Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 158 (2004) 662–677
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Decision Aiding

Social multi-criteria evaluation: Methodological foundations and operational consequences

Giuseppe Munda

Department of Economics and Economic History, Building B and Institute for Environmental Sciences and Technologies, Universitat Autonoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Received 19 April 2003; accepted 12 May 2003
Available online 11 August 2003

Abstract

In order to address contemporary issues economics and decision sciences need to expand their empirical relevance by introducing more and more realistic (thus more complex) assumptions in their models. One of the most interesting research directions in the field of public economics is the attempt to introduce political constraints, interest groups and collusion effects explicitly (J.J. Laffont, Incentives and Political Economy, 2000). The main argument developed here is the proposal of the concept of social multi-criteria evaluation (SMCE) as a possible useful framework for the application of social choice to the difficult policy problems of our millennium, where, as stated by Funtowicz and Ravetz, "*facts are uncertain, values in dispute, stakes high and decisions urgent*". This paper starts from the following main questions:

1. Why "social" multi-criteria evaluation?
2. How should such an approach be developed?

The foundations of SMCE are set up by referring to concepts from complex system theory and philosophy, such as reflexive complexity, post-normal science and incommensurability.

To give some operational guidelines on the application of SMCE basic questions to be answered are

1. *How is it possible to deal with technical incommensurability?*
2. *How can we deal with the issue of social incommensurability?*

To answer these questions, using theoretical considerations and lessons learned from real-world case studies, is the main objective of the present article.

Keywords: Multi-criteria analysis; Economics; Complexity theory; Environment; Social choice; Post-normal science; Incommensurability; Ethics

E-mail address: giuseppe.munda@uab.es (G. Munda).

0377-2217/$ - see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/S0377-2217(03)00369-2

1. Methodological foundations of social multi-criteria evaluation: Complexity, post-normal science and incommensurability

The world is characterized by deep *complexity*. This obvious observation has important implications on the manner policy problems are represented and decision-making is framed. As a consequence, one may decide to adopt a reductionistic approach trying to tackle one of the many possible dimensions or simply to deal with real-world complexity. This second approach is the one adopted in the present article. My firm conviction is that any representation of a complex system reflects only a sub-set of the possible representations of it. A system is complex when the relevant aspects of a particular problem cannot be captured using a single perspective (Funtowicz et al., 1999; O'Connor et al., 1996; Rosen, 1977).

To make things more difficult, systems including humans are *reflexively* complex. Reflexive systems present two peculiar aspects: "*awareness*" and "*purpose*", both requiring an additional "jump" in describing complexity. The presence of self-consciousness and purposes (*reflexivity*) means that these systems can continuously add new relevant qualities/attributes that should be considered when explaining, describing or forecasting their behaviour (i.e. human systems are learning systems).

Moreover, the existence of *different levels and scales* at which a hierarchical system can be analyzed implies the unavoidable existence of non-equivalent descriptions of it (Giampietro, 1994). Even a simple "objective" description of a geographical orientation is impossible without taking an arbitrary subjective decision on the relevant system scale. In fact the same geographical place, e.g., in the USA, may be considered to be in the north, south, east or west according to the scale chosen as a reference point (the whole USA, a single state and so on) [1] (Giampietro and Mayumi, 2000a,b). Therefore, the problem of *multiple-identities* in complex systems cannot be interpreted solely in terms of *epistemological plurality* (non-equivalent observers), but also in terms of *ontological characteristics* of the observed system (non-equivalent observations).

The implications of scale for multi-criteria evaluation are very important. For example, in generating evaluation criteria (e.g., in evaluating the impacts building a ski infrastructure in a mountain region, who are the relevant social actors? The inhabitants of the mountain region, the potential users in urban areas or even the ecological preservationists all around the world all are reasonable answers) or in computing the impact scores (e.g., a contamination indicator has to be computed locally, or should it be computed at a larger scale? The use of hydrogen cars inside cities is clearly good at a local level, but it is not that clear at a global level, where the emissions depend on the technology by which hydrogen is produced, since hydrogen is an energy carrier and not an energy source) or in choosing the weight factors.

A consequence of these deep subjectivities is that in any normative exercise connected to a social decision problem, one has to choose an operational definition of "*value*" in spite of the fact that social actors with different interests, cultural identities and goals have different definitions of "value" (O'Neill, 1993). That is, to reach a ranking of policy options, there is a previous need for deciding about *what is important* for different social actors as well as *what is relevant* for the representation of the real-world entity described in the model.

In general, these concerns have not been considered very relevant by scientific research in the past. On the other hand, the new nature of the problems faced in this third millennium (e.g., mad cow, genetic modified organisms, ...), implies that very often when deciding on problems that may have long term consequences we are confronting issues "*where facts are uncertain, values in dispute, stakes high and decisions urgent*" (Funtowicz and Ravetz, 1991, 1994).

In this case, scientists cannot provide any useful input without interacting with the rest of society and the rest of the society cannot perform any sound decision making without interacting with the

[1] These multiple-identity/multiple-scale systems can be defined as "*Learning Holarchies*". A "*holon*" is a whole made of smaller parts (e.g. a human being made of organs, tissues, cells, atoms) and at the same time it forms a part of a larger whole (an individual human being is a part of a household, a community, a country, the global economy) (Koestler, 1969).

scientists. That is, the question of "how to improve the quality of a policy process" must be put, quite quickly, on the agenda of "scientists", "decision makers" and indeed the whole society. This extension of the "peer community" is essential for maintaining the quality of the process of decision making when dealing with reflexive complex systems. In relation to this objective Funtowicz and Ravetz have developed a new epistemological framework called "*post-normal science*", where it is possible to better deal with two crucial aspects of science in the policy domain: *uncertainty* and *value conflict*. The name "post-normal" indicates a difference from the puzzle-solving exercises of normal science, in the Kuhnian sense (Kuhn, 1962).

Post-normal science can be characterized in relation to other, complementary, scientific strategies according to the diagram in Fig. 1, which is based on two axes: "*systems uncertainties*" and "*decision stakes*". When both uncertainty and stakes are small, we are in the realm of "normal" academic science, where it is safe to rely on "codified expertise". When either uncertainty or stakes are in the medium range, then the application of routine techniques and standardized and generalized knowledge is no longer enough. In these cases, skill, judgement, and sometimes even courage, are required to adjust the "general knowledge" available to the "special situation". Funtowicz and Ravetz call this "professional consultancy", with the examples of the surgeon or the senior engineer facing a critical situation. Finally we arrive to cases, in which conclusions are not completely determined by scientific facts;

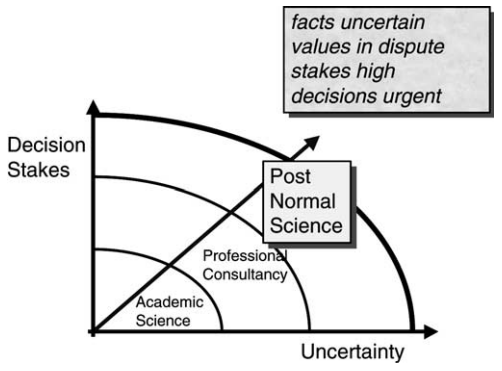

Fig. 1. Graphical representation of post-normal science.

inferences will (naturally and legitimately) be conditioned by the values held by the agents. When the stakes are very high (as when an institution is seriously threatened by a policy) then a defensive tactic will involve challenging every step of a scientific argument (this applies even to those cases in which systems uncertainties are actually small). Such a tactic should be considered wrong only when is conducted covertly, as by scientists who present themselves as impartial judges when, in reality, they are actually committed advocates of one view. When legitimate contrasting views are openly used to challenge scientific arguments, we are in the realm of post-normal science.

The previous discussion can be synthesised by using the philosophical concept of *weak comparability* (Martinez-Alier et al., 1998; O'Neill, 1993). Weak comparability implies incommensurability i.e. there is an irreducible value conflict when deciding what common comparative term should be used to rank alternative actions. Remembering that the presence of multiple-identities in complex systems can be explained in terms of epistemological plurality and in terms of ontological characteristics of the observed system, I think that it is possible to further distinguish the concepts of social incommensurability and technical incommensurability. *Social incommensurability* can be derived from the concepts of reflexive complexity and post-normal science and refers to the existence of a multiplicity of legitimate values in society, that is, just in one word to democracy. *Technical incommensurability* comes from the multidimensional nature of complexity and refers to the issue of representation of multiple identities in descriptive models.

At this point, if we accept that real-world systems are multidimensional in nature, we have also to accept that the evaluation of public plans or projects has to be based on procedures that explicitly require the integration of a broad set of various and conflicting points of view. Consequently, multi-criteria evaluation is in principle an appropriate policy framework.

For example, the concept of sustainable development has a wide appeal mainly because it does not set economic growth and environmental preservation in sharp opposition. Rather sustainable development carries the ideal of a harmonisation

or *simultaneous* realisation of economic growth and environmental concerns. Unfortunately, multi-criteria decision theory teaches us that a consequence of taking into account various dimensions simultaneously is that it is impossible to optimise all the objectives at the same time. So that we should learn how to look for "*compromise solutions*" i.e. the balance between conflicting incommensurable values and dimensions (Bromley, 1998; Faucheux and O'Connor, 1998; Munda, 1997). The arguments developed in this section imply that at least there could be two different compromise solutions: a *social compromise solution* coming from value conflicts and a *technical compromise solution* coming from conflicting non-equivalent representations of the same policy options.

At this stage, basic questions to be answered are

1. *How is it possible to deal with technical incommensurability?*
2. *How can we deal with the issue of social incommensurability?*
3. *Which are the main consequences of technical and social incommensurability in a SMCE framework?*

To answer these questions is the aim of the rest of the present article.

2. Technical incommensurability and multi/interdisciplinarity

An effective policy exercise should consider not merely the measurable and contrastable dimensions of the simple parts of the system, that even if complicated may be technically simulated (technical incommensurability). To be realistic it should also deal with the higher dimensions of the system. Those dimensions in which power relations, hidden interests, social participation, cultural constraints, and other "soft" values, become relevant, and unavoidable variables that heavily, but not deterministically, affect the possible outcomes of the strategies to be adopted (social incommensurability).

One should note that the construction of a descriptive model of a real-world system depends on very strong assumptions about (1) the *purpose* of this construction, e.g. to evaluate the sustainability of a given city, (2) the *scale* of analysis, e.g. a block inside a city, the administrative unit constituting a Commune or the whole metropolitan area and (3) the set of dimensions, objectives and criteria used for the evaluation process. A reductionist approach for building a descriptive model can be defined as the use of just *one measurable indicator* (e.g. the monetary city product per person), *one dimension* (e.g. economic), *one scale of analysis* (e.g. the Commune), *one objective* (e.g. the maximisation of economic efficiency) and *one time horizon*. If one wants to avoid reductionism, there is a clear need to take into account incommensurable dimensions using different scientific languages coming from different legitimate representations of the same system. This is what Neurath (1973) called the need for an "*orchestration of sciences*".

It is clear that a multi-criteria approach, being multidimensional in nature, seems an interesting framework to make Neurath's idea operational. A real world case study involving the water supply system of the city of Palermo in western Sicily (South Italy) can help clarifying this point. This problem was part of a project which was commissioned by the Sicily region and executed in the frame of the European Commission DGXVI structural funds. This case study was developed in two years of interaction mainly between a multidisciplinary team and the management body of the water supply system of the city of Palermo (plus some social actors involved in the final step of the study) (for more information on this case study see POP Sicily, full final report European Commission contract no. 10122-94-03 TIPC ISP I or for a shorter version Munda et al. (1998)).

Water resource management is characterized by the presence of a strong competition among different categories of consumptive water uses and, as a consequence, among various interest groups. Such a competition also exists between consumptive uses as a whole and "ecological uses" which aim at limiting water diversion for off-stream uses in order to preserve the ecological equilibrium of ecosystems. This permanent condition of competition

may become a real conflict under drought conditions, i.e. when there is a temporary reduction of available water resources due to a long and severe decrease of rainfall (compared to mean or median natural values). The problem of water shortages due to drought is particularly relevant in Southern Europe. In Sicily, the water distribution issue has deep historically routs. The *mafia* started from the fighting over water control.

Water shortages not only depend on hydrological drought which in turn follows from meteorological drought, but also depend on water supply system characteristics and demand levels, which are both affected by different drought mitigation measures. As a consequence, the pure technical hydrological solutions cannot be separated from their consequences on the socio-economic system. Although this was not evident in the beginning of the project, after a few meetings, hydrologists accepted that an economist could be of some help for this kind of problems. However, it was still very difficult to find a common language and to understand which contribution each could give to progress towards a possible solution (or at least a better understanding) of a such as complex problem.

The water system of Palermo provides water to municipal, agriculture and industrial users by using surface water and groundwater; a reservoir is also used for energy production.

It was agreed that alternative management options under drought conditions can be divided into two main groups:

- alternatives that try to satisfy 100% of the water demands,
- alternatives that do not satisfy completely the water demands.

To specify the alternatives, it was necessary to understand the structure of the Palermo water supply system and, given the technicalities involved, it was immediately clear that this was the job of hydrologists. However, these alternatives had to be evaluated for the longest historic drought experienced in the water supply system (four years) according to a set of criteria including the economic dimension (e.g. connected financial costs and benefits of the company managing the water supply system, the energy production company, and so on), the social dimension (e.g. hygienic risk and social discomfort) and the environmental dimension (e.g., the in-stream flow requirement defined as the discharge which maintains a stream ecosystem or aquatic habitat). At this point the advantage of the multi-criteria structuring of the problem was evident. Each expert suddenly knew her/his comparative advantage.

From the experience of this case study, a first lesson can be learned. The use of a multi-criteria framework is a very efficient tool to implement a multi/inter-disciplinary approach. The experts involved had various backgrounds (mainly in engineering, economics and mathematics). While in the beginning, the communication process was very difficult, when it was decided to structure the problem in a multi-criterion fashion, it was astonishing to realize that immediately a common language was created. In terms of inter-disciplinarity, the issue is to find agreement on the set of criteria to be used; in terms of multi-disciplinarity, the issue is to propose and compute an appropriate criterion score. The efficiency of the interaction process may greatly increase and its effectiveness too. [2]

In the Palermo case study, it was also experienced that taking explicitly distribution issues into account increases the transparency of the study and makes possible a process of interaction with various social actors in an effective way. This second lesson leads to the issue of social incommensurability and public participation.

3. Social incommensurability: Public participation, ethics and transparency

At this point in the discussion, one question arises, who is making the decisions? Some critics of

[2] Here I refer to the idea of orchestration of sciences as a combination of multi/inter-disciplinarity. Multi-disciplinarity: each expert takes her/his part. Inter-disciplinarity: methodological choices are discussed across the disciplines (this definition has been discussed with R. Strand).

multi-criteria evaluation say that *in principle*, in cost-benefit analysis, votes expressed on the market by the whole population can be taken into account (of course with the condition that the distribution of income is accepted as a means to allocate votes). [3] On the contrary, multi-criteria evaluation can be based on the priorities and preferences of some decision-makers only (we could say that the way these decision-makers have reached their position is accepted as a way to allocate the right to express these priorities. This criticism may be correct if a "technocratic approach" is taken, where the analyst constructs the problem relying only upon experts' inputs (by experts meaning those who know the "technicalities" of a given problem).

For the formation of contemporary public policies, it is hard to imagine any viable alternative to *extended peer communities* (Corral-Quintana et al., 2001; De Marchi and Ravetz, 2001; Funtowicz and Ravetz, 1991, 1994; Funtowicz et al., 1999; Gowdy and O'Hara, 1996; Kasemir et al., 2003). They are already being created, in increasing numbers, either when the authorities cannot see a way forward, or know that without a broad base of consensus, no policies can succeed. They are called "citizens' juries", "focus groups", or "consensus conferences", or any one of a great variety of names; and their forms and powers are correspondingly varied. But they all have one important element in common: they assess the quality of policy proposals, including the scientific and technical component. And their verdicts all have some degree of moral force and hence political influence. Here the quality is not merely in the verification, but also in the *creation*; as local people can imagine solutions and reformulate problems in ways that the accredited experts, with the best will in the world, do not find natural.

However, even a participatory policy process can always be conditioned by heavy value judgements such as, have all the social actors the same importance (i.e. weight)? Should a socially desirable ranking be obtained on the grounds of the majority principle? Should some veto power be conceded to the minorities? Are income distribution effects important? And so on.

The management of a policy process involves many layers and kinds of decisions, and requires the construction of a *dialogue process* among many stakeholders, individual and collective, formal and informal, local and not. This need has been more and more recognized in a multi-criteria decision-aid (MCDA) framework. Banville et al. (1998) offers a very well structured and convincing argumentation on the need to extend MCDA by incorporating the notion of stakeholder. This is the reason why a social multi-criteria process must be as *participative* and as *transparent* as possible; although in my opinion, participation is a *necessary* condition but not a *sufficient* one. This is the main reason I propose the concept of social multi-criteria evaluation (SMCE) in substitution of participative multi-criteria evaluation (PMCE) or stakeholder multi-criteria decision aid (SMCDA) (Banville et al., 1998). To clarify this very important point, the experience of the so-called VALSE project (see VALSE full final report, Chapter 9, European Commission ENV4-CT96-0226, or for a synthesis De Marchi et al., 2000) is instructive.

Troina is a small town (10,000 inhabitants) in the North-eastern Sicily, Italy. On the one hand, it seems there is a common assumption that there is an actual water shortage, which could be remedied by more effective use of existing resources (paradoxically, although real water shortage is common in Sicily, Troina is an exception). On the other hand, there is a complex and heterogeneous collection of interests in the Troina water issue, who have hitherto had no effective dialogue. Hence an effective structuring of the water problem at this early stage is an important task, so that eventual negotiations among social actors can have a better chance of a positive outcome. The steps of the overall evaluation process followed are schematised in Fig. 2.

One has to note that policy evaluation is not a one-shot activity; on the contrary, it takes place as a *learning process*. It has to be realised that the evaluation process is usually highly dynamic, so

[3] One should note that indeed cost-benefit analysis can be easily criticised both from the distributive and environmental points of view (see e.g., Munda, 1996; Spash and Hanley, 1995). However I prefer not to deal with this issue here.

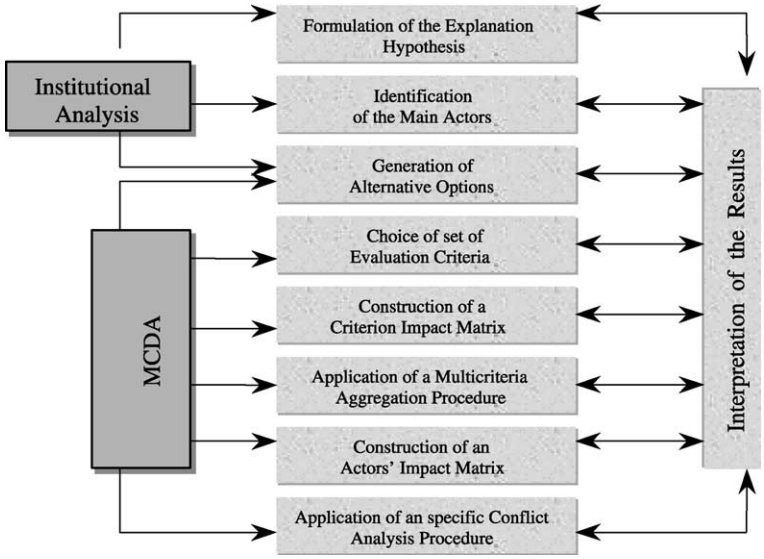

Fig. 2. Scheme of the evaluation process in the Troina case study.

that judgements regarding the political relevance of items, alternatives or impacts may present sudden changes, hence requiring a policy analysis to be flexible and adaptive in nature. This is the reason why evaluation processes have a *cyclic nature*. By this is meant the possible adaptation of elements of the evaluation process due to continuous feedback loops among the various steps and consultations among the actors involved (Nijkamp et al., 1990).

The first question to be answered is the following: *is "business as usual" a possible option in the long run?* Business as usual is a situation where power and water management are fragmented among the main actors and where infrastructure actions are the only ones not requiring agreements. This can be considered the classic case of non-cooperative resource exploitation.

For example, the Municipality of Troina is trying to become self-sufficient for its drinking water needs using its own spring water sources, even if this could be perceived as inefficient. To evaluate the business as usual option properly, it has to be compared to a set of different possible options on the basis of some evaluation criteria. At this point, an issue immediately arises: alternatives and criteria for whom? This leads to a need to take into account the preferences of some of the actors playing an important role in the problem at hand.

Initially, only the actors playing an important role in the community of Troina (as a result of the institutional analysis) were taken into account. Later on, as a surprising feedback of the process of generation of alternative options, it was clear to everybody that additional interest groups outside Troina also had to be taken into account. This learning process was very interesting particularly for the local administrators of Troina, who fully realised the importance of Troina water resources outside their own territory. As the Mayor acknowledged, such a process of structuring the problem at hand was extremely useful for understanding the hierarchy of interests that is behind the exploitation of local natural resources.

During the study, the top position of a course of action proposing an *information campaign* was an unexpected surprise. The response to this surprise was the idea of implementing, within a very short time horizon, an exposition on water management issues in the town of Troina. The Mayor and the municipal administration thought that the implementation cost of such a policy measure was quite low and the positive impacts on the community could be very high. Of course, the political risks

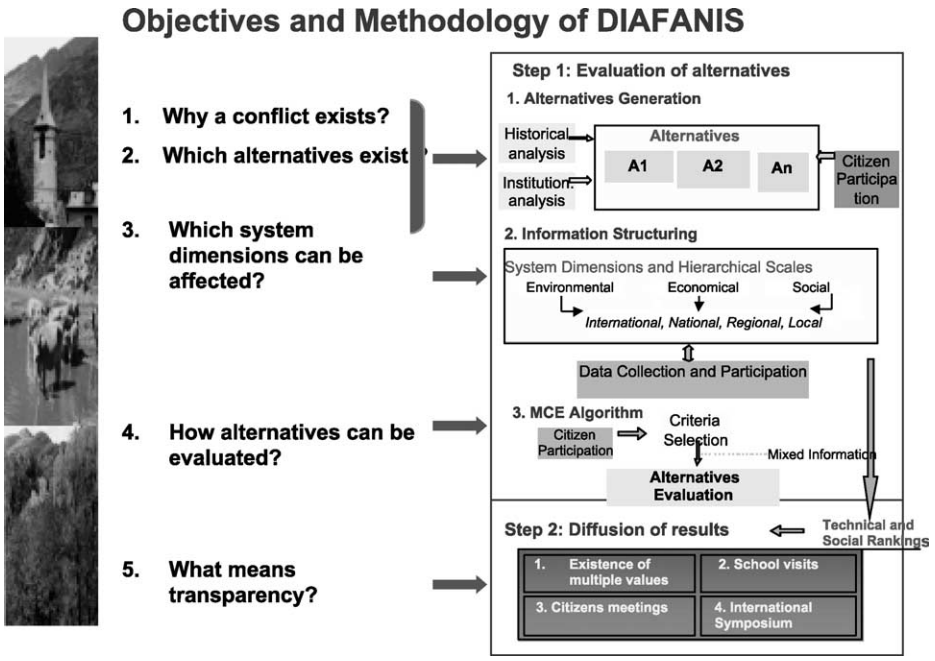

Fig. 3. Structure of the evaluation process in the DIAFANIS project.

for the administration can also be very high, since it was clear that a lot of powerful actors work hard to keep the status quo. [4] This point leads us to the initial and principal question, is business as usual a defensible option?

One should note that business as usual was ranked almost on the bottom of the multi-criteria ranking. While in the NAIADE conflict analysis (Munda, 1995), it was in a low position for some actors and in a high/medium position for all the others. Almost all the powerful social actors of Troina community belong to this second group. We could say that the status quo is a compromise solution among the opposite internal interests. This can explain why nobody is willing to change the present situation (though it is very risky for the community at large). However, this situation looks much more as an *impasse* than as a real equilibrium.

In this study it was attempted to avoid the pitfalls of the technocratic approach, by applying different *methods of sociological research*. The "*institutional analysis*", performed mainly on historical, legislative and administrative documents, provided a map of the relevant social actors. Much insight was offered by "*participant observation*" as some contributors to the study were also members of the community and knowledgeable of its internal dynamics. The possible biases of this "insider perspective" were checked against the information obtained from some "*in-depth interviews*" with key local actors. Finally a "*survey*" by using a questionnaire was performed on a random sample of the resident population, so as to explore their perception of the water issue in Troina.

To better clarify the arguments I want to develop in this section, I will refer to another case study the DIAFANIS project (project financed by the Spanish Ministry of Environment, see final report (in Spanish and Catalan) and Martí (2001) (in Catalan)). This project was named "diafanis" to indicate that the emphasis of the approach is on the transparency issue (Fig. 3).

The problem dealt with was the possible expansion of a ski infrastructure in the Catalan

[4] Actually, I must say that the Mayor and his administration lost the next elections ….

Pyrenees (north-east Spain). It was very clear from the beginning that the choice of the geographical scale would determine the policy option considered desirable. In fact, local people living close to the area think that the expansion would bring more tourists and as a consequence more economic welfare. This perception changes as long as one leaves the immediate neighborhood of the zone affected by the expansion project. Thus for example in Barcelona preservationists, since the area in question is close to a natural park and even declared by the autonomous government of Catalonia as a possible natural area of European interest, are quite against the project. Thus, which is the appropriate scale? Local people close to the affected area, all the Pyrenees area, all Catalonia or even the whole Europe?

To understand if other possible courses of actions exist, an institutional analysis was carried out and consequently some participatory techniques were undertaken. By means of focus groups it was possible to have an idea of people's desires and it was then possible to develop a set of policy options. A limitation of the focus group technique, immediately evident, was that at the local scale, some people were not willing to say publicly what they really thought, since they were afraid on the consequences for their everyday life (social exclusion in small communities is considered a tragedy, or sometimes they saw their jobs at danger, since e.g., they were working for an important hotel owner who was absolutely in favor of the ski infrastructure). For this reason anonymous questionnaires and personal interviews are an essential part of the participatory process. When far from the immediate vicinity of the affected area, this component of social control was almost not-existing.

The selection of evaluation criteria was also based on what it was learned through the participation process. However, at this stage a problem immediately arose: the evaluation criteria should come directly from the public participation process or they should be "translated" by the research team? It was soon understood that the rough material collected during interviews and focus groups could be used as a source of inspiration but the technical formulation of criteria having properties such as "non-redundancy", "legibility" and so on (see Bouyssou, 1990) is a clear job of the researchers. Of course in this step, subjectivity is unavoidable (for example, in the team there were a lot of discussions to limit the biases of some members who have strong ecologist convictions).

The same criticism of use and sometimes abuse of the subjective component of the research team can be easily done when synthesizing the impacts of the various courses of actions on the different social actors (e.g., to build the NAIADE conflict analysis procedure). This is obviously true, although the social scientists involved in the study appreciated a lot the possibility to work with an operational framework which allows synthesizing the big amount of non-formalized information collected during their field investigations.

Being conscious of the subjective and sometimes even arbitrary components inherent in the study, a widespread information campaign was planned on the assumptions and conclusions of the study including local people, regional and national authorities, international scientists and even children at school.

From these case studies some interesting lessons can be learned.

(1) In synthesis, one should not forget that the classical schematised relationship decision-maker/analyst is indeed embedded in a social framework, which is of a crucial importance in the case of public policy.
(2) The combination of various participatory methods, which has been proved powerful in sociological research, becomes even more so when integrated with a multi-criterion framework.
(3) The use of a cyclic evaluation process allows incorporating the concept of learning of the scientific team on the case study tackled. It is extraordinary important that different participatory and interaction tools are used in different points in time. This allows for continuous testing of assumptions and unavoidable biases of the study team.
(4) According to the geographical scale chosen, the relevant social actors with an interest at stake can be found thanks to institutional analysis. Institutional analysis is an essential

step to identify possible "stakeholders" for a participative process. However, besides the unavoidable mistakes that may happen in carrying out an appropriate institutional analysis, I think there are even stronger reasons why I do not believe desirable a pure participatory study.

(5) In synthesis, the scientific team cannot simply accept uncritically the inputs of a participatory process, since:
   (a) In a focus group, powerful stakeholders may influence deeply all the others.
   (b) Some stakeholders might not desire or be able to participate, but ethically the scientific team should not ignore them.
   (c) The notion of stakeholder [5] only recognises relevant organised groups; this is the reason why I prefer the term "*social actor*".
   (d) Focus groups are never meant to be a representative sample of population. As a consequence, they can be a useful instrument to improve the knowledge of the scientific team of the institutional and social dimensions of the problem at hand, but never a way for deriving consistent conclusions on social preferences.

These conclusions lead to the following personal (and thus arguable) *convictions*:

(1) Transparency is an essential component to guarantee the quality of any study based on science for policy. In fact all these studies should be *accountable* (accountability is a concept recently proposed by the European Commission in the White Paper on Governance [6]) to the public at large for peer-reviewing.

(2) Multi-criteria methods supply a powerful framework for policy analysis since this type of evaluation processes can be very effective since it accomplishes the goals of being *inter/multi-disciplinary* (with respect to the research team), *participatory* (with respect to the local community) and *transparent* (since all criteria are presented in their original form without any transformations in money, energy or whatever common measurement rod).

(3) Since decision-makers search for legitimacy [7] of the decisions taken, it is extremely important that public participation or scientific studies do not become instruments of political de-responsibility. I strongly believe that the deontological principles of the scientific team and policy-makers are essential for assuring the quality of the evaluation process. Social participation does not imply that scientists and decision-makers have no *responsibility* of policy actions defended and eventually taken.

(4) As a consequence, ethics matters. Let us imagine the extreme case where a development project in the Amazon forest will affect an indigenous community with little contact with other civilizations yet. Would it be ethically more correct to invite them in a focus group... or ethically compulsory to take into account the consequences of the project for their survival? The importance of the inclusion of ethical considerations in mathematical modelling and decision-making has been recently discussed also by Kleijnen (2001) and Rauschmayer (2000).

(5) A positive externality of participatory approaches is that sometimes the results obtained by the research team, i.e. data, findings, interpretations and insights, can also be returned to the community which may use them not as just given, but rather as an input for deliberative democracy. In synthesis a participatory approach can also simply be an educational tool to learn what democracy is.

A clear example of the difference between a participatory multi-criteria study and a social multi-criteria one can be found in the determination of criterion weights. Can we have an elicitation of weights from all the social actors involved to be used in the evaluation process?

The issue of weights in single/multi actor frameworks is a highly discussed one (see e.g. Bana e Costa, 1990; Munda, 1993; Roy, 1985, 1996;

[5] Banville et al. (1998) discuss the notion of a stakeholder and its ambiguities deeply.

[6] I owe this information to B. De Marchi.

[7] On the issue of legitimacy see also Roy and Damart (2002).

Nijkamp et al., 1990; Vansnick, 1986). Often, a decision made pragmatically is not to use weights at all when multi-actor situations are present. In fact under these circumstances normally weights are a bottleneck for the decision process. This is the solution adopted for example, in ELECTRE IV by B. Roy and following Roy by me in the NAIADE method. However, maybe this is not the most adequate solution to the problem of weights in "social" multi-criteria evaluation.

Let us start with an example; in Spain about 40 years ago, there was an important policy criterion: safety of the north frontier with France. Nowadays nobody even remembers the existence of this Franco's attitude towards frontiers. What I want to emphasise here, is the fact that policy criteria are the consequence of the social and political framework existing in a given historical period. To give another example, at the moment the environmental dimension is becoming more and more important in evaluation projects while this was almost irrelevant 40 years ago.

As we know in society there are different legitimate values and points of view. This creates social pressure for taking into account various policy dimensions, e.g. economic, social and environmental. [8] These dimensions are then translated by analysts into objectives and criteria. [9] At this point a question arises who should attach criterion weights and how? To answer this question we have to accept a basic assumption: to weigh different criteria implies to give weights to different groups in society. This assumption has the following main consequences:

1. In social decision processes, weights cannot be derived as inputs coming from participatory techniques. This is *technically* very difficult (e.g., which elicitation method has to be used? Which statistical index is a good synthesis of the results obtained? Do average values of weights have meaning at all?), *pragmatically* not desirable (since strong conflicts among the various social actors are very probable to occur) and even *ethically* unacceptable (if one accepts the arguments developed in the previous section).

2. *A plurality of ethical principles* seems the only consistent way to derive weights in a SMCE framework.

3. Weights in the framework I am proposing are clearly meaningful only as *importance coefficients* and not as trade-off (since different ethical positions leads to different ideas on criterion importance). This also implies that the aggregation conventions used should be non-compensatory mathematical algorithms (Bouyssou and Vansnick, 1986; Roberts, 1979). Non-compensability implies that minorities represented by criteria with smaller weights can still be very influent. This is for example clear in the use of the discordance index in the ELECTRE methods (Roy, 1985, 1996).

4. *Sensitivity and robustness analysis* have a complete different meaning with respect to the case of single person and technical decisions. [10] In fact in the case of SMCE, weights derive only from a few clear cut ethical positions. This means that sensitivity or robustness analysis have to check the consequences on the final ranking of only these positions and not of all the possible combinations of weights. Sensitivity and robustness analysis are then a way to improve transparency. [11]

[8] By *dimension*, here I mean the highest hierarchical level of analysis which indicates the scope of objectives and criteria.

[9] This hierarchical structure of a multi-criteria problem is similar to some extent to the one proposed by Saaty (1980). However, I use only the basic idea of hierarchy but no technical proposals of AHP.

[10] I have discussed this point with Serafin Corral-Quintana. I think that he correctly maintains that in a policy framework, sensitivity analysis should consider the willingness of social actors to implement a given course of action more than combinations of weights (Corral-Quintana, 2001). I disagree on the fact that actors should be considered in function of the power they have to support or fight a policy action. I think this has a descriptive content but not a normative one. This is the reason why I insist on the ethical dimension of a normative exercise as SMCE is.

[11] On this point I disagree with Kleijnen (2001), who claims that "modellers should try to develop robust models", in the sense that models should not be very sensitive to modellers' assumptions. Some ethical positions might be very different and thus lead to different rankings of the policy options. What is essential in a social framework is then transparency on these assumptions.

4. Conclusion: Social multi-criteria evaluation as a framework for applied social choice

The pioneering research developed by Arrow and Raynaud (1986) showed that the relationships between multi-criteria decision theory and social choice are clear and relevant. In my opinion, the main directions of cross-fertilization between these research fields are two:

1. Multi-criteria decision theory can be an adequate framework for applied social choice.
2. Social choice can supply interesting theoretical results for assuring the axiomatic consistency needed by multi-criterion aggregation conventions.

The main research issue considered in the present article has been the first one, which was not considered at all by Arrow and Raynaud. These authors explicitly state that their interest is the so-called "industrial outranking problem", whose aim is to help decisions of business-people. In my opinion the substantial meaning of multi-criteria evaluation in a social context is simply *tolerance and democracy*.

Historically, the first stage of the development of multi-criteria decision theory was characterized by the so-called methodological principle of *multi-criteria decision making* (MCDM) whose main aim is to elicit clear subjective preferences from a decision-maker and then try to solve a well-structured mathematical decision problem thanks to a more or less sophisticated algorithm. In this way a multi-criterion problem can be still presented in the form of a classical optimisation problem (Keeney and Raiffa, 1976).

Complex systems i.e. all real-world systems, present multiple possible descriptions all of them correct. Complexity is then a property of the appraisal process rather than a property inherent to the system itself. As a consequence, any model is the representation of reality resulting from a number of arbitrary assumptions, implying the existence of two or more different correct representations of the same real-world system. With these arguments I stress that, in a multi-criteria framework, what really matters is *the process* since the problem structuring will determine the result. Thus the method as such is just a framework, which of course has to be as consistent and above all transparent as possible, but please remember *a computation is not a decision*! [12] The limitations of the classical concept of an optimum solution and the consequential importance of the decision process has recently been emphasised by different authors (e.g., Clímaco, 2000; Moreno-Jiménez, 1999; O'Connor et al., 1996; Roy, 1985; Simon, 1976).

According to Simon (Simon, 1976, 1983), a distinction must be made between the general notion of rationality as an adaptation of available means to ends, and the various theories and models based on a rationality which is either substantive or procedural. This terminology can be used to distinguish between the rationality of a decision considered independently of the manner in which it is made (in the case of substantive rationality, the rationality of evaluation refers exclusively to the results of the choice) and the rationality of a decision in terms of the manner in which it is made (in the case of procedural rationality, the rationality of evaluation refers to the decision-making process itself). "*A body of theory for procedural rationality is consistent with a world in which human beings continue to think and continue to invent: a theory of substantive rationality is not*" (Simon, 1976).

Roy (1985, 1990, 1996) states that in general, it is impossible to say that a decision is a good one or a bad one by referring only to a mathematical model. Thus, it becomes impossible to find the validity of a procedure either on a notion of *approximation* (i.e. discovering pre-existing truths) or on a mathematical property of *convergence* (i.e. does the decision automatically lead, in a finite number of steps, to the optimum a*?). The final solution is more like a "*creation*" than a discovery. In *multiple-criteria decision aid* (MCDA) (Roy, 1985), the principal aim is not to discover a solution, but to construct or create something which is

[12] This difference has been pointed out to me by J. Kay. For a philosophical discussion on the concept of decision see Munda (1993).

viewed as liable to help "*an actor taking part in a decision process either to shape, and/or to argue, and/or to transform his preferences, or to make a decision in conformity with his goals*" (*constructive or creative approach*) (Roy, 1990).

This classical schematised relationship decision-maker/analyst and the related concept of "decision aid" as a learning process for the actors involved seems to me more adequate in situations such as the ones defined as applied science and professional consultancy by Funtowicz and Ravetz (see Fig. 4). Since this process seems more adequate for the search of a technical compromise solution, I call it a "*technocratic approach*".

Expansions of MCDA to the social domain have recently been attempted by various scientists (e.g., Banville et al., 1998). For the reasons I discussed earlier in this article, I think that a MCDA participatory approach is mainly under the conditions of medium uncertainty and medium decision stake ranges.

All the arguments and convictions discussed in this article have led me to the development of the concept of social multi-criteria evaluation (SMCE) whose very essence is the recognition that (see Fig. 5):

- The use of a multi-criteria framework is a very efficient tool to implement a *multi/inter-disciplinary* approach.
- Science for policy implies a *responsibility* of the scientists towards the whole society and not just towards a mythical decision-maker.
- *Public participation* is a necessary component but not a sufficient one. Participation techniques are a tool for improving the knowledge of the problem at hand and not for receiving inputs to be used uncritically in the evaluation process. Social participation does not imply lack of responsibility.
- *Ethical judgements* are unavoidable components of the evaluation exercise. These judgements always influence heavily the results. As a consequence, *transparency* on the assumptions used is essential.
- In this framework, of course mathematical aggregation conventions play an important role, i.e. to assure that the rankings obtained are *consistent* with the information and the assumptions used. [13]

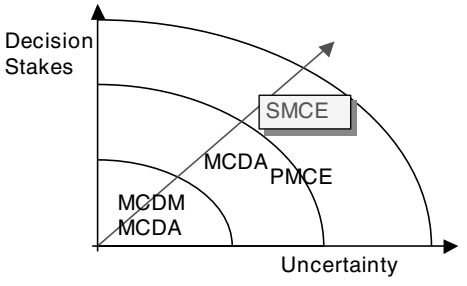

Fig. 4. Multi-criteria approaches in relationship to Funtowicz–Ravetz classification of science for policy.

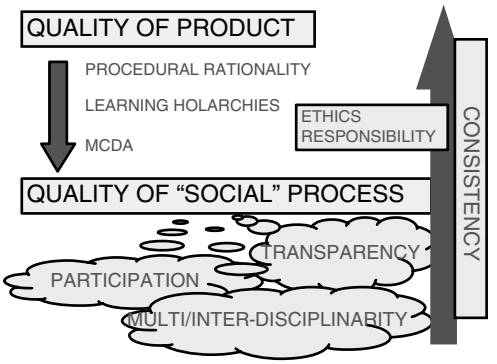

Fig. 5. Synthesis of a social multi-criteria evaluation process.

This discussion leads to the need of defining the concept of evaluation as the combination of *representation, assessment and quality check* connected to a given policy problem in relation to a given objective. [14] This is the reason why I use the term "multi-criteria evaluation" and not "multi-criteria decision" when a social context is implied.

[13] I insist on the importance of the algorithmic component in SMCE. Indeed I used the term "non-algorithmic" multi-criteria evaluation as an implementation tool for the incommensurability principle (Martinez-Alier et al., 1998). This term was intended to emphasize the importance of the decision process however I think was an unfortunate choice since it gives the impression that the algorithmic component is not useful at all.

[14] This definition has been developed thanks to discussions with M. Giampietro.

Let's finally have a look at the consequences of social and technical incommensurabilities on the so-called axiomatization issue. When different conflicting evaluation criteria are taken into consideration, a multi-criteria problem is mathematically ill-defined. The consequence is that a complete axiomatization of a multi-criterion aggregation convention is quite difficult. To deal with this problem, two main approaches can be distinguished.

1. The attempt to check under which specific circumstances each method could be more useful than others, i.e. the search of the right method for the right problem (e.g., see Guitouni and Martel, 1998).
2. The attempt of looking for a complete set of formal axioms that can be attributed to a specific method (e.g., Arrow and Raynaud, 1986; Vincke, 1994).

Here, I will try to isolate some properties that may be considered desirable for a discrete multi-criteria method in the framework of SMCE. Of course in another framework, e.g. stock exchange investments, these properties can easily be irrelevant or even undesirable. A deeper discussion on this topic can be found in Munda (2003).

The idea of *technical incommensurability* makes the following properties desirable:

- Indifference and preference thresholds should be explicitly taken into account.
- Mixed information of the widest type should be addressed in a consistent way.
- Simplicity, meaning the use of as less parameters as possible, is a very desirable property to guarantee transparency.
- The hierarchical dimension of a policy problem should be explicitly considered.

The idea of *social incommensurability* makes the following properties desirable:

- Weights in this framework are meaningful only as importance coefficients and not as trade-off. As a consequence, complete compensability cannot be implemented.
- Sensitivity and robustness analysis have to check the consequences on the final ranking of only some clear ethical positions and not of all the possible combinations of weights.
- Conflict analysis procedures explicitly looking for social compromises should integrate a SMCE exercise.
- In a policy framework, to have a complete ranking of all the alternatives is more useful than just to select one alternative only; this implies that dominated alternatives cannot be excluded *a priori*.

Acknowledgements

Comments by the anonymous referees are gratefully acknowledged. This research has been partly financed by the European Commission, research project: Development and Application of a Multi-Criteria Decision Analysis Software Tool for Renewable Energy Sources (MCDA_RES).

References

Arrow, K.J., Raynaud, H., 1986. Social Choice and Multicriterion Decision Making. MIT Press, Cambridge, MA.

Bana e Costa, C.A., 1990. An additive value function technique with a fuzzy outranking relation for dealing with poor intercriteria preference information. In: Bana e Costa, C.A. (Ed.), Readings in Multiple Criteria Decision Aid. Springer-Verlag, Berlin, pp. 351–382.

Banville, C., Landry, M., Martel, J.M., Boulaire, C., 1998. A stakeholder approach to MCDA. Systems Research and Behavioral Science 15, 15–32.

Bouyssou, D., Vansnick, J.C., 1986. Noncompensatory and generalized noncompensatory preference structures. Theory and Decision 21, 251–266.

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

Bromley, D., 1998. Searching for sustainability: The poverty of spontaneous order. Ecological Economics 24, 231–240.

Clímaco, J.C.N., 2000. Uma reflexao critica da decisao óptima. Abertura Solene das Aulas na Universidade de Coimbra, Coimbra.

Corral-Quintana, S.A., 2001. Una metodología integrada de exploración y comprensión de los procesos de elaboración de políticas públicas, unpublished Ph.D. Thesis, Univ. La Laguna, Dep. Economics, Spain.

Corral-Quintana, S., De Marchi, B., Funtowicz, S., Gallopín, G., Guimarães-Pereira, Â., Maltoni, B., 2001. The Visions Project at the JRC, European Commission, Joint Research Centre, Institute for the Protection and Safety of the Citizens, EUR 19926 EN [http://alba.jrc.it/visions].

De Marchi, B., Funtowicz, S.O., Lo Cascio, S., Munda, G., 2000. Combining participative and institutional approaches with multi-criteria evaluation. An empirical study for water issue in Troina, Sicily. Ecological Economics 34 (2), 267–282.

De Marchi, B., Ravetz, J., 2001. Participatory approaches to environmental policy. Concerted Action EVE, Policy Research Brief, No. 10.

Faucheux, S., O'Connor, M. (Eds.), 1998. Valuation for Sustainable Development: Methods and Policy Indicators. Edward Elgar.

Funtowicz, S.O., Ravetz, J.R., 1991. A new scientific methodology for global environmental issues. In: Costanza, R. (Ed.), Ecological Economics. Columbia, New York, pp. 137–152.

Funtowicz, S.O., Ravetz, J.R., 1994. The worth of a songbird: Ecological economics as a post-normal science. Ecological Economics 10, 197–207.

Funtowicz, S., Martinez-Alier, J., Munda, G., Ravetz, J., 1999. Information tools for environmental policy under conditions of complexity, European Environmental Agency, Experts' Corner, Environmental Issues Series, No. 9.

Giampietro, M., 1994. Using hierarchy theory to explore the concept of sustainable development. Futures 26 (6), 616–625.

Giampietro, M., Mayumi, K., 2000a. Multiple-scale integrated assessment of societal metabolism: Introducing the approach. Population and Environment 22 (2), 109–154.

Giampietro, M., Mayumi, K., 2000b. Multiple-scale integrated assessment of societal metabolism: Integrating biophysical and economic representations across scales. Population and Environment 22 (2), 155–210.

Gowdy, J.M., O'Hara, S., 1996. Economic Theory for Environmentalists. Saint Lucie Press.

Guitouni, A., Martel, J.M., 1998. Tentative guidelines to help choosing an appropriate MCDA method. European Journal of Operational Research 109, 501–521.

Kasemir, B., Gardner, M., Jäger, J., Jaeger, C. (Eds.), 2003. Public Participation in Sustainability Science. Cambridge University Press, Cambridge.

Keeney, R., Raiffa, H., 1976. Decision with Multiple Objectives: Preferences and Value Trade-offs. Wiley, New York.

Kleijnen, J.P.C., 2001. Ethical issues in modelling: Some reflections. European Journal of Operational Research 130, 223–230.

Koestler, A., 1969. Beyond atomism and holism: The concept of the holon. In: Koestler, A., Smythies, J.R. (Eds.), Beyond Reductionism. Hutchinson, London, pp. 192–232.

Kuhn, T.S., 1962. The Structure of Scientific Revolutions. University of Chicago Press, Chicago.

Martí, N., 2001. Processos de decisió i instrumentalització de l' avaluació d' actuacions en el territori. Una proposta metodólogica d' avaluació integrada a l' entorn del Parc Nacional d' Aigüestortes i estany de Sant Maurici: el cas de l' estudi DIAFANIS. Master Dissertation. Doctoral Programme in Environmental Sciences, Univ. Autonoma de Barcelona, Spain.

Martinez-Alier, J., Munda, G., O'Neill, J., 1998. Weak comparability of values as a foundation for ecological economics. Ecological Economics 26, 277–286.

Moreno-Jiménez, J.M., Aguaron, J., Escobar, T., Turon, A., 1999. Multicriteria procedural rationality on SISDEMA. European Journal of Operational Research 119 (2), 388–403.

Munda, G., 1993. Multiple-criteria decision aid: Some epistemological considerations. Journal of Multi-Criteria Decision Analysis 2, 41–55.

Munda, G., 1995. Multicriteria Evaluation in a Fuzzy Environment. In: Contributions to Economics Series. Physica-Verlag, Heidelberg.

Munda, G., 1996. Cost-benefit analysis in integrated environmental assessment: Some methodological issues. Ecological Economics 19 (2), 157–168.

Munda, G., 1997. Environmental economics, ecological economics and the concept of sustainable development. Environmental Values 6 (2), 213–233.

Munda, G., Paruccini, M., Rossi, G., 1998. Multicriteria evaluation methods in renewable resource management: The case of integrated water management under drought conditions. In: Beinat, E., Nijkamp, P. (Eds.), Multicriteria Evaluation in Land-use Management: Methodologies and Case Studies. Kluwer, Dordrecht, pp. 79–94.

Munda, G., 2003. An axiomatic solution for the discrete multicriterion problem in a public choice framework, unpublished manuscript.

Neurath, O., 1973. Empiricism and Sociology. Reidel, Dordrecht.

Nijkamp, P., Rietveld, P., Voogd, H., 1990. Multicriteria Evaluation in Physical Planning. North-Holland, Amsterdam.

O'Connor, M., Faucheux, S., Froger, G., Funtowicz, S.O., Munda, G., 1996. Emergent complexity and procedural rationality: Post-normal science for sustainability. In: Costanza, R., Segura, O., Martinez-Alier, J. (Eds.), Getting Down to Earth: Practical Applications of Ecological Economics. Island Press/ISEE, Washington, DC, pp. 223–248.

O'Neill, J., 1993. Ecology, Policy and Politics. Routledge, London.

Rauschmayer, F., 2000. Ethics of multicriteria analysis. International Journal of Sustainable Development 3 (1), 16–25.

Roberts, F.S., 1979. Measurement Theory with Applications to Decision Making, Utility and the Social Sciences. Addison-Wesley, London.

Rosen, R., 1977. Complexity as a system property. International Journal of General Systems 3, 227–232.

Roy, B., 1985. Méthodologie Multicritère d' Aide à la Decision. Economica, Paris.

Roy, B., 1990. Decision aid and decision making. In: Bana e Costa, C.A. (Ed.), Readings in Multiple Criteria Decision Aid. Springer-Verlag, Berlin, pp. 17–35.

Roy, B., 1996. Multicriteria Methodology for Decision Analysis. Kluwer, Dordrecht.

Roy, B., Damart, S., 2002. L'analyse coûts-avantages, outil de concertation et de légitimation? Metropolis, N. 108/109, pp. 7–16.

Saaty, T.L., 1980. The Analytic Hierarchy Process. McGraw Hill, New York.

Simon, H.A., 1976. From Substantive to procedural rationality. In: Latsis, J.S. (Ed.), Methods and Appraisal in Economics. Cambridge University Press, Cambridge.

Simon, H.A., 1983. Reason in Human Affairs. Stanford University Press, Stanford.

Spash, C., Hanley, N., 1995. Preferences, information, and biodiversity preservation. Ecological Economics 12, 191–208.

Vansnick, J.C., 1986. On the problem of weights in multiple criteria decision making (the non-compensatory approach). European Journal of Operational Research 24, 288–294.

Vincke, Ph., 1994. Recent progresses in multicriteria decision-aid. Rivista di Matematica per le Scienze Economiche e Sociali 2, 21–32.