Social Multi-Criteria Evaluation (SMCE)

Giuseppe Munda, EC-JRC/IPSC[1]

1. The Issue of representation of Real-World Complex Systems

Real world is characterised by deep *complexity*. This obvious observation has important implications on the manner policy problems are represented and decision-making is framed. 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). To make things more difficult, human systems are *reflexive* complex systems. Reflexive systems have two peculiar properties: "*awareness*" and "*purpose*", which imply an additional "jump" in describing complexity. In fact, 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 and describing their behaviour (i.e. human systems are learning systems). One important feature of reflexivity is that the human representation of a given *policy problem* necessarily reflects perceptions, values and interests of those 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 (Giampietro, 1994; Giampietro and Mayumi, 2000). As discussed by Giampietro even a simple "objective" description of a geographical orientation is 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 chosen as a reference point (the whole USA, a single state and so on). 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 building 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 and 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?)

Therefore, the problem of *multiple-identities* in complex systems cannot only be interpreted in terms of epistemological plurality (non-equivalent observers), but also in terms of *ontological characteristics* of the observed system (non-equivalent observations). A consequence of these deep subjectivities is that in any normative exercise connected to a public decision problem, one has to

[1] The views expressed are those of the author and do not represent necessarily those of the European Commission.

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". 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. One should note that the representation 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*.

An outcome of this discussion is that the political and social framework must find a place in multi-criteria decision aid. To give 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.

In general, these concerns have not been considered very relevant by scientific research in the past (where the basic implicit assumption was that time was an infinite resource). On the other hand, the new nature of the policy problems faced in this third millennium (e.g., the mad cow, genetic modified organisms, …), implies that very often when using science for policy-making, long term consequences may exist and scientists and policy-makers 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 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 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"* (the name "post-normal" indicates a difference from the puzzle-solving exercises of normal science, in the Kuhnian sense), where it is possible to better deal with two crucial aspects of science in the policy domain: *uncertainty* and *value conflict*. When cases in which conclusions are not completely determined by scientific facts exist; 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.

2. Social Multi-Criteria Evaluation as a Tool for Aiding Policy Processes in Reflexive Complex Systems

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 argue that it is possible to further distinguish the concepts of social incommensurability and technical incommensurability (Munda, 2002a). *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 of complexity and refers to the issue of representation of multiple identities in descriptive models.

If one wants to implement technical incommensurability, 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 multi-dimensional in nature, seems an interesting framework in which operationalize Neurath's ideal. This virtue of multi-criteria approaches has been corroborated in a great number of real-world case studies employing a variety of methods (see e.g. Beinat, 1997; Espelta *et al.,* 2003; Janssen, 1992; Maystre *et al.,* 1994; Moreno-Jiménez *et al,* 1999; Stewart and Joubert, 1998). A real world case study involving the water supply system of the city of Palermo in western Sicily (Southern Italy) can help to clarify the point.

The case was part of a project commissioned by the region of Sicily and executed in the frame of the European Commission's DGXVI structural funds. This case study was developed during two years of interaction mainly between a multi-disciplinary 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 strong competition among different categories of consumptive water uses and, as a consequence, among various interest groups. Such competition also exists between consumptive uses as a whole and "ecological uses" which aim to limit

water diversion for off-stream uses in order to preserve the ecological equilibrium of ecosystems. This permanent state of competition may intensify into real conflict under drought conditions, i.e. when there is a temporary reduction of available water resources due to a long and severe decrease in 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 historical roots. Indeed *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 on water supply system characteristics and demand levels, which are both affected by different drought mitigation measures. As a consequence, the purely technical hydrological solutions cannot be separated from their consequences for 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 with this kind of problem. However, it was still very difficult to find a common language and to understand which contribution each could make towards a possible solution (or at least a better understanding) of such a complex problem.

The water system of Palermo provides water to municipal, agricultural 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 could be divided into two main groups:

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

To establish the alternatives, it was necessary to understand the structure of the Palermo water supply system and, given the technicalities involved (as one can see from its description in Figure 3.1), 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 a set of criteria which included the economic dimension (e.g. associated financial costs

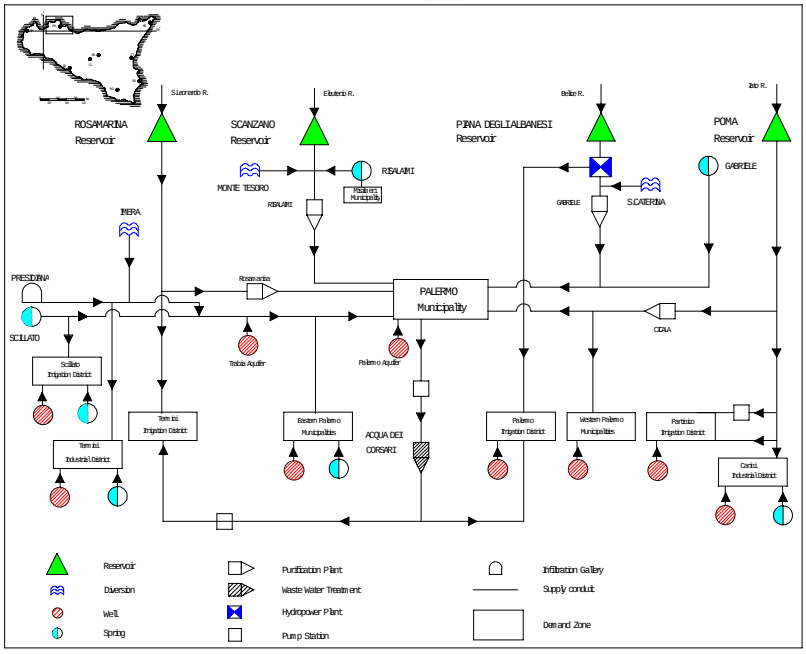

Figure 1 Scheme of the Palermo Water Supply System
(Source: Elaborations made by G. Rossi and his team at Catania University)

and benefits for 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 became evident. Each expert suddenly knew her /his comparative advantage.

From the experience of this case study, a first lesson could be learned: *a multi-criteria framework is a very efficient means of implementing a multi-inter-disciplinary approach.* The experts involved had diverse backgrounds (mainly in engineering, economics and mathematics). While the communication process was initially very difficult, once it had been decided to structure the problem in a multi-criterion fashion, it was astonishing to recognize that a common language had immediately been 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, it is to propose and compute an appropriate criterion score. The efficiency of the interaction process may also increase greatly[2].

In the Palermo case study, it was also experienced that explicitly taking distribution issues into account increases the transparency of the study and

[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).

facilitates an effective process of interaction between various social actors. This second lesson leads on to the issue of social incommensurability and public participation.

For the formation of contemporary public policies, it is hard to imagine any viable alternative to *extended peer communities*. 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 (De Marchi and Ravetz, 2001; Gowdy and O'Hara, 1996).

This need of public participation has been more and more recognized in a multi-criteria 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) in substitution of "Participative Multi-criteria Evaluation" (PMCE) or "Stakeholder Multi-criteria Decision Aid" (SMCDA).

One should not forget that even a participatory policy process can always be conditioned by heavy value judgements. 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?

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. 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?

To answer this question we have to accept a basic assumption: to weigh different criteria implies to ponder 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.

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. 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.
4. *Sensitivity and robustness analysis* have a complete different meaning with respect to the case of single person and technical decisions (Roy, 2002; Rosenhead, 2002). In fact in the case of SMCE, weights derive 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.

The main principles of Social Multi-criteria Evaluation can be summarised as follows (Munda, 2004):

(1) 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 the assumptions used.

(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 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 Book on Governance) 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* of the decisions taken (Roy and Damart, 2002), 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?

(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 my opinion the substantial meaning of multi-criteria evaluation in a social context is simply tolerance and democracy. Complexity is 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, as pointed out by authors such as B. Roy (1985) and H. Simon (1976), in a multi-criteria context what really matters is *the process* since the problem structuring will determine the result. 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 "multi-criteria decision" when a social context is implied. Of course this does not mean that mathematical models are useless. On the contrary, I strongly believe that they play the fundamental

role of guaranteeing consistency between the assumptions used and the results obtained, in terms of rankings of the available policy options. For this reason I think that multi-criteria algorithms to be used in a social context should be as simple as possible (i.e. with the minimum number of exogenous parameters) and that their axiomatization should be complete and clear.

References

Banville C., Landry M., Martel J.M., Boulaire C. (1998) – A stakeholder approach to MCDA, *Systems Research and Behavioral Science*, Vol. 15, pp. 15-32.

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

Funtowicz S.O., Ravetz J. R. (1991) - A new scientific methodology for global environmental issues, in R. Costanza (ed.)- *Ecological Economics,* New York, Columbia, 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, pp.197-207.

Funtowicz S., Martinez-Alier J., Munda G. and 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. (2000) – Multiple-scale integrated assessment of societal metabolism: introducing the approach, *Population and Environment,* Vol. 22, No. 2, pp. 109-154.

Giampietro M., Mayumi K. (2000) – Multiple-scale integrated assessment of societal metabolism: integrating biophysical and economic representations across scales, *Population and Environment,* Vol. 22, No. 2, pp. 155-210.

Gowdy, J.M., and O'Hara S. (1996) - *Economic Theory for Environmentalists.* Saint Lucie Press.

Kuhn T.S. (1962) - *The structure of scientific revolutions,* University of Chicago Press, Chicago.

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

Munda G. (1995) - *Multicriteria evaluation in a fuzzy environment,* Physica-Verlag, Contributions to Economics Series, Heidelberg.

Munda G. (2004) – "Social multi-criteria evaluation (SMCE)": methodological foundations and operational consequences, *European Journal of Operational Research*, vol. 158/3, pp. 662-677.

Munda G. (2005) – Multi-Criteria Decision Analysis and Sustainable Development, in J. Figueira, S. Greco and M. Ehrgott (eds.) –*Multiple-criteria decision analysis. State of the art surveys*, Springer International

Series in Operations Research and Management Science, New York, pp. 953-986.

Munda G. (2005) – "Measuring sustainability": a multi-criterion framework, *Environment, Development and Sustainability_,* Vol 7, No. 1, pp. 117-134.

Munda G. (2006) – Social multi-criteria evaluation for urban sustainability policies*, Land Use Policy* Volume 23, Issue 1, January, Pages 86-94.

Munda G. (2007) – Social multi-criteria evaluation for a sustainable economy, Springer-Verlag, Heidelberg, New York.

Neurath O. (1973) – *Empiricism and sociology*, Reidel, Dordrecht.

O'Connor M., Faucheux S., Froger G., Funtowicz S.O., Munda G. (1996) - Emergent complexity and procedural rationality: post-normal science for sustainability, in R. Costanza, O. Segura and Martinez-Alier J. (eds.) -*Getting down to earth: practical applications of ecological economics,* Island Press/ISEE, Washington D.C., pp. 223-248.

O'Neill, J. (1993) - *Ecology, Policy and Politics*, Routledge, London.

Rosenhead J. (2002) – Robustness Analysis, *Newsletter of the European Working Group «Multicriteria Aid for Decisions», Forum, Series 3, No. 6.*

Roy B. (1985) - *Méthodologie multicritère d'aide à la décision,* Economica, Paris.

Roy B. (2002) – Robustesse de quoi et vis-à-vis de quoi mais aussi robustesse pourquoi en aide à la décision ? *Newsletter of the European Working Group «Multicriteria Aid for Decisions», Opinion Makers Section, Series 3, No. 6.*

Roy B., Damart S. (2002) – L'analyse Coûts-Avantages, outil de concertation et de légitimation? *Metropolis,* N. 108/109, pp. 7-16.

Simon H.A. (1976) - From Substantive to procedural rationality. In: J.S. Latsis (Editor) *Methods and appraisal in economics,* Cambridge: Cambridge University Press