	?	P	P
DC-N		?				I	?	?	?	P
DI+E				?	?	?	I	P	P	P
DI+N					?	?		I	?	P
DI-E				?		?		?	I	P
DI-N										I

Consistent judgements

Figure 7. Pairwise comparison process: dominance judgements.

Now, the best facilitation strategy is to look for pairs of profiles that only differ on two dimensions, and to try to identify indifference judgements. This occurred between the profiles *(D C + N)* and *(D I + E)*, and between the profiles *(D C – N)* and *(D I – E)*. Note that this means that the pair of dimensions "Effects on demographic trends" and "Infrastructural potential" is not preferentially independent of the other two (that is the preference order for two combinations of levels of the two dimensions depends on the levels at which the other dimensions are fixed – see Keeney, 1992, section 5.2). A "P" was then introduced in the respective cells (Figure 8. b), and each of the two pairs of indifferent profiles put together (Figure 8. c). The next step was to point out to the evaluator that, for consistency reasons, the comparison between indifference profiles and any other profile should be identical. For example, as *(D C + N)* was previously judged as more attractive than *(D C – N)*, consistency requires that *(D I + E)* should also be judged as more attractive than *(D C – N)*. Generalising, this is the same as saying that the lines and columns of indifferent profiles should be completely identical (see Figure 8. d). As can be seen, now only four more comparisons need to be done, all of them requiring holistic trade-offs. These comparisons were first made for pairs of profiles differing only in two dimensions: *(D C + E)* was judged more attractive than

(A I + E), *(D C + N)* – and, consequently, also *(D I + E)* – was judged more attractive than *(D C – E)*, with no inconsistency problems arising (see Figure 8. d). The two remaining comparisons involve pairs of profiles that differ on three dimensions. Suppose that the evaluator judged *(D C – E)* to be more attractive than *(A I + E)*. This is inconsistent with previous judgements, as detected by the software, because it gives rise to cycles that contradict transitivity, as shown in Figure 9. The suggestion given by the software was followed and the judgement was reversed. Finally, *(D C – E)* was judged more attractive than *(D I + N)*, which does not lead to inconsistencies, therefore completing the ranking process.



Consonance with local development strategi...

	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI+N	DI-E	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	?	P	P	P	P	P	P	P
AI+E		?	I	?	P	?	?	P	P	P
DC+N			?	I	?	?	P	P	?	P
DI+E				?	I	?	?	P	P	P
DC-E			?	?	?	I	P	?	P	P
DC-N			?		?		I	?	?	P
DI+N						?	?	I	?	P
DI-E				?			?	?	I	P
DI-N										I

Consistent judgements

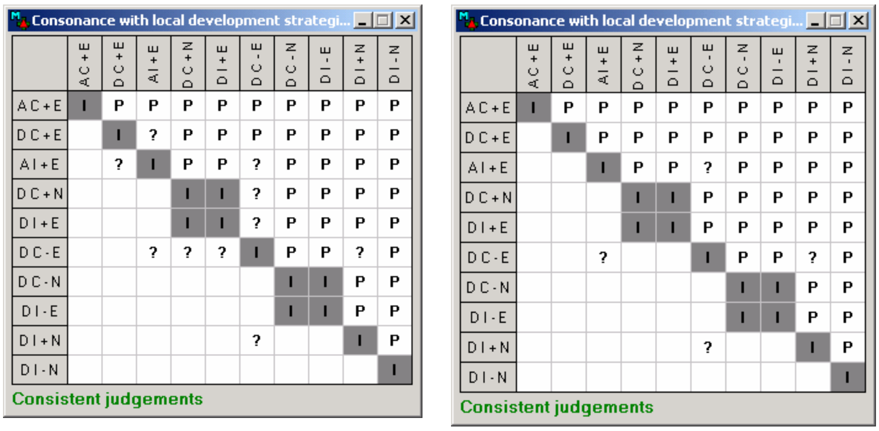
a) Reordering after dominance

Consonance with local development strategi...

	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI+N	DI-E	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	?	P	P	P	P	P	P	P
AI+E		?	I	?	P	?	?	P	P	P
DC+N			?	I	I	?	P	P	?	P
DI+E				I	I	?	?	P	P	P
DC-E			?	?	?	I	P	?	P	P
DC-N			?		?		I	?	I	P
DI+N						?	?	I	?	P
DI-E				?			I	?	I	P
DI-N										I

Consistent judgements

b) Indifference judgements



Consonance with local development strategi...

	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI-E	DI+N	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	?	P	P	P	P	P	P	P
AI+E		?	I	P	P	?	P	P	P	P
DC+N				I	I	?	P	P	P	P
DI+E				I	I	?	P	P	P	P
DC-E			?	?	?	I	P	P	?	P
DC-N							I	I	P	P
DI-E							I	I	P	P
DI+N						?			I	P
DI-N										I

Consistent judgements

c) Reordering after indifference and identical indifference lines and columns

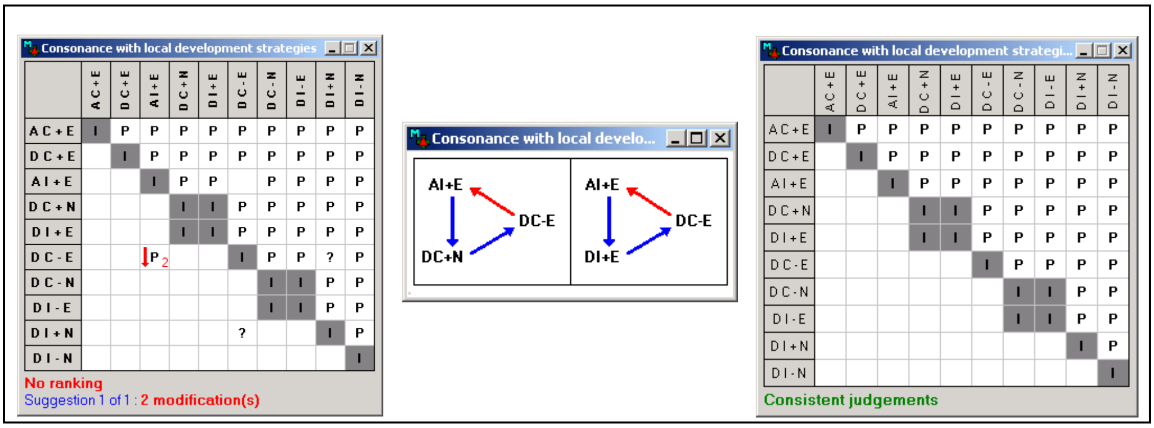
Consonance with local development strategi...

	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI-E	DI+N	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	P	P	P	P	P	P	P	P
AI+E			I	P	P	?	P	P	P	P
DC+N				I	I	P	P	P	P	P
DI+E				I	I	P	P	P	P	P
DC-E			?			I	P	P	?	P
DC-N							I	I	P	P
DI-E							I	I	P	P
DI+N						?			I	P
DI-N										I

Consistent judgements

d) Single trade-off holistic judgements

Figure 8. Development of the pairwise comparison process with MACBETH.



	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI-E	DI+N	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	P	P	P	P	P	P	P	P
AI+E			I	P	P		P	P	P	P
DC+N				I	I	P	P	P	P	P
DI+E				I	I	P	P	P	P	P
DC-E			↓P2			I	P	P	?	P
DC-N							I	I	P	P
DI-E							I	I	P	P
DI+N						?			I	P
DI-N										I

No ranking
Suggestion 1 of 1 : 2 modification(s)

	AC+E	DC+E	AI+E	DC+N	DI+E	DC-E	DC-N	DI-E	DI+N	DI-N
AC+E	I	P	P	P	P	P	P	P	P	P
DC+E		I	P	P	P	P	P	P	P	P
AI+E			I	P	P	P	P	P	P	P
DC+N				I	I	P	P	P	P	P
DI+E				I	I	P	P	P	P	P
DC-E						I	P	P	P	P
DC-N							I	I	P	P
DI-E							I	I	P	P
DI+N									I	P
DI-N										I

Consistent judgements

Figure 9. Inconsistent judgement, cycles and judgemental revision.

Using rules to develop multidimensional scales: the determinants technique

Even if the number of reference levels of each dimension is small, it still can happen that the total number of combinations is very high, making it unrealistic or undesirable to define all the possible combinations and follow the procedure suggested in Table 5.

In such cases, it could be interesting to adopt a “two-dimensions-at-a-time” approach, consisting of a series of factorial designs, each one involving two dimensions only (see Green and Srinivassan, 1978). The aim is to generate impact-profiles which differ only in two attributes, keeping all remaining dimensions at fixed levels. The pitfall of this approach is that it requires each pair of dimensions to be preferentially independent of the remaining ones. This is far from being the usual case, as revealed by the indifference judgements of the case just described above. And, if the independence condition held, why not then take each of the dimensions as a key-concern? The argument of avoiding “error accumulation” when defining too many key-concerns does not compensate for the violation of a fundamental methodological principle.

An alternative would be to opt for a “fractional design”. Fractional designs are used to reduce the number of combinations to a subset of combinations. As this means a reduction in the number of the samples of the decision space, a careful scheme is important in order to maintain an insight into the complete range of all the dimensions of the key-concern. Fractional designs are commonly used in experimental design in statistics, and a number of approaches are available (*cf.*, among others, Winer *et al.*, 1991; Cochran and Cox, 1957). The orthogonal design is a typical fractional design used for holistic judgement in multidimensional contexts (Currim and Sarin, 1983 and 1984). Each level of each dimension is combined with every other level of the other dimensions once and only once. Suppose a cluster of four dimensions, each one described by three levels (see the example in Barron and Person, 1979). The application of the orthogonal design leads to 9 reference impact-profiles (against the 3^4 =81 of the corresponding full-factorial design). The problem with using this procedure, as an initial step to developing a multidimensional descriptor, is that nothing ensures that a significant number of the selected combinations is not actually unrealistic and or indifferent, which would give rise to a descriptor formed by a number of reference impact-profiles (the few remaining plausible combinations) not sufficiently spread along the multidimensional range.

What can the facilitator do in such cases? An interesting procedure is to establish a few reference multidimensional impact-levels indirectly, that is defined by some empirical rules. There exist two basic types of rules: compensatory and non-compensatory.

Let as take the "example panel 5.2" in (Belton and Stewart, 2002, p. 130), in which the issue is to define a "quality of life" scale. This key-concern is decomposed into six dimensions – climate, standard of living, ease of adaptation to culture, quality of social/ cultural life, quality of the environment (pollution, noise, etc.), safety considerations (crime level, etc.) – "each of which is rated as: unfavourable, acceptable or favourable" (verbal descriptions of these reference impact-levels are offered by the authors). Instead of exhaustively analysing all the 3^6 possible combinations, a mixture of (*ad hoc*) compensatory and non-compensatory rules was used to define the following five-levels scale (from most attractive to least attractive):

- "All factors are favourable
- Balance of factors is better than all acceptable
- All factors are acceptable or at most one unfavourable factor may be balanced by a favourable factor
- Balance of factors is worse than all acceptable
- No factors are favourable and three or more factors are unfavourable."

Note that this is not equivalent to establishing an arbitrary scoring system (for instance to assign the value 1, 0 and –1 to each reference level and sum the scores). Nevertheless, the problem is that the multidimensional descriptor thus constructed does not necessarily possess one of the fundamental characteristics that a qualitative scale should have: to be "justifiable: an independent observer could be convinced that the scale is reasonable" (Belton and Stewart, 2002, p. 129).

A step forward in this direction is to start by clearly defining combination rules, as suggested by (Bana e Costa *et al.*, 2002). They report the extensive use of the so-called "Determinants Technique" in the framework of facilitating the evaluation of bids in a public call for tenders, to construct descriptors for key-concerns involving many elementary aspects, like "technical quality". It consists of the sequential application of rules like:

1. Establish two reference levels, "satisfactory (+)" and "neutral (o)", in each one of the dimensions;
2. classify each dimension as "determinant" (D), "important" (I) or "secondary" (S). A dimension will be "determinant" if an impact being negative (worse than neutral) in that dimension is a necessary and sufficient condition for an option to be considered negative (worse than neutral) in the dimension (this means that a determinant dimension has a non-compensatory nature);
3. define a "good" level of the descriptor by, for instances: all its determinant dimensions are satisfactory and, a majority of its important dimensions are satisfactory; define a "neutral" level by, for instances: "a majority of its determinant and important dimensions are neutral, without any dimension negative";
4. use similar rules to define at least one level better than "good" (for instances, "all determinant and important dimensions are satisfactory, without any secondary characteristic negative), and one level worst than "neutral" (for instances, "at least one determinant characteristic is negative"). Of course, other intermediate levels can be defined whenever appropriate.

Constructed indices

A constructed index is an analytical combination of two or more quantitative variables. A vast literature on indices is available especially in the fields of environmental management and protection.

The purpose of an index is to simplify, offering a parsimonious representation that conveys the necessary meaning (see Beinat 1994b and 1997; Ott, 1978). An index maps one or more variables into a single number that retains the necessary meaning. They are always a compromise between scientific accuracy and concise information (Gilbert and Feenstra, 1994).

As an example, the U.S. Environmental Protection Agency (1999) suggests the use of the Air Quality Index (AQI) to measure and report air quality in urban areas. The index is calculated for individual pollutants first and then the maximum value (the worst case) of all pollutants is used to assess air quality. The formula for each pollutant is the following (see U.S. Environmental Protection Agency, 1999, p. 16):

$$I_P = I_{LO} + (I_{HI}-I_{LO})(C_P-BP_{LO})/(BP_{HI}-BP_{LO})$$

where: I_p = the index for pollutant_p
I_{LO} = the AQI value corresponding to BP_{LO}.
I_{HI} = the AQI value corresponding to BP_{HI}
C_P = the rounded concentration of pollutant $_P$
BP_{HI} = the breakpoint that is greater than or equal to C_P
BP_{LO} = the breakpoint that is less than or equal to C_P

An accompanying table provides for each pollutant considered for the AQI (ground-level ozone, particle pollution or particulate matter, carbon monoxide, sulfur dioxide, and nitrogen dioxide) the breakpoint values used in the AQI formula.

The AQI is mainly a tool for an understandable, meaningful, and easy to compute air quality index suitable for communication between the media and the lay public. “The U.S. Environmental Protection Agency (EPA) and others are working to make information about outdoor air quality as easy to understand as the weather forecast. A key tool in this effort is the Air Quality Index, or AQI. EPA and local officials use the AQI to provide you with simple information on local air quality, the health concerns for different levels of air pollution, and how you can protect your health when pollutants reach unhealthy levels” (http://www.epa.gov/airnow/aqibroch/aqi.html).

Babcock and Nagda (1972) proposed an earlier version of the AQI : the so-called ORAQI air pollution index for carbon monoxide, sulphur oxides, nitrogen dioxide, oxidants and particulate matter ($x_1,....,x_5$). The analytical expression of the index is:

$$ORAQI(x_1,...,x_5) = \left(5.7 \cdot \sum_{i=1}^{5} \frac{x_i}{x_i^{st}}\right)^{1.37}$$

where x_i are the pollution concentrations and x_i^{st} are the corresponding standards. The ORAQI maps air pollution profiles into dimensionless scores which increase as pollution levels increase. The coefficient 5.7 and the exponent 1.37 are scaling constants which anchor the index at the value of 10 when the concentrations are at the background levels and at the value of 100 when they reach the standard simultaneously. The exponent 1.37 also indicates that pollution effects increase more than proportionally. In addition, there is some laboratory evidence to show that the adverse effects of these air pollutants are linked through an additive-exponential rule.

An index often embodies scientific knowledge and values. For instance, the index above is anchored between the background concentrations and the standard levels. However, the combinations (x_1^{st}, x_2^{st}, x_3^{st}, x_4^{st}, x_5^{st}) and ($5x_1^{st}$, 0, 0, 0, 0), among many others, have the same score of 100 implying a specific substitution rate among the pollutants. It is unclear whether this substitution rate is an approximation of the interaction between the substances or is a representation of the value judgements on the effects of the substances. This outcome has often to be considered as a *result* of the analytic form used to specify the index, rather than a *known* relation among pollutants.

Another example is the index EJC = ¼.JDC+JAC (where EJC means "equivalent number of jobs created", JDC means "number of jobs during construction", and JAC means "expected number of new jobs after construction") taken as the descriptor of the key-concern "employability" in the context of the construction of a new infrastructure. The formula shows that a trade-off is necessary to establish the relative value of the two types of jobs.

As a further example, in the context of the evaluation of measures to protect Venice from high water (Collegio di Esperti Internazionali, 1998), the frequency of floods in Venice was measured by the number of times per year that the tidal peak in the Venice lagoon reaches +100 cm above the reference level at the location "Punta della Salute – Venezia". Also in this case the index requires a preliminary modelling of the flood dynamics and of the full measurement of city layout and pavement elevation bathimetry. The level +100 is reasonable, but arbitrary, and any other level between +90 and +130 could be used as most floods fall within these limits.

An interesting discussion about this type of constructed descriptors is offered by Bouyssou (1990), around the measurement of the impact in terms of noise of the construction of a new airport on the neighbouring population. He discusses several decisions that have to be made in building an adequate descriptor, and presents several formulas of different levels of complexity that can be adopted, depending on the hypotheses made about the notion of what defines a neighbour, the consideration of the distance to the infrastructure or to the air corridors, the time frame for impact appraisal, the expected variation of the numbers of the affected population in the time frame, the evolution of aircraft technology in terms of noise, etc.

In the UK the official procedure for the appraisal of transport investments was broadened in 1998 with the introduction of the New Approach to Appraisal (NATA). This introduced a series of indicators, in addition to the strict financial or cost-benefit appraisal used hitherto, in order to support the government's integrated transport policy. The appraisal is based on five general key-concerns: environment, safety, economy, accessibility and integration, each of which uses a set of both quantitative and qualitative indicators. There is no formal multicriteria analysis based on these indicators, but rather they are used to draw up an Appraisal Summary Table (AST) which is used as a transparent way of summarizing all the available information (see Vickerman, 2000 for a more detailed assessment). The indicators used do, however, present the same problems as discussed above and this can be seen with reference to two of the more difficult indicators.

Part of the economy key-concern is whether a scheme aids the economic regeneration of an area, but measuring this in advance is problematic since actual regeneration depends on much more than transport, which is likely to be a facilitator of economic development rather than an independent determinant and which runs the danger of double-counting other more direct effects. In the absence of any more rigorous analysis a simple indicator is used which identifies whether:

- a proposal is potentially beneficial for designated regeneration areas; and,
- there are significant developments *within or adjacent to the regeneration area* which are likely to be dependent upon the proposal being approved.

Proposals are categorised as having a potentially beneficial effect if they satisfy two criteria:

- Firstly, the proposal should serve an area with recognised regeneration status (such as Assisted Area, Single Regeneration Budget, European Structural Fund) such that either:
    - the road which the proposal would affect passes through the regeneration area and there is, or will be, a junction(s) within that area; or
    - an access road from the key development site within the area feeds or is intended to feed directly onto the road; and
- Secondly the scheme is consistent with the strategy to achieve the local regeneration objectives, as outlined in evidence from the relevant Government Office.

This gives a simple yes/no indicator, but the guidance indicates that this should be supported by descriptive evidence on the socio-economic characteristics of the regeneration area, along with a view on the extent to which the proposal contributes to the local regeneration plan and the nature of any investments dependent upon the scheme along with the number of associated jobs. Even with this added evidence it remains a very crude way of trying to represent a potentially major impact of a scheme which simply biases investments towards designated regeneration areas (regardless of their actual impact) and away from non-regeneration areas (where their overall impact could be much more substantial). A scheme within a regeneration area can include the impact on adjacent areas, but a scheme in an adjacent area cannot include the possible impacts on the regeneration area.

A second indicator, under the accessibility key-concern, is whether a new scheme contributes to a reduction or increase in the degree of severance experienced by a community, particularly to pedestrians. The indicator uses a three point assessment scale: slight, moderate and severe/substantial and includes estimates of the numbers of people who may be affected either positively or negatively. An assessment score is derived as follows (in each case, the assessment is *beneficial* if relief is greater than new severance, *adverse* otherwise):

- the overall assessment is likely to be *Neutral* (if new severance is broadly balanced by relief of severance) or *Slight* where:
    - new severance and/or relief from existing severance is *slight*
    - or the total numbers of pedestrians affected across all levels of severance is *low* (less than 200 per day, say)
- the overall assessment is likely to be *Large* where:
    - new severance and/or relief from existing severance is *severe/ substantial*, and affects a moderate or high number of pedestrians

    - or the total numbers of pedestrians affected across all levels of severance is high (greater than 1000, say)
- the overall assessment is likely to be *Moderate* in all other cases.

Note that this indicator depends on the analyst's judgement of severity and an arbitrary indication of the numbers of people affected. There is a very arbitrary trade-off in which however severe the impact of the scheme, if it affects less than 200 people it is deemed to have a slight overall impact and however slight the impact if it affects more than 1000 then it has a large overall impact.

9 Impact analysis

Estimation of impacts

The definition of the descriptors has to bear in mind that they must be operational for the estimation and analysis of the impacts of options. This assumes that, for each one of the key-concerns, it will be possible to appraise the impacts, for each of the involved policy-units, of implementing each option. Formally, the "impact" $g_j(a)$ of an option a in terms of the key-concern j is a sub-set of the set of impact-levels that define the respective descriptor X_j. All impacts together form what we will call the "impacts-profile" ($g_1(a)$, …, $g_j(a)$, …, $g_n(a)$) of the option (in terms of the family of n key-concerns).

There is little practical sense in selecting a descriptor, even if apparently the "best" one, if afterwards it is not ultimately suitable for the estimation of the impacts of options, due to the lack of information, or time, or any other necessary resources. An example of this difficulty was operationalising the key-concern "air pollution effects" in the Lisbon road case (see Bana e Costa, 2001). A natural descriptor would be the "number of people affected by pollutant emissions", in each policy unit, if a given package of new roads is constructed. This would require the use of a model to estimate pollutant emissions. However, neither was a rigorous mathematical model available, nor did financial resources exist to purchase one, and, moreover, such a rigorous impact estimate was not seen as necessary. Therefore, as shown in the fourth line Table 9, a much simpler qualitative descriptor was constructed: {"decrease in emissions", "no significant change in emissions", "increase in emissions"} and the impacts of the options were defined by an expert analysing the new roads included in each package.

In conclusion, the descriptor definition is only complete if a procedure for estimating the impacts of the options is also associated with it. Such a procedure can consist of the use of a mathematical (optimisation, simulation, etc.), econometric, or technical model or formula, an empirical rule, a panel of experts, a survey technique, an experimental process, etc. (see examples and discussions in Roy, 1996, and Hobbs and Meier, 2000); the last column of Table 9 summarises the impact estimation procedures followed in the Lisbon road case.

The most desirable result of the impact estimation process is when each impact of each option in each key-concern is a single impact-level of the respective descriptor (see Table 10). However, for certain key-concerns, this was actually a simplification.

As a matter of fact, in general, the estimation of impacts may be substantially affected by different types of uncertainty (Roy, 1989, distinguishes "inaccurate determination, uncertainty and imprecision"), making it difficult, or even impossible, to describe the impact of an option on a key-concern in terms of a unique impact-level of its descriptor. Typical situations are when impacts are dispersed through time and/or through space, or can only be estimated with explicit reference to diverse scenarios.

Table 9. Descriptors of the key- concerns, reference levels, and impact estimates of the Lisbon road case (adapted from Bana e Costa, 2001).

Key-concerns	Descriptors	Reference levels			Impact estimates
		Least attractive	neutral	Most attractive	
C1 Accessibility	X1 Average travel time decrease (%)	-5	0	20	At the municipal level: Ratio between the estimate of the travel time gains on all the trips with origin on that municipality that result from the implementation of each package of projects, and the estimate of travel time for the status quo. At the regional level: Ratio between the sum of travel time gains of all municipalities and the sum of travel times for the status quo.
C2 Connectivity	X2 Shortest paths distance decrease (%)	0	0	20	The basis for these computations is the difference between the shortest paths matrices for the status quo and the matrices of each package. This results in a difference matrix for each option. The connectivity value of each municipality is the sum of its line; in other words, the sum of the reduction on all paths with origin or destination on that municipality. The connectivity value for the whole network is the sum of the values of all municipalities.
C3 Noise Pollution	X3 Increase in people exposed to a noise level > 65 dbA (thousand inhabitants)	60	0	0	On a GIS platform, a buffer of 150m and one of 500m around the roads were defined. The buffer area is then multiplied by population density for each municipality. The value for the region is the average of the values for the municipalities.
C4 Air Pollution	X4 HC and CO: changes in emissions Qualitative scale: + Decrease o No significant change – Increase	–	o	+	Emissions are a function of volume of traffic and speed. For each municipality the impact for each package (decrease, no change, or increase) was defined by an expert by analysing the new roads included in the package. The impact for the LMR considers all new roads of the package.
C5 Urban land use	X5 Urban land use affected (%)	8	0	0	On a GIS platform, a buffer of 50m around the roads was defined. Urban uses were calculated for each municipality. The value for the region is the sum of the values for the municipalities.
C6 Non-urban land use	X6 Non-urban land use affected (%)	5	0	0	On a GIS platform, a buffer of 50m around the roads was defined. The non-urban uses were calculated for each municipality. The value for the region is the sum of the values for the municipalities.
C7 Urban development potentialities	X7 Increase in expected dwellings in Master Plans (%)	0	0	60	On a GIS platform, taking into account the rates of urban development presented in each municipal Master Plan, the potential urban areas in radius of 1 from each road junction was measured for each municipality. The values for LMR are the sum of the municipal areas.

Table 10. Crisp impacts-profiles of six alternative packages of news road of the Lisbon road case (adapted from Bana e Costa, 2001).

Key-concerns	Impact units	Direction of preference	*Status quo* (neutral)	Options (packages of new roads)					
				P1	P2	P3	P4	P5	P6
C1 – Accessibility	percentage	maximise	0	4	3	4	4	3	6
C2 – Connectivity	percentage	maximise	0	3	3	3	3	4	5
C3 – Noise pollution	thousand inhabitants	minimise	0	9.4	10.1	10.1	6.9	7.5	8.4
C4 – Air pollution	qualitative scale	best: +	O	+	+	+	+	+	+
C5 – Urban land use	percentage	minimise	0	1.6	0.8	0.8	0.8	0.8	0.8
C6 – Non-urban land use	percentage	minimise	0	0.4	0.3	0.4	0.6	0.3	0.4
C7 – Urban development potentialities	percentage	maximise	0	21	21	19	26	24	19

Bouyssou (1989) offers an interesting discussion of what he considers to be the four main sources of uncertainty: “the ‘map’ is not the ‘territory’” (meaning that an impacts-profile of an option can not be taken as the option in itself), “the ‘future’ is not a ‘present’ to come” (referring to the classical notion of uncertainty in Decision Analysis), “the data are not the result of exact measurement”, and “the model is not the description of a real entity independent of the model” (highlighting, in a constructive perspective, that the way the modelling process is conducted influences the model in itself).

Figure 10 shows the impact profile of an option a in a hypothetical situation involving four key-concerns. The descriptors X_1 and X_3 are continuous scales, while X_2 and X_4 are discrete scales: $g_1(a)$ and $g_2(a)$ are crisp impacts, $g_3(a)$ is an interval (for example, experts’ judgements like the impact of option a is between the level x and y of X_3) and $g_4(a)$ is a sub-set of (not necessarily consecutive) impact levels (for example, expert 1 estimates that the impact of a is equivalent to level N_x of X_4, whereas the estimate for expert 2 is level N_{x+2}). Both $g_3(a)$ and $g_4(a)$ highlight that some form of imprecision or uncertainty can be present in impact estimates. In such cases, it is useful that some qualitative or quantitative indication of the likelihood of the estimate be specified (like, for instances, a subjective probability). The extent to which it is worthwhile to formalise and detail the modelling of uncertainty phenomena is indissoluble from what is a “requisite model” (*cf.* Phillips, 1984) in the specific decision-context: that is, what is the necessary and sufficient information that should be incorporated in the evaluation model to resolve the issues at hand.[9]

Impact analysis can yield important insights about the relative and intrinsic attractiveness of the options. Relative analysis consists in comparing impacts of different options, namely using the concept of “dominance” discussed earlier (see Figure 7); analysis of intrinsic attractiveness can be technically reduced to a relative analysis consisting of the individual comparison of each option’ impact with reference profiles.

---

[9] See (Roy, 1996, section 8.2) for a deep discussion of alternative ways to model “imprecision, uncertainty, and inaccurate determination” in impact estimation.

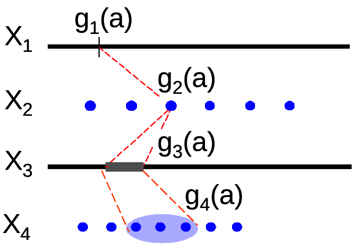


Figure 10. Types of impacts.

Relative impact analyses

The simplest relative impact-analysis is strict dominance analysis: option x strictly dominates y if and only if the impact of y is not more attractive than the impact of x in any key-concern and the impact of x is more attractive than the impact of y in at least one key-concern. For instances, referring to Table 10, it is easy to see that packages P5 and P6 both (strictly) dominate package P2 and package P6 also (strictly) dominates package P3. Although this is a poor result, one can however conclude that packages P2 and P3 can never be the most attractive of the six alternatives, as each one of them is dominated by at least another package for the given family of key-concerns. Packages P1, P4, P5 and P6 are non-dominated, that is, they are "efficient" options; therefore, for the given family of key-concerns, one of these four should be the best among the six alternatives.

If uncertainty is present, another type of pairwise comparison of options, called "significant dominance" (in Hobbs and Meier, 2000), can be very useful. It consists, for a given key-concern j, in defining two thresholds, an indifference threshold (q_j) below which the difference of impact is not significant enough to differentiate between the options, given the uncertainty affecting impact estimation, and a preference threshold (p_j) above which the difference is considered significant enough to state that one option is better than the other, even if the impacts are uncertainly defined. Then, a significantly dominates b if the impact of b never exceeds the impact of a by more than q_j, in all key-concerns j, and the impact of a exceeds the impact of b by more than p_j in at least one key-concern j. For a given key-concern j, the two thresholds q_j and p_j and the impacts $g_j(a)$ of all options a define what Roy (1996, p. 192) called a pseudo-criterion. Strict dominance corresponds to the particular case $q_j = p_j = 0$, for all j, which is actually only appropriate when no uncertainty is present. Suppose that uncertainty was present in the procedure followed in the Lisbon road case to estimate the impacts of the options in terms of "urban development potentialities (key-concern C7 in Table 9), to such a degree that it would not be realistic to differentiate impacts differing by less than 3%: under this condition, option P6 would now "significantly dominate" P1, contradicting the conclusion that P1 is an efficient option.

It may happen that the recommendation of a best option can be drawn just from comparisons of impacts-profiles; or, at least, it can possibly highlight a small sub-set of best options, very much simplifying the choice task. For example, the dominance analyses above would justify the elimination of P1, P2 and P3 (the dominated options) from the choice set. Moreover, the number of key-concerns could also be restricted to a sub-family. Suppose that the indifference threshold of 3% was valid for all impacts estimated in percentage units in Table 10 (that is, the impacts in terms of the key-concerns C1, C2, C5, C6 and C7): the problem could then be significantly simplified, for one could discard C2, C5, and C6 (and, of course, C4) for all packages would be "indifferent" in terms of each of these four key-concerns. The initial multicriteria problem represented in Table 10 could be reduced to the analysis of the simplified problem, represented in Table 11, involving only two options and three criteria. Note also that, now, package P4 would (strictly) dominate package P5, which could then be eliminated. Moreover, P4 and P6 would be indifferent in C1 (given the indifference threshold of 3%), so C1 could be removed. Finally, P4 could be taken as the “best” option as it would dominate P6. Actually it was not this simple, because the impact estimates were much more reliable!

Table 11. Sample of simplified impact-table as a result of dominance analyses.

Key-concerns	Efficient options		
	P4	~~P5~~	P6
~~C1 (minimise)~~	~~4%~~	~~3%~~	~~6%~~
C3 (minimise)	6900 inhab.	~~7500 inhab.~~	8400 inhab.
C7 (maximise)	26%	~~24%~~	19%

Care should be taken, however, in eliminating a dominated option from further analysis, because it can be a “brilliant second-best option” that may turn out to be the best one at the end of the day, not only if a new key-concern favourable to that option is included later on in the family, but also because sensitivity analyses around uncertain impacts can reverse the result of the initial strict dominance analysis.

It is worth mentioning that trade-off analysis of non-dominated options can also provide useful lessons about the relative attractiveness of the options. Suppose that, in Table 11, the impacts of P4 and P6 in C3 were actually reversed, that is 8400 and 6900 inhabitants, respectively. P4 would no longer significantly dominate P6. However, to resolve the choice problem nothing more would be needed than to judge if the decrease from 26% to 19% in C7 would be worthwhile or not against an increase of 1500 inhabitants, from 6900 to 8400, affected by noise pollution (C3).

The wise use of dominance and trade-off analyses to eliminate options and criteria step-by-step is the core of the development of the so-called “Even Swaps” procedure proposed in (Hammond *et al.*, 1998).

Intrinsic impact-analysis: definition of levels of reference “neutral” and “good”

Intrinsic impact-analysis refers to the comparison of the impacts-profile of an option with some references of intrinsic value. For instance, identifying a “neutral” level (an impact level that is neither attractive nor unattractive) in the descriptor of each key-concern enables the highlighting of the “pros” and “cons” of each option by simply analysing if its impacts are more or less attractive than the neutral levels in the several key-concerns.

For example, in the Lisbon road case, the reference neutral profile was taken as the impacts-profile of the *status quo*, shown in Table 10. The comparison of the impact-profiles of the packages with the neutral profile provides qualitative knowledge about the acceptability of each package for each policy unit.

Another useful intrinsic reference is the “good-profile”, which is often formed by aspiration levels in the several key-concerns. Comparing the impacts of an option with these references is used to appraise the extent to which the option enables the achievement of the desired levels of performance.

As mentioned in (Bana e Costa *et al.*, 2002; see also Bana e Costa *et al*., 2003), there are at least three more reasons for recommending the identification of two reference levels (“good” and “neutral”) with intrinsic value in the descriptor of each key-concern, that operationalise the idea of a good option and a neutral option:

1. Experience has revealed that the effort required to identify good and neutral levels contributes significantly to the intelligibility of the key-concern. It is one thing to say that an option is better than another in price, for example, yet, it is quite another to specify what is meant by a good price or a neutral price.
2. An explicit statement regarding good and neutral levels of reference makes it possible to make objective (in the phase of evaluation) the notion of intrinsic attractiveness of each option, assigning it to one of the following categories:
   - ✓ *very attractive option*, when it is at least as attractive as the profile “good all over”;
   - ✓ *attractive option*, if it is at least as attractive as profile “neutral all over”, but less attractive than the profile “good all over”;
   - ✓ *unattractive option*, if it is less attractive then the profile “neutral all over”.

   Making the reference levels good and neutral explicitly permits the determination of the intrinsic overall value of each option, which helps avoiding situations in which an inappropriate option is chosen, even if it is the best option (in this case, the best of a set of unattractive options).
3. Defining the two reference levels good and neutral allows the use of a weighting procedure that simultaneously is valid in the theoretical framework of the application of an additive aggregation model and avoids the pitfalls of classic weighting procedures.

10 Conclusions

This paper provides an overview of techniques and methods for model structuring in public decision aiding. We focused on structuring for multicriteria models and address the part of a multicriteria decision aid exercise which arguably finds least emphasis in the multicriteria literature. In our experience of decision facilitators we often faced situations where problem structuring takes the best and largest part of entire decision aid process. In most cases, participants recognise that structuring is the most important activity, which often encompasses the entire justification and validity of the facilitation exercise. In spite of this, the vast majority of the decision aid literature focuses on ranking or weighing methods which can be applied under the assumption that a proper structuring has been already carried out. The issue is that proper structuring is not an external activity but it is an essential, if not the only necessary one, activity in any decision aid process. In many practical cases, the added value of decision aid methods boils down completely in the structuring phase. If structuring has served to address the open issues of the decision context, the options available, the concerns of those involved in a clear and comprehensible way, then it may even be argued that the following steps in the process (weighing and raking) are a formality rather than a substantial addition to the process. This is clearly a different approach to standard decision aid literature which emphasises issues such as weighting or aggregation procedures, often with mathematical sophistication and algorithmic complexity that are completely out of touch with the reality of most decision aid situations. In this paper we shift the emphasis on the earlier part of the process, which is a prerequisite from any further analysis, and that in our experience of decision facilitators, makes the difference between a useful and a formal decision aid exercise.

References

Ariëns, E.J., Simonis, A.M., Offermeier, J. (1976), Introduction to General Toxicology, Academic Press, New York.

Babcock, L.R., Nagda, N.L. (1972), Indices of air quality, in W. Thomas (ed.), Indicators of Environmental Quality, New York, Plenum, 183-197.

Bana e Costa, C.A. (1990), Structuration, Construction et Exploitation d'un Modèle Multicritère d'Aide à la Décision, PhD Thesis, Technical University of Lisbon, IST, Lisbon.

Bana e Costa, C.A. (1996), Les problématiques de l'aide à la décision: Vers l'enrichissement de la trilogie choix–tri–rangement, Recherche Opérationnelle, 30, 2, 191-216.

Bana e Costa, C.A. (2001), The use of multicriteria decision analysis to support the search for less conflicting policy options in a multi-actor context: case-study, Journal of Multi-Criteria Decision Analysis, 10, 2, 111-125.

Bana e Costa, C.A., Antão da Silva, P., Correia, F.N. (2004), Multicriteria evaluation of flood control measures: The case of Ribeira do Livramento, Water Resources Management, 18, 3, 263-283.

Bana e Costa, C.A., Corrêa, E.C. (1999), Final report of the decision aid process for the comparison of alternative sites for the new Lisbon airport, Instituto Superior Técnico, Lisbon (in Portuguese).

Bana e Costa, C.A., Corrêa, E.C., De Corte, J.M., Vansnick, J.C. (2002), Facilitating bid evaluation in public call for tenders: a socio-technical approach, OMEGA: The International Journal of Management Science, 30, 3, 227-242.

Bana e Costa, C.A., De Corte, J.M., Vansnick, J.C. (2003), MACBETH, Working Paper LSEOR 03.56, London School of Economics, London.

Bana e Costa, C.A., De Corte, J.M., Vansnick, J.C. (2005a), On the mathematical foundations of MACBETH, in J. Figueira, S. Greco, M. Ehrgott (eds.), Multiple Criteria Decision Analysis: The State of the Art Surveys, Springer, New York (409-442).

Bana e Costa, C.A., Ensslin, L., Costa, A.P. (1998), Structuring the Process of Choosing Rice Varieties at the South of Brazil”, in E. Beinat, P. Nijkamp (eds.), Multicriteria Analysis for Land-Use Management, Kluwer, Dordrecht, 33-45.

Bana e Costa, C.A., Fernandes, T.G., Correia, P.V.D. (2005b), Prioritisation of public investments in social infra-structures using multicriteria value analysis and decision conferencing: A case-study, Working Paper LSEOR 05.78, London School of Economics, London (forthcoming in International Transactions in Operational Research).

Bana e Costa, C.A., Nunes da Silva, F., Vansnick, J.C. (2001), Conflict dissolution in the public sector: A case-study, European Journal of Operational Research, 130, 2, 388-401.

Bana e Costa, C.A., Stewart, T.J., Vansnick, J.C. (1997), Multicriteria decision analysis: Some thoughts based on the tutorial and discussion sessions of the ESIGMA meetings, European Journal of Operational Research, 99, 1, 28-37.

Barron, F.H., Person, H.B. (1979), Assessment of multiplicative utility functions via holistic judgments, Organizational Behaviour and Human Performances, 24, 147-166.

Batterman, S., Amann, M., Hettelingh, J.P., Hordijk, L., Kornai, G. (1988), Optimal SO2 abatement policies in Europe: Some examples, System Analysis-Modeling-Simulation, 5, 533-559.

Beinat, E. (1994), Decision aid and decision support systems for environmental management, Syllabuses, IAL Ambiente, Pordenone (in Italian).

Beinat, E. (1997), Value Functions for Environmental Management, Kluwer, Dordrecht.

Beinat, E., Nijkamp, P., Rietveld, P. (1994a), Indices of soil quality: a multicriteria value function approach, Journal of Environmental Systems, 23, 1, 1-20.

Beinat, E., Nijkamp, P., Rietveld, P. (1994b), Value functions for environmental pollutants: a new technique for enhancing the assessment of expert judgments, Environmental Assessment and Monitoring, 30, 9-23.

Beinat, E, van Drunen, M. (1998), Spatial conflicts in transport policies: an exploration of the perspectives of regional and local actors, in E. Beinat and P. Nijkamp, Multicriteria analysis for land use management, Kluwer, Dordrecht, 139-162.

Belton, V., Stewart, T. (2002), Multiple Criteria Decision Analysis: An Integrated Approach, Kluwer, Dordrecht.

Bouyssou, D. (1989), Modelling inaccurate determination, uncertainty, imprecision using multiple criteria, in A.G. Lockett, G. Islei (eds.), Improving Decision Making in Organisations, Springer, Heidelberg, 78-87.

Bouyssou, D. (1990), Building criteria: a prerequisite for MCDA, in C.A. Bana e Costa (ed.), Readings in Multiple Criteria Decision Aid, Springer, Berlin, 58-80.

Briassoulis, H. (1995), Environmental criteria in industrial facility siting decisions: an analysis, *Environmental Management*, 19, 2, 297-311.

Brownlow, S.A., Watson, S.R. (1987), Structuring multiattribute value hierarchies, Journal of the Operational Research Society, 38, 309-318.

Buede, D.M. (1986), Structuring value attributes, Interfaces, 16, 2, 52-62.

Cochran, W.G., Cox, G.M. (1957), Experimental Designs, 2nd edition, Wiley, New York.

Collegio di Esperti Internazionali (1998), Report on the mobile gates project for the tidal flow regulation at the Venice lagoon inlets, Consorzio Venezia Nuova, Venice.

Collins, J.P., Glysson, E.A. (1980), Multiattribute utility theory and environmental decisions, Journal of the Environmental Engineering Division, 106(EE4), 815-830.

Currim, I.S., Sarin, R.K. (1983), A procedure for measuring and estimating consumer preferences under uncertainty, Journal of Marketing Research, 20, 249-256.

Currim, I.S., Sarin, R.K. (1984), A comparative evaluation of multiattribute consumers preference models, Management Science, 30, 5, 543-561.

Damásio, A.R. (1994), Descartes' Error. Emotion, Reason and the Human Brain. G.P. Putnam's Sons, New York.

DETR (2000), Multi-Criteria Analysis: A Manual, Department of the Environment, Transport and the Regions, London.

Eden, C. (2004), Analyzing cognitive maps to help structure issues or problems, European Journal of Operational Research, 159, 673-686.

Eden, C., Ackermann, F. (1998), Making Strategy: The Journey of Strategic Management, Sage Publications, London.

Eden, C., Ackermann, F. (2001), SODA – Journey making and mapping in practice, in Rosenhead, J., Mingers, J. (eds.), Rational Analysis for a Problematic World Revisited (2nd edition), Wiley, Chichester, 43-60.

Edwards, W., Barron, F.H. (1994), SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement, Organizational Behavior and Human Decision Processes, 60, 306-25

European Commission (2004), Trans-European transport network: Implementation of the guidelines 1998-2001, Luxembourg: Eur-OP.

Fischer, G.W., Damodaran, N., Laskey, K.B., Lincoln, D. (1987), Preferences for proxy attributes, Management Science, 33, 2,198-214.

Gilbert, A.J., Feenstra, J.F. (1994), A sustainability indicator for the Dutch environmental policy theme “Diffusion”: Cadmium accumulation in soil, Ecological Economics, 9, 253-265.

Green, P.E., Srinivassan, V. (1978), Conjoint analysis in consumer research: issues and outlook, Journal of Consumer Research, 5, 103-123.

Haasstrup, P. (1994), Overview of problems of risk management of accidents with dangerous chemicals in Europe, European Journal of Operational Research, 75, 488-498.

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.

Hobbs, B.F., Meier, P. (2000). Energy Decisions and the Environment: A Guide to the Use of Multicriteria Methods, Kluwer, Boston.

Keeney, R. (1988), Value-Focused Thinking and the Study of Values, in D.E. Bell, H. Raiffa, A. Tversky (eds.), Decision Analysis: Descriptive, Normative, and Prescriptive Aspects, Cambridge University Press, 465-494.

Keeney, R.L. (1992), Value-Focused Thinking: A Path to Creative Decisionmaking, Harvard University Press, Cambridge.

Keeney, R.L., Raiffa, H. (1976), Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Wiley, New York.

Kerstholt, J.H. (1992), Information search and choice accuracy as a function of task complexity and task structure, Acta Psychologica, 80, 185-197.

Kirkwood, C.W. (1997), Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets, Duxbury Press, Belmont, CA.

Kraak, M.H.S., Schoon, H., Peeters, W.H.M., van Straalen, N.M. (1993), Chronic ecotoxicity of mixtures of Cu, Zn and Cd to the Zebra Mussel Dreissena Polymorpha, Ecotoxicology and Environmental Safety, 25, 315-327.

Krischer, J. (1976), Utility structure of a medical decision-making problem, Operations Research, 24, 951-972.

Miller, G.A. (1956), The magical number seven, plus or minus two: some limits on our capacity for processing information, Psychological Review, 63, 2, 81-97.

Moriarty, F. (1990), Ecotoxicology, 2nd edition, Academic Press, London.

Nardini, A. (1998), Improving decision making for land use management: key ideas for an integrated approach based on MCA negotiation forums, in E. Beinat, P. Nijkamp (eds.), Multicriteria analysis for land use management, Kluwer, Dordrecht, 199-228.

Ott, W.R. (1978), Environmental Indices, Theory and Practice, Ann Arbor Science Publications, Ann Arbor.

Phillips, L.D. (1984), A theory of requisite decision models, Acta Psychologica, 56, 29-48.

Phillips, L.D. (1990), Requisite decision modelling for technological projects, in C. Vlek, G. Cvetkovich (eds.), Social Decision Methodology for Technological Projects, Kluwer, Dordrecht, 95-110.

Phillips, L.D., Bana e Costa, C.A: (2005), Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing, Working Paper LSEOR 05.75, London School of Economics, London (forthcoming in Annals of Operations Research).

Rosenhead, J. (1996), What’s the problem? An introduction to problem structuring methods, Interfaces, 26, 117-131.

Rosenhead, J., Mingers, J (eds.) (2001), Rational Analysis for a Problematic World Revisited, John Wiley and Sons, Chichester.

Roy, B. (1985), Méthodologie Multicritère d’Aide à la Décision, Economica, Paris.

Roy, B. (1989), Main sources of inaccurate determination, uncertainty and imprecision in decision models, Mathematical and Computer Modelling, 12, 10/11, 1245-1254.

Roy, B. (1996), Multicriteria Methodology for Decision Aiding, Kluwer, Dordrecht.

Saaty, T.L. (1980), The Analytical Hierarchy Process, McGraw-Hill, New York.

U.S. Environmental Protection Agency (1999), Guideline for reporting of daily air quality - air quality index (AQI), Office of Air Quality Planning and Standards, North Carolina.

Vickerman, R. (2000), Evaluation methodologies for transport projects in the UK, Transport Policy, 7, 7-16.

Vickerman, R., Norman, C. (1999), Local and regional implications of trans-European transport networks: The Channel Tunnel Rail Link, Environment and Planning A, 31, 705-718.

Watson, S., Buede, D.M. (1987), Decision Synthesis: The Principles and Practice of Decision Analysis, Cambridge University Press, Cambridge.

Wenstøp, F., Carlsen, A.J. (1998), Using Decision Panels to Evaluate Hydropower Development Projects, in E. Beinat, P Nijkamp (eds.) Multi Criteria Evaluation in Land Use Management, Kluwer, Dordrecht.

Wenstøp, F., Seip, K.L. (2001), Legitimacy and quality of multi-criteria environmental policy analysis: A met analysis of five MCE studies in Norway, Journal Multi-Criteria Decision Analysis, 10, 2, 53-64.

Winer, B.J., Brown, D.R., Michels, K.M. (1991), Statistical Principles in Experimental Design, McGraw-Hill, New York.

von Winterfeldt, D., Edwards, W. (1986), Decision Analysis and Behavioral Research, Cambridge University Press, Cambridge.