“Social Multi-Criteria Evaluation” (SMCE): Methodological Foundations and Operational Consequences

Giuseppe Munda
Universitat Autonoma de Barcelona
Dept. of Economics and Economic History, building B
08193 Bellaterra (Barcelona, Spain)
e_mail: giuseppe.munda@uab.es

Abstract: The main argument developed here is the proposal of the concept of “Social Multi-Criteria Evaluation” (SMCE) as a possible useful framework for the difficult policy problems of our Millennium, “where 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 such an approach should be developed?

The foundations of SMCE are set up by referring to concepts coming 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?
3. Which is the role of weights?
4. Which are the implications of SMCE for the mathematical axiomatization of multi-criteria aggregation conventions?

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

KEY WORDS: Multi-Criteria Analysis, Complexity Theory, Environment, Post-Normal Science, Incommensurability, Ethics

The main starting point of my discussion is that real world is characterised by deep complexity. 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 is reflecting only a sub-set of the possible representations of it. A system is then complex when the relevant aspects of a particular problem cannot be captured when using a single perspective (Funtowicz et al., 1999; O’Connor et al., 1996; Rosen, 1977).

To make things more difficult, human systems are reflexive complex systems. Reflexive systems have two peculiar properties: “awareness” and “purpose”. These properties imply an additional “jump” in complexity when trying to describe them. In fact, the presence of self-consciousness and purposes (reflexivity) entails that these systems can continuously add new relevant qualities/attributes that should be considered when explaining and describing their behaviour (i.e. they are learning systems). One important feature to keep in mind here is that reflexivity implies that the way humans represent a given policy problem to be solved, necessarily reflect the perceptions, values and interests of those who are structuring the problem.

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. Therefore the problem of multiple-identities in complex systems cannot only interpreted in terms of epistemological plurality due to the unavoidable existence of non-equivalent observers, but also in terms of ontological characteristics of the observed system (Giampietro, 1994; Giampietro and Mayumi, 2000). As discussed by Giampietro even a simple “objective” description of a geographical orientation results indeed to be impossible without taking an arbitrary subjective decision on the system scale considered relevant. 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 level chosen (the whole USA, a single state and so on)\(^1\).

The implications for multi-criteria evaluation of the scale issue are very important. For example, in generating evaluation criteria (e.g., in evaluating the impacts of the building of a ski infrastructure in a mountain region, who are the relevant social actors to interact with? The inhabitants of the mountain region, the potential users in urban areas or even the ecological preservationists all around the world might sound 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 means hydrogen is produced - being hydrogen an energy carrier and not an energy source -) or in choosing the weight factors (this issue will be discussed more in depth later on in this article).

A consequence of these deep subjectivities in the representation of policy problems imply that each time policy options are ranked, one has to chose 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”. That is, when ranking policy options, there is a

\(^1\) 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).
need for making a decision 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., the 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 the society as well as the rest of the society cannot perform any sound decision making without interacting with the scientists. That is, the question on “how to improve the quality of a policy process”, must be put, quite quickly, on the agenda of “scientists”, “decision makers” and indeed of 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" wants to indicate 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 given in Figure 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, 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; inferences will (naturally and legitimately) be conditioned by the values held by the agent. 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.

These ideas coming from complex system theory and Post Normal Science can be further corroborated by 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 argue 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. Technical incommensurability comes from the multidimensional nature 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 (Munda, 1997). The arguments developed in this section imply that at least there could be 2 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.

An attempt of operationalizing these two concepts of a compromise solution has been made in the so-called NAIADE approach (Munda, 1995). A peculiarity of NAIADE is the use of conflict analysis procedures to be integrated with the multi-criteria results. NAIADE can give the following information:

- ranking of the alternatives according to the set of evaluation criteria (i.e. technical compromise solution/s);
- indications of the distance of the positions of the various interest groups (i.e. possibilities of convergence of interests or coalition formations);
- ranking of the alternatives according to actors’ impacts or preferences (social compromise solution).

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 is the role of weights?
4. Which are the implications of SMCE for the mathematical axiomatization of multi-criteria aggregation conventions?

To answer these questions will be the aim of the rest of the present article.

2. Technical Incommensurability and Multi/Inter-Disciplinarity

An effective policy exercise, in order to be realistic, 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); 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). In this section the discussion will focus on the technical dimensions of a complex system; in the next two sections the issue of social incommensurability will be tackled.

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 the building of 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 for taking 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. To clarify this point I will refer to one real world case study I was involved, which seems to me an interesting success story. The problem was about the water supply system of the city of Palermo in western Sicily (South Italy). 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 characterised 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 and casual 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 is so deep that historically even the phenomenon of mafia started from the fighting for 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 as me 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 of us could give to progress towards a possible solution (or at least a better understanding) of a so complex problem.

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

![Figure 2. Scheme of the Palermo Water Supply System](image)

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.

At the moment to specify the alternatives, it was necessary to work on the structure of the Palermo water supply system shown in Figure 2 and… 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 (4 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 was evident the advantage of the multi-criteria structuring of the problem at hand, each expert suddenly knew which was her/his comparative advantage and which box could easily fill out!

From the experience made in 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-criteria fashion, it was astonishing to realize that immediately a common language was created. In terms of inter-disciplinarity, the issue is to find an 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.

In the Palermo case study, it was also experienced that the possibility of taking explicitly into account distribution issues 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 us to the issue of social incommensurability and public participation.

3. Social Incommensurability: Public Participation and Transparency

At this point of the discussion, one question arises, that is who is taking the decisions? Some critics of 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). 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 (Funtowicz and Ravetz, 1991, 1994; Funtowicz et al., 1999). 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, one should not forget that even a participatory policy process can always be conditioned by heavy value judgements such as, have all the social actors the same

---

3 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.
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 (see e.g., De Marchi and Ravetz, 2001; Guimaraes-Pereira, 2001, Kasemir et al., 2002). This need has been more and more recognized in a multicriteria decision-aid (MCDA) framework too. Banville et al., 1998 offers a very well structured and convincing argumentation in this direction. I agree with them on the need of extending 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 I argue that 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) and not “Participative Multi-criteria Evaluation” (PMCE) or “Stakeholder Multicriteria Decisions Aid (SMCDA) (Banville et al., 1998). To clarify this very important point, I will again use recent empirical examples in which I was involved I will start with 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).

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 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 Figure 3.

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 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 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. This point leads us to the initial and principal question, is business as usual a defensible option?

![Figure 3. Scheme of the Evaluation Process in the Troina Case Study](image)

According to the ranking, business as usual was almost on the bottom of the multicriteria ranking. While in the NAIADE conflict analysis, 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.

---

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

![Figure 4. Structure of the Evaluation Process in the DIAFANIS Project](image)

The problem dealt with was the possible expansion of a ski infrastructure in the Catalan 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, it was developed an institutional analysis 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 that was immediately evident of the focus group technique is 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, it was planned a widespread campaign of information 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-criteria 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. I firmly think that 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 Book on Governance) to the public at large for peer-reviewing.

2. I resolutely believe that 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 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’s imagine the extreme case where a development project in Amazon will affect an indigenous community with no 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.

All these arguments and convictions have lead me to the development of the concept of Social Multi-Criteria Evaluation whose very essence is the recognition that:

- 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.

---

\(^5\) Banville et al., 1998 discuss the notion of a stakeholder and its ambiguities deeply.
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. This is the reason why I use the term “multi-criteria evaluation” and not “multicriteria decision” when a social context is implied.

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? To answer this question will be the aim of the next section.


Are weights of any use in the framework of “social” multi-criteria evaluation? 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; 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’s start with an example; in Spain about 30 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 30 years ago.

These reflections can be synthesised graphically in Figure 5. 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. These dimensions are then translated by analysts into objectives and criteria. At this point a question arises who should attach criterion weights and how?

Let’s start with the pragmatic solution of no criterion weighting. This approach would probably reduce conflicts in the problem structuring step, but is it normatively correct? Indeed the fact that all the criteria have the same weight does not guarantee at all that objectives, dimensions at above all social groups have the same weight. This would be guaranteed only under the condition that all the dimensions have the same number of criteria. This of course is quite unnatural and artificial and even dangerous. Analysts could be tempted to choose the same number of criteria for each dimension although these criteria were completely redundant. A better solution can be to give the same weight to each dimension and to split each weight among the criteria of any dimension proportionally. We arrive then at the conclusion that by giving the same weight to all the criteria the different social dimensions have different weights (since any dimension will be weighted according to its number of criteria). On the contrary different criterion weights can guarantee that all the dimensions are considered equal!

---

6 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.
Thus from a normative (since at least there is a clear decision to weight all the dimensions equally) and maybe also pragmatic (since this decision may reduce social conflicts) points of view it seems better to attach weights to different criteria than not to attach them at all. But at this point one should answer to another question: why different social dimensions should be weighted equally?

To answer this question we have to accept a basic assumption: to weight different criteria implies to weight different groups in society (see figure 5). Can we find any normative justification for the weighting of social groups? I think that at least 4 different ethical justifications, capable of being transformed into social weights, exist:

1. Political democracy,
2. Economic democracy,
3. Sustainability,
4. Precautionary principle.

Let’s then discuss them briefly.

Political democracy. The basic idea here is that the majority of population has the right to influence decisions more than minorities. In this case, the derivation of social weights is relatively easy: each dimension should have a weight equal to the proportion of population supporting the values represented by this dimension. This approach presents two main problems (Moulin, 1988). One pragmatic in nature, i.e. to know the exact percentage of population can be a difficult and expensive process (for instance to go for a kind of referendum each time there is a complex decision that affects a community). The other more theoretical, i.e. minorities always loose without any compensation. For these reasons, economists have proposed the concept of economic democracy.
The main idea of economic democracy is that it is possible to avoid any type of referendum simply because it is possible to derive people’s preferences by looking at their behaviour on the market (Munda, 1996). Moreover, in this way it is possible to know also the intensity of preference of the economic agents (while this is not possible in the political democracy since the principle is one person, one vote) by looking at their consumer surplus and to compensate minorities by means of the Kaldor-Hicks compensation principle.

Problems connected with this approach are at least two:

1. Distributional issues. Very often willingness to pay depends on ability to pay, and not on preferences. Income distribution is then the basic principle one should accept to arrive at social decisions.

2. There are goods and services for which markets do not exist, e.g. environmental goods. For this kind of goods and services one should then invent artificial markets or derive implicit markets to obtain a so-called shadow price. The most general valuation method is contingent valuation.

Sustainability may imply that the economic, social and environmental dimensions should be taken into consideration equally. As a consequence this is the only principle under which all the dimensions should have the same weight. One should note that under this principle, the environmental dimension should weigh like the economic one, although e.g. only 5% of society has some environmental concerns. One way to explain this apparent paradox is that one should accept that this 5% of population is also speaking for animals and future generations who cannot be taken into account by both political and economic democracy principles. Another way to defend this position is that humans should behave according to a Kantian principle of universality such that any negative impact of their actions on the planet should be avoided. Thus even if no social group is willing to support the environmental dimension this should anyway weigh like the other dimensions or even more. This last position of a higher weight to the environmental dimension could be defended by using the so-called precautionary principle.

The precautionary principle maintains that in all situations where there is a simultaneous presence of uncertainty and irreversibility, prudence is the correct driving principle of social actions. For example, in the case of global warming, due to the fact that there is no clear scientific evidence of what may happen with the present trend of greenhouse gas emissions, it is better to be prudent and reduce these emissions now, even if the economic cost could be very high. A consequence of the adoption of this principle is that the environmental dimension has a higher weight or that it must be operationalized by using constraints instead of criteria.

This discussion has the following main consequences:

---

7 Consumer’s surplus is the difference between what an individual has to pay for a good (the market price) and what an individual would be willing to pay for each unit of the good rather than to go without it. Hence consumer’s surplus is defined as the area under the demand function and above the price line. A necessary condition for an effective calculation of the aggregate consumer surplus is knowledge of the demand curves for the elements of the project at hand.

8 The social cost of a given output is defined as the sum of money which is just adequate when paid as compensation to restore to their previous level of utility all who lose as a result of production of the output in question. In other words, the Kaldor-Hicks principle declares a social state y “socially preferable” to an existing social state x if those who gain from the move to y can compensate those who lose and still have some gains left over. Such a situation is consistent with a Pareto improvement since we have x indifferent to y for the losers (once they are compensated) and y preferred to x for the gainers (if they can over-compensate). It is just this principle which underlies cost-benefit analysis. If the monetary value of benefits exceeds the monetary value of costs, then the gainers can hypothetically compensate the losers and still have some gains left over. The excess of gains over required compensation is equal to the net benefits of the project.

9 At least if one accepts that sustainability has three main columns to be integrated simultaneously: economic efficiency, inter/intra generational equity and impacts of the society on environmental quality (Munda, 1997).
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. Of course, I do not think that the list of ethical principles I provided is the only possible one, it is just a first attempt to provide a plausible one. Actually, I do think there are other possibilities but however, I have the firm conviction that this list of possibilities is a finite and not that big one.

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

4. Sensitivity and robustness analysis have a complete different meaning with respect to the case of single person and technical decisions. 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.

However, one should note that the structure presented in Figure 5 is useful for a theoretical discussion, but it is not an operational one. In real-world applications, multi-criteria rankings derive from a set of evaluation criteria generated by the research team and searching for the “technical compromise solution”. As I argued before, I think this is a correct approach. The technical translation of social preferences, interests, values and desires into evaluation criteria has necessarily to be done by an inter/multi-disciplinary team. Thus the direct connection of social actors with dimensions and criteria is somewhat lost. This is the reason why I think that the use of 2 different matrices in parallel, one technical and one social can give useful information. In fact even when criteria are coming directly from the affected actors (very extreme case in my opinion), it is important to remember that the quantification of the criterion scores is anyway a technical issue; by contrast, the impact score of each alternative policy option to each social actor is much more direct. Of course the basic issue of weighting of social groups remains the same. This aspect was not then correctly addressed in the first version of NAIADE where no weight to social group was considered.

In conclusion regarding the issue of weight, I argue the following:

1. Contrary to what I maintained some years ago (Munda, 1993), I think that the use of weights in policy problems is desirable and important.

---

10 I have deeply 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.
2. Weights should come from ethical positions and not from participatory techniques or even worst from decision-maker’s preferences.
3. Since incommensurability can be divided into a technical one and a social one the parallel use of a multi-criteria impact matrix and of a conflict matrix showing clearly the distributional issues at stake is desirable.
4. Since a multi-criteria framework is very useful for implementing inter/multi-disciplinarity, a pragmatic solution can be to weight the dimensions used equally. This may reduce the probability of conflicts inside the scientific team.
5. In the conflict analysis matrix, the weights attached to the various social groups should be derived from ethical principles and connected sensitivity or robustness analyses should always be carried out.

Summarizing the arguments developed till now, it is possible to conclude that social multi-criteria evaluation can be considered as an approach which is:
1. Inter/Multi-disciplinary, to respect the plurality of scientific points of view.
2. Participative, to get as much knowledge as possible from the general public.
3. Transparent, to make clear the ethical assumptions adopted in the study and the responsibilities taken.
4. Consistent, to assure that the results are a real consequence of the assumptions adopted.

The properties from 1 to 3 are clearly dependent on the evaluation process; property 4 is more connected to the mathematical models used. Next section is then devoted to the discussion of the consistency property and connected mathematical aggregation conventions.

5. How to Choose an Adequate Aggregation Convention for SMCE?

In practical applications of multi-criteria evaluation for policy analysis, the most common approach is the use of discrete multi-criteria methods. For this reason I will focus on this topic here (Beinat and Nijkamp, 1998, Janssen and Munda, 1999, Nijkamp et al., 1990; Romero and Rehman, 1989; Roy and Bouyssou, 1993). In this context, there is a range of multicriteria problem formulations, which may take one of the following forms (Roy, 1985; 1996):

(α) the aim is to identify one and only one final alternative;
(β) the aim is the assignment of each alternative to an appropriate predefined category according to what one wants it to become afterwards (for instance, acceptance, rejection or delay for additional information);
(γ) the aim is to rank all feasible alternatives according to a total or partial preorder;
(δ) the aim is to describe relevant alternatives and their consequences.

Clearly the steps required by such a process need a number of arbitrary unavoidable subjective decisions. The degree of the subjective component may be higher or lower but it is always present. 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-criteria aggregation convention i.e. a multi-criteria method is quite difficult (Arrow and Raynaud, 1986). To deal with the problem correctly pointed out by Arrow and Raynaud, 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.

Let’s start with a principle that carries a lot of epistemological implications: the principle that dominated alternatives can be ignored in an evaluation exercise. This principle, often presented as a simple technical step, needs the acceptance of the following assumptions:

1) The assumption that all the relevant criteria have been identified needs to be accepted. If relevant criteria are omitted, there are potential opportunity costs associated with assuming that it is safe to ignore dominated alternatives. In all cases in which uncertainty and genuine ignorance can be assumed to play an important role in the problem structuring, the concept of “efficient solution” may be useful, but only in the short term, whereas it may be potentially very dangerous in the long term. Any algorithmic definition of an option dominating all the others on all the selected criteria must be based on a closed and finite information space. This implies the unavoidable existence of: (1) other goals/objectives (reflecting the identity of other social groups) not considered in the existing analysis; and (2) other “problem structuring” able to catch other views and dynamics of the system’s behaviour under analysis.

2) The assumption that only one alternative considered the best has to be identified needs to be accepted (α problem formulation). Since the "second best" may have been eliminated during the technical screening, if more than one action has to be found, the elimination of the "inefficient" action may result in an opportunity loss (one has to note that if the best action is removed from the set of feasible alternatives, then the second best becomes a member of the non dominated set). If one is interested in the γ problem formulation, then dominated alternatives cannot be eliminated. It has to be noted that in public policies, it is often much more useful to have a ranking of policy options than to select just one alternative. In fact sometimes the first alternative in the ranking can also be the most controversial from a social point of view. Thus in the framework of a social dialectics, it is more useful to implement the policy option that maybe ranks second (so technically it is not so bad) but socially might reduce conflicts.

3) A third problem is connected to the question: how relevant are "irrelevant" alternatives? Arrow's axiom of "the independence of irrelevant alternatives" (Arrow, 1951) states that the choice made in a given set of alternatives A depends only on the ordering made with respect to the alternatives in that set. Alternatives outside A (irrelevant since the choice must be made within A) should not effect the choice inside A. Empirical experience does not generally support this axiom; thus to exclude some actions already inside A can have even less justification. However, the issue of the independense of irrelevant alternatives is particularly important and tricky when pair-wise comparisons are used, since rank-reversal phenomena may occur.

4) Finally, a dominated action may be slightly worst than an efficient action, if indifference and/or preference thresholds are used (which from a descriptive point of view, in the light of Luce’s paradox (Luce, 1956) can be considered highly recommended), then the two actions could present an indifference relation.

As a conclusion of this discussion we can state then that in SMCE applications it is better to use aggregation procedures that do not exclude dominated alternatives a priori.

Another sensitive point is the one of compensability. Some critics of multi-criteria evaluation often say that to compute some kind of "utility" requires making trade-offs and thus there is no real difference between multi-criteria methods and conventional cost-benefit analysis. I share this opinion, but I would like to stress that this apply only to utility based compensatory multi-criteria methods (see Munda, 1996 for a deep discussion of this issue).
The aggregation of several dimensions implies taking a position on the problem of compensability. Compensability refers to the existence of trade-offs, i.e. the possibility of offsetting a disadvantage on some attribute by a sufficiently large advantage on another attribute, whereas smaller advantages would not do the same. Thus a preference relation is non-compensatory if no trade-off occurs and is compensatory otherwise. (Bouyssou, 1986; Bouyssoy and Vansnick, 1986; Keeney and Raiffa, 1976; Vansnick, 1986; Vincke, 1992).

From the discussion on weights of the previous section we already know that weights as trade-offs are not desirable for SMCE. As a consequence methods allowing for the use of weights as importance coefficients should be preferred. From this point of view, ELECTRE II (Roy and Bertier, 1973) and ELECTRE III (Roy, 1978) methods are probably the most interesting methods. A method that for sure needs weights as importance coefficients is also the REGIME method when only ordinal criterion scores are used (Hinloopen et al., 1983). However, when mixed information is considered (Hinlopen and Nijkamp, 1990), weights are more likely to be considered as trade-offs and not importance coefficients anymore.

Moreover, compensability has also consequences for the so-called weak/strong sustainability debate (Faucheux and O'Connor M., 1998). The so called "weak sustainability" concept states that an economy can be considered sustainable if it saves more than the combined depreciation of natural and man-made capital. "We can pass on less environment so long as we offset this loss by increasing the stock of roads and machinery, or other man-made (physical) capital. Alternatively, we can have fewer roads and factories so long as we compensate by having more wetlands or mixed woodlands or more education (Turner et al., 1994, p. 56)". The concept of "strong sustainability" is based on the assumption that certain sorts of natural capital are deemed critical and not readily substitutable by man-made capital. In particular, the characterisation of sustainability in terms of the "strong" criterion of non-negative change over time in stocks of specified natural capital provides a strong justification for development of non-monetary indicators of ecological sustainability based on direct physical measurement of important stocks and flows. It is clear at this stage that non-compensatory and partial compensatory multi-criteria aggregation conventions are the only ones that may allow an implementation of the strong sustainability concept (see Martinez-Alier et al., 1998 for a deeper discussion of this point).

Let's now try to find other desirable properties for SMCE. It has been argued that the presence of qualitative information in evaluation problems concerning socio-economic and physical planning is a rule, rather than an exception (Nijkamp et al., 1990). Thus, the idea of technical incommensurability implies that there is a clear need for methods that are able to take into account information of a "mixed" type (both qualitative and quantitative criterion scores). For simplicity, I refer to qualitative information as information measured on a nominal or ordinal scale, and to quantitative information as information measured on an interval or ratio scale. Moreover, ideally, this information should be precise, certain, exhaustive and unequivocal. But in reality, it is often necessary to use information which does not have those characteristics so that one has to face the uncertainty of a stochastic and/or fuzzy nature present in the data (Munda et al., 1995; Munda, 1995). In conclusion, multi-criteria methods able to tackle consistently the widest types of mixed information should be considered as desirable ones. Examples of this kind of methods are REGIME (Hinlopen and

12 However, this property, very important in our context, is disputed by some people who argue that indeed weights depend also on intensity since the complete formalisation of importance depend on both the weight and the veto threshold corresponding to a given criterion. This has two main implications: 1) when changing a weight on a criterion, also its veto threshold must change, (2) the fact that weight and veto threshold are connected means that the criterion weight depends also on intensity; how can we be sure it is not a trade-off? At least at my knowledge, at this moment no precise mathematical answer exists to this problem, and then I just give the warning and leave the reader free to create a personal opinion.
Another important point is the issue of how deep the interaction with a method should be. To give an example, the implementation of the philosophy of “Multi-Criteria Decision Aid” (MCDA) by means of e.g. an ELECTRE method leads to a need of establishing a large number of ad hoc parameters, i.e. indifference and preference thresholds, concordance threshold, discordance thresholds, weights. It seems to me that according to Roy these parameters allow the decision-maker to express her/his preferences and to learn about them. In this way, the concept of decision aid is implemented; in this framework, probably the PROMETHEE methods (Brans et al., 1984) are the best attempt to introduce various parameters with physical or economic interpretation for a decision-maker. I also used these suggestions in building the NAIADE method. However, it must be admitted that a big amount of parameters may cause a loss of transparency and consistency in the model. I think that this approach is correct in the case of single person or very technical decisions, but it is not desirable in a social framework. In this latter case, the role of the multi-criteria method should be to guarantee the consistency between the information presented in the impact matrix and the rankings obtained without any further manipulation (at least if it is agreed that transparency is an important feature of this type of approach). For this reason I believe that simplicity is a very desirable property for multi-criteria methods to be used in a SMCE framework. One should however note that this wish may conflict with the desire of taken into account indifference and preference thresholds. May be in this case, sensitivity or robustness analyses in the classical sense are the only possible solution.

Finally, one should note that the possibility of explicitly tackling the hierarchy of scale issue would be a quite interesting and desirable property (as discussed in section 1). At my knowledge, the only method which explicitly deals with the hierarchy issue is the Analytic Hierarchy Process (AHP) method (Saaty, 1980). However, in the framework of SMCE, this method presents the problem of needing weights as trade-offs.

As a conclusion of this section, I think it is possible to state quite safely that a problem of all multi-criteria approaches is that a lot of different mathematical aggregation conventions (or methods) exist. In synthesis, the information contained in the impact matrix useful for tackling one of the possible decision problem formulations is the following:

- Intensity of preference (when quantitative criterion scores are present).
- Number of criteria in favour of a given alternative (when a pair-wise comparison is carried out).
- Weight attached to each single criterion.
- Relationship of each single alternative with all the other alternatives.

Combinations of this information generate different methods. For example the use of weight with intensity of preference originates compensatory multi-criteria methods and gives to the weights the meaning of trade-offs. For the case of SMCE, the following considerations can be useful to select appropriate methods.

The idea of social incommensurability makes the following properties desirable in a social multi-criteria method:

- Multicriteria methods must be as simple as possible to guarantee transparency.
- Weights in this framework are clearly meaningful only as importance coefficients and not as trade-off. As a consequence, complete compensability cannot be implemented.

---

13 One should note that an issue here is the meaning of weights. Since in the concordance index of PROMETHEE they are multiplied by a kind of credibility degree which looks very much as information on the preference intensity; in principle their meaning should be of trade-offs and not importance coefficients.
• 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 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.

The idea of technical incommensurability makes the following properties desirable in a social multi-criteria method:

• Partial compensability is an essential consistency requirement.

• 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.

• The hierarchical dimension of a policy problem should be explicitly considered.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MAUT</td>
<td>--</td>
<td>--</td>
<td>+</td>
</tr>
<tr>
<td>ELECTRE 2</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>ELECTRE 3</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
</tr>
<tr>
<td>REGIME (H,N,R,1983)</td>
<td>+++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>REGIME (H,N,1990)</td>
<td>+++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>NAIADE</td>
<td>+++</td>
<td>++</td>
<td>---</td>
</tr>
<tr>
<td>AHP</td>
<td>+++</td>
<td>-</td>
<td>---</td>
</tr>
<tr>
<td>EVAMIX</td>
<td>+++</td>
<td>-</td>
<td>---</td>
</tr>
<tr>
<td>PROMETHEE</td>
<td>+++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Martel &amp; Zaras method</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1. Example of Evaluation of Some Multicriteria Methods According to Proposed Desirable Properties for SMCE
As an example, I present here a table (see Table 1) where the above cited methods are confronted with the proposed desirable properties. The scale +++/-/- is ordinal in nature (+ is more desirable than -; +++ is more desirable than + and --- is less desirable than -) and subjectively determined by my knowledge of these methods. However, even if quite rough, the information contained in Table 1 leads to an important conclusion no one of these methods achieves all the properties considered desirable.

6. Conclusion: Social Multi-criteria Evaluation as an Integration of Mathematical Tools with the Social Process

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 it-self. 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 want just to remind 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!

The importance of the decision process has recently been emphasised by different authors. 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: all aspects of the whole decision process which leads to a given decision also contribute to its quality and success. According to Roy, Multi-Criteria Decision Making (MCDM) approaches assume that the decision-maker’s preferences are made perfectly explicit, so that the only thing left to do is consider a well-formulated mathematical model. "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: organisational, pedagogical and cultural aspects of the whole decision process which leads to a given decision also contribute to its quality and success..."(Roy, 1990). Thus, it becomes impossible to found 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 (MCDM) (Roy, 1985, 1990, 1996), the principal aim is not to discover a solution, but to construct or create something which is 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 Figure 6). 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 still under the conditions of medium uncertainty and medium decision stake ranges.

![Figure 6. Multi-Criteria Approaches in Relationship to Funtowicz-Ravetz Classification of Science for Policy](image)

All the arguments and convictions discussed in this article have lead me to the development of the concept of Social Multi-Criteria Evaluation whose very essence is the recognition that (see Figure 7):

- 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\(^\text{14}\).

\(^{14}\) 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.
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. This is the reason why I use the term “multicriteria evaluation” and not “multicriteria decision” when a social context is implied.

However, one should remember that we are in a second best world. A useful analogy here is with Flatland, the classic Victorian science fiction and social parody (Abbott, 1935). There, the inhabitants of spaces with more dimensions had a richer awareness of themselves, and also could see beyond and through the consciousness of the simpler creatures inhabiting fewer dimensions. At this stage it is not unfair to reveal the dénouement of the story, namely that the Sphere of three-dimensional space showed himself to be just another Flatlander at heart, when he angrily refused to accept the reality of higher dimensions of being…

![Figure 7. Synthesis of a Social Multi-Criteria Evaluation Process](image)

**ACKNOWLEDGEMENTS**

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)*, Contract NNE5-2001-273.

**References**

Arrow K.J. (1951) - *Social choice and individual values*, Wiley, New York.


