Technical Report on Environmental Sciences

Integrated Assessment of Agricultural Sustainability: the Pros and Cons of Reductionism

Mario GIAMPIETRO

Affiliations:

ICREA Research Professor, Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Spain.

Contact: Mario GIAMPIETRO <Mario.Giampietro@uab.cat>

Date: *09.11.2010*

Author: Mario Giampietro
year: 2010
Title: Integrated Assessment of Agricultural Sustainability: the Pros and Cons of Reductionism

Reports on Environmental Sciences ##
URL:

Institut de Ciència i Tecnologia Ambientals (ICTA)
Edifici Cn, Campus UAB
08193 Cerdanyola del Vallès, Spain
Tel: (+34) 935812974
http://icta.uab.cat
icta@uab.cat

Integrated Assessment of Agricultural Sustainability: The Pros and Cons of Reductionism

Mario Giampietro
ICREA Research Professor, Institut de Ciència i Tecnologia Ambientals (ICTA), Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Spain.

ABSTRACT

In this paper I review a series of theoretical concepts that are relevant for the integrated assessment of agricultural sustainability but that are not generally included in the curriculum of the various scientific disciplines dealing with quantitative analysis of agriculture. I first illustrate with plain narratives and concrete examples that sustainability is an extremely complex issue requiring the simultaneous consideration of several aspects, which cannot be reduced into a single indicator of performance. Following, I justify this obvious need for multi-criteria analysis with theoretical concepts dealing with the epistemological predicament of complexity, starting from classic philosophical lessons to arrive to recent developments in complex system theory, in particular Rosen´s theory of modelling relation which is essential to analyze the quality of any quantitative representation. The implications of these theoretical concepts are then illustrated with applications of multi-criteria analysis to the sustainability of agriculture. I wrap up by pointing out the crucial difference between "integrated assessment" and "integrated analysis". An integrated analysis is a set of indicators and analytical models generating an analytical output. An integrated assessment is much more than that. It is about finding an effective way to deal with three key issues: (i) legitimacy – how to handle the unavoidable existence of legitimate but contrasting points of view about different meanings given by social actors to the word "development"; (ii) pertinence – how to handle in a coherent way scientific analyses referring to different scales and dimensions; and (iii) credibility – how to handle the unavoidable existence of uncertainty and genuine ignorance, when dealing with the analysis of future scenarios.

KEY WORDS: Sustainable Agriculture; Multi-Criteria Analysis; Integrated Assessment; Modelling Relation; Hierarchy Theory; Post-Normal Science; Science for Governance

TABLE OF CONTENTS

1. Introduction

The science of agricultural sustainability has been forced to acknowledge the existence of persistent and growing problems and challenges:

1) The chronic problem of inadequate food supply; more than one billion people are malnourished according to the latest FAO estimates (FAO, 2009);

2) Growing environmental problems; the existing trend of damage to natural ecosystems and of loss of biodiversity at the world level is scaring, according to the Millennium Ecosystem Assessment project (MEA, 2005);

3) Serious socioeconomic problems; the increasing collapse of social fabric in rural areas all over this planet has been described by the International Assessment of Agricultural Science and Technology for Development (IAASTD, 2008). In developed countries, the chronic crisis of the agricultural sector is reflected in the endless discussions on the amounts and modalities of farm subsidies.

This list of growing problems in agriculture has justified a general call for swift ad-hoc actions rather than theoretical discussions; call that has been endorsed by politicians and scientists alike. However, I believe that a reflexion on the scientific framework employed in the science of agricultural sustainability is direly needed. As a matter of fact, there is a growing literature questioning the direction taken by technological “progress” in agriculture (Röling, 1996; Altieri, 1996, 1999, 2004; Gliessman, 1998; Waltner-Toews and Lang, 2000; Jordan, 2001; Pretty *et al*., 2003; Giampietro, 2003, 2008; Rosset, 2006; IAASTD, 2008). In my view, the persistent failure of the existing paradigm of industrial agriculture to improve the sustainability of farming systems clearly suggests that the pre-analytical narratives that have been adopted thus far to justify and implement this paradigm are no longer valid. The paradigm of industrial agriculture is defined here as the uncontested agreement over the set of narratives used to define an “improvement” in agriculture when envisioning current technological progress. If this pre-analytical vision is wrong, it is unwise to continue proposing a strategy of technical innovations with the goal of doing “more of the same” in spite of the fact that nobody is happy with the actual performance of agriculture, neither in developed nor in developing countries. The paradigm of industrial agriculture is presently producing, by stressing the environment and eliminating traditional rural communities, food surplus without a demand in developed countries and a too-expensive food supply in developing countries (Giampietro, 2008; Giampietro and Mayumi, 2009).

In fact technological progress within the paradigm of industrial agriculture follows the double goal of getting rid of farmers and of local agro-ecosystems, by substituting human labour with machines and natural flows of nutrients and ecological services with technical inputs. For this reason agricultural production is increasingly perceived as an activity that is incompatible with the preservation of nature. Not surprising then that recent policies for the preservation of biodiversity increasingly favour the creation of 'human-free' reserves. These policies seem to assume that humans and nature cannot interact in a sustainable way. The combined effect of technical progress in agriculture and human-free conservation policies is the progressive elimination of those human communities that still know how to live in harmony with nature by respecting the natural patterns of ecological processes. This pattern is extremely evident in Latin America where the expansion of *La Republica Unida de la Soya* (a new geographic entity proposed by Syngenta in an advertisement published in 2003 in the Argentinean newspaper La Nacion – Joensen, 2004) is eliminating the various forms of traditional farming systems across Argentina, Paraguay, Bolivia, Uruguay and Brazil. Indigenous communities that learned how to co-evolve with natural ecosystems for millennia are first marginalized, then disempowered and finally forced out from their traditional territories.

These trends point at a major flaw in the use of scientific analyses for governance in relation to agricultural sustainability. This flaw is especially evident when dealing with quantitative analysis and hence I believe it justifies the argument presented in this paper. In technical jargon we may say that this paper deals with an overview of *epistemological challenges* associated with the integrated assessment of sustainable farming and the analytical tools developed for dealing with them. From personal experience, I know that the term epistemology is unwelcome by quantitative scientists, it being considered some sort of idle philosophical distraction. As a matter of fact, I must confess that I myself blindly crunched numbers for decades, before I eventually came to realize that a good understanding of the semantics used for generating quantitative analysis is as crucial as the rigor in the gathering and processing of data. When dealing with sustainability, robust data processed in flawless calculations within *meaningless* models or *misleading* narratives are the major source of harm to the cause.

To introduce the issue, I first present examples of the epistemological challenge encountered when defining the performance of a production system on multiple scales and in relation to different dimensions of analysis. The main point is that dealing with sustainability requires the simultaneous use of different narratives, which are not reducible to each other with mathematical models. Studying agricultural sustainability entails the adoption of multi-criteria analysis. Following, I provide an overview of theoretical concepts useful for understanding the roots of this epistemological challenge, including the theory of modelling relation proposed by Robert Rosen. I then illustrate the challenges faced by those willing to apply multi-criteria analysis to the sustainability of agriculture. Finally, in the last section I wrap up the discussion by pointing out the difference between integrated analysis and integrated assessment, and present the conclusions.

2. The challenge of quantitative analysis of farming system performance

2.1 Strength and Weakness of Quantitative Analysis

In modern life, numbers have become key inputs required for virtually any daily routine. Be it shopping, driving, hiring personal, using a mobile phone, a microwave oven, or a TV recorder, we have to organize the input of information over numbers. Human beings seem to have become fully dependent on numbers for handling any flow of information. No wonder then that quantitative analysis is perceived by many as the *only* way to generate *true* and *useful* information.

However, this complete dependency on numbers represents an Achilles heel when dealing with complex issues. In fact, the modern scientific approach leading to quantitative description is aptly described by the term *reductionism*, defined as a process that allows us to simplify the complexity associated with any real situation, by focusing only on few relevant aspects of a complex issue at the time. Indeed, it is only after having carried out such a simplification that the magic power of numbers can be unleashed. Therefore, reductionism implies a clear predicament: If we want to describe a complex phenomenon that can only be perceived and represented using simultaneously several dimensions and scales of analysis, we must focus only on a limited subset of these possible scales and dimensions in order to get a simple quantitative representation, i.e. crisp numbers. This is exactly the meaning of the term used to describe this operation: reductionism.

Any quantitative model in order to be operational in finite time must be based on the use of a limited set of attributes for the representation of the observed system, which translate into a finite set of variables that can be included in the inferential system (Giampietro *et al.*, 2005). This pre-analytical choice of a limited number of variables entails a dramatic reduction in the set of the potential perceptions and representations of the given complex situation. This is the price to pay in order to gain a robust quantitative representation. For this reason, the scientist must be aware that the choice of going for a quantitative analysis translates into the adoption of only *one* of the possible narratives about the events to be represented. The gain obtained by the quantification is paid for by the impossibility of perceiving and representing within the same quantitative representation other aspects of the investigated system. These ignored aspects, however, may result relevant to other analysts or social actors, who have other purposes or disagree with the problem structuring adopted in the proposed quantitative analysis. Finally, there is an additional standard predicament affecting models of living system: the unavoidable presence of genuine ignorance. Quantitative analysts may become aware of the ignorance generated by their initial choice of reduction because of painful experiences – e.g. the discovery of prions in relation to mad cow disease.

2.2 Representing and Measuring System Performance in Agricultural Production in Substantive Terms

Scientific literature on technological progress in agriculture tells us about the impressive achievements of science-driven innovations in the field of production of food and fibres. For example, consider the technical coefficients characterizing chicken production: in advanced production systems less than 1.6 kg of feed input is required to produce 1 kg of poultry meat (output) in 35 days (Poultryhub.org, 2009), whereas chicken raised in the traditional way may require up to 8 kg of feed per kg of poultry meat produced (van Eekeren *et al*., 2006). This represents an increase in efficiency of 5 times! Clearly, this progress was not obtained by keeping all other elements of the production process equal; increased efficiency was obtained by moving chicken production from open-air yards to animal factories. "*With the use of incubator, chicken are no longer produced in parallel as in the old system dictated by nature . . . "Chicken farm" has thus become a misnomer: the situation calls for replacing it by "chicken factory". Because of the new system a pound of chicken sells in the United States for less than a pound of any other kind of meat, while in the rest of the world, where the old system still prevails, chicken continues to be "the Sunday dinner"* (Georgescu-Roegen, 1971, p. 253).

An even more spectacular result has been obtained with milk productivity of cows. The net milk produced per cow per day (yearly average) differs a lot among countries and production systems: It is below 1 litre in pastoralist societies, less than 5 litres in pre-industrial farms, more than 50 litres in modern "milk factories", and it can reach peaks of more than 75 litres in cases where production is boosted with hormones.

In these examples of chicken and milk production, we can readily calculate a number, i.e. a quantitative index characterizing, respectively, an increase in efficiency (an output/input ratio of chicken meat produced per unit of feed) and an increase in productivity (a given pace of milk flow per cow per day). But how useful and effective are these numerical indices in defining whether or not we are witnessing an *improvement*? Under which assumptions should we consider an increase in productivity or efficiency of a system of production an absolute improvement? This is an important question since *improvement* is a value-loaded term, the use of which implies normative consequences − scientists and politicians have to go for improvements! We can also formulate our question as follows: Given an index of performance, is it always true that *more of the same* is equal to *better*? That is, if we base our assessment only on the use of a single numerical index, how do we know *if* and *when* we are getting *too much of a good thing*?

To find the answer, we must ask ourselves another question: Is it only the final farm output that matters or should we also care for possible side effects generated by the process of production? Indeed, we must be aware that a higher efficiency in relation to just one of the production factors would represent an improvement only if the *ceteris paribus* assumption remains valid. That is, *if everything else remains the same, then* producing more per unit of production factor can truly be viewed as an improvement. Unfortunately, experience tells us that the *ceteris paribus* assumption is never valid for important technological changes in agriculture. That is, in complex adaptive systems the solution of a problem defined at a given scale, in the long term, will generate other types of problems. One of the names given to this systemic predicament is “Jevons paradox” (Polimeni *et al*., 2008), an issue that should be addressed in any sound discussion of technological progress. Of course, this does not mean that increases in productivity or efficiency in agriculture are necessarily bad. It simply implies that we must be careful when studying the performance of agricultural systems and always look for an integrated set of indicators that allows us to better understand the pros and cons in terms of side effects of the technical changes we want to introduce. In other words, we must *contextualize* the indications given by any single numerical index of performance.

That is, in which sense is one kilogram of poultry meat produced in a chicken factory *the same* as one kilogram of poultry produced in the farm backyard in the traditional way? What set of criteria (indicators of performance) could we use to answer this question? Similarly, the same search for side effects should be done when assessing the consequences of moving from an average production level of a few litres of milk per day by the family cow milked by hand in the backyard to an average production level of over 50 litres of milk per day per cow achieved in high-tech, large-scale dairy farms. If we are unable to account for all the side effects entailed by this move, then how reasonable is it to assume that a litre of milk is *the same* regardless of the system used to produce it? What is the meaning of the concept of *sameness* in this discussion?

Do not make the mistake of thinking that this abstract discussion is irrelevant for practical purposes. For example, the selection of criteria to define this very concept of *sameness* for the purpose of labelling food products did cause heated ideological debates. This debate was about the implementation of the controversial *principle of substantial equivalence* (Giampietro, 2002). This latter concept has played and still plays an especially important role in the regulation of genetically modified organisms (GMOs) in food production and the use of bovine growth hormone (BGH) to enhance milk productivity in cows (Giampietro, 2002).

In real life, innocent questions like whether a tomato produced from genetically modified plants is *the same* as a tomato produced from traditional and local varieties, or whether the milk produced by cows injected with a synthetic hormone is the same as milk produced by traditional cows, have led to ferocious ideological battles. Because of the high stakes involved, every single term used has been fiercely debated and it turned out to be necessary to get into philosophical questions such as:

- Is it possible by only comparing a few selected chemical substances, chosen from among thousands, to declare the substantive equivalence of two batches of milk?

- How reasonable is the claim of reductionism that it is only the chemical composition of the edible parts of a food product that matters?

This type of philosophical questions has been instrumental in defining and implementing the concept of substantial equivalence and in the final decision of the U.S. Food and Drug Administration (FDA) to not require labelling for milk produced with BGH or food products obtained from genetically modified plants. Thus epistemological discussions like those above may result crucial for the future survival of small organic farmers and dairy-farms and in relation to the health of our children.

We may ask ourselves then: How defendable is the use of reductionism in making these decisions? Imagine that by considering just a few chemical elements, such as the amino acids making up meat protein, a scientist concludes that pork is substantially equivalent to beef. Given this conclusion, should we no longer label meat as pork or beef simply because they were considered equivalent according to some arbitrary reductionist criterion? Should then Jews and Muslims no longer have the option to avoid eating pork because of the lack of a label (Giampietro, 2002)?

This example raises two important questions:

1) *What* criteria should be used for defining improvements in agriculture?
2) *Who* should choose these criteria and *how*?

These questions are addressed in section IV of this paper.

2.3 The Existence of Multiple Narratives about the Performance of Milk Production

I now elaborate the example of milk production in order to illustrate the need for multiple narratives in the description of agricultural performance. The example deals with the side effects that are necessarily associated (i.e., they must be expected!) with milk production in a ‘milk factory’ with an average production of over 50 litres of milk per day per cow. The example illustrates that if we want to describe an integrated set of expected features with a set of quantitative indicators and models (looking for possible causal relations) then we are forced to adopt an integrated set of different narratives referring to perceptions dealing with different levels of organization and different dimensions of analysis. Indeed, an integrated characterization of the performance of milk production has to be carried out using *simultaneously* various narratives defined across different scales and disciplines.

Six narratives that are relevant for assessing the performance of milk production are presented in Box 1. Be aware that these are narratives, not scientific facts! By definition they entail the expression of a given point of view about the issue to be tackled. Note that in the first narrative of this list, which focuses exclusively on the quality of the final product, it is assumed that the quality of one litre of milk is always ‘the same’ regardless of its source of production. However, as noted earlier, even if we accept this reductionist narrative, which is behind the claim of substantial equivalence, it is obvious that the relative process of production can (and should) be characterized anyhow by considering several other attributes of performance, which are also relevant for the issue of sustainability.

****

BOX 1 Narratives describing relative attributes of the performance of a milk production system with an average productivity of more than 50 litres of milk per day per cow.

NARRATIVE A: focusing only on the product

For this exercise we assume by default that the composition of milk is the same across different processes of milk production. Thus we assume that the quality of the milk produced by a cow with a low productivity (< 10 liters/day) and living in 100% natural conditions (grazing on open pastures, not receiving any drugs or other chemicals) *is exactly the same* as the milk produced by a cow with a huge productivity (> 50 liters/day) living on a high-tech dairy farm. This assumption is of course doubtful to say the least. For example, at high levels of production cows tend to develop mastitis and consequently traces of antibiotics and other drugs may be found in the milk (Millstone *et al*., 1994).

NARRATIVE B: regarding the physiology of the cow

A relevant attribute in relation to this criterion is the life expectancy of the cow. The life expectancy of a dairy cow used to be about 20 to 25 years for non-stressed animals. However, in modern high-tech dairy farms, the productive life span of milk-producing cows has dropped to only 2 to 4 years. In fact, an extremely high level of milk production entails that cows produce milk at the expense of their own body and, consequently, productive life is limited by induced calcium deficiency. Moreover, the excessive intake of feed protein tends to transform these cows from ruminants into monogastric animals, with the concomitant generation of ketosis. In conclusion, highly productive cows risk being affected by lameness and becoming “downed animals”. (The narrative and relative data are from Farm Sanctuary, 2005).

NARRATIVE C: Regarding the typology of the farm

Modern milking parlors are linked to modern systems of refrigeration and storage of milk, and this translates into high fixed investment costs. To get full advantage of both economies of scale and the edge provided by technology, dairy factories have to invest a huge quantity of capital per worker. For this reason high-tech dairy farming is not an economic activity that provides a lot of employment! Economies of scale also entail a very large number of cows per farm. In 2007 almost 50% of U.S. dairy farms had more than 500 cows (USDA, 2007). Large ‘milk factories’ may count up to tens of thousands of cows. Consequently, these ‘factories’ also consume a huge flow of high-quality, protein-rich feed, which calls for a total openness of the system. Feed must be imported from the market and this implies full dependency on market prices for the farm's economic viability. Finally, the huge flow of feed input consumed in the milk factory generates a corresponding huge flow of manure to be disposed off. This flow of waste is extremely concentrated and generates problems of ecological compatibility in relation to several chemical elements; the most notorious problem being the excess of phosphorous and nitrogen in the water table.

NARRATIVE D: Regarding the economic performance of the agricultural sector

High levels of capitalization within the agricultural sector have translated into a systemic lack of economic viability. In fact, it is well known that the agricultural sector in modern economies has the lowest return on the investment and among the highest requirements of investment per worker (Giampietro, 2008; Giampietro and Mayumi, 2009). This weakness in economic performance translates into the necessity of heavy subsidies from central governments for dairy farms, a well-known economic and political issue in both the European Union (EU) and the United States (USA). Clearly, exceptions exist and it is possible to find special niches in which dairy farms enjoy comparative advantages due to the endowment of natural resources and low demographic pressure.

NARRATIVE E: Regarding international trade

Given the common characteristic of openness of high-tech milk production systems and the systemic problems of economic viability of agricultural production, it is unavoidable that a massive production of milk (and animal products) in developed countries increasingly relies on imported feed from developing countries. The rapid expansion of *la Republica Unida de la Soja* in South America is a clear example of this trend. To study this side effect of high-tech dairy farms in highly-developed countries on local traditional farming systems in less-developed countries and the concomitant massive changes in land use, one must resort to scientific disciplines, such as ecology, sociology, and anthropology, that are not focused on the technical aspects of agricultural production.

NARRATIVE F: Regarding human health

Considerations of human health are of two types, one regarding the milk-producing countries and the other the animal feed-exporting countries. In developed, milk-producing countries the quality of the diet is seriously affected by *excesses*, most notoriously the excessive intake of saturated fat from dairy products and red meats and the excessive intake of simple sugars from corn syrup, poured indiscriminately into the US food system. The omnipresence of corn syrup in the US food system (from sodas to sausages) is a side effect of the use of corn-silage for animal feed. Sugar, in the form of corn syrup, must be extracted from corn silage during animal feed production to obtain a sufficiently high protein/energy ratio. Since corn syrup does not crystallize, it is difficult to market as such, and as a consequence, to get rid of it, it has to be incorporated in food products.

In feedstock exporting countries the massive production of feed crops increasingly induces adverse effects on human health due to impoverishment in the local diet (the best land is used to produce feed crops for export rather than for local food) and negative side effects associated with intensive production of monocultures, such as exposure to residues from fumigation and other types of pesticides and deterioration of water quality (Wasley, 2009). To study the causal effects within this narrative one should use the expertise from several scientific disciplines (nutrition, epidemiology, medicine, eco-toxicology, ecology) that are not traditionally focused on the technical aspects of agricultural production.

NARRATIVE F: Regarding the environment

Also in relation to the environment, effects are different for developed and developing countries. In developed countries, where feed production only takes place on good productive land where full advantage can be taken of technological inputs, the main environmental problems are the excessive load of phosphorous and nitrogen in the water table and contamination of the environment with pesticides residues. A serious complication of the former problem is the formation of black spots in the sea near the mouth of major rivers discharging the surplus of nutrients from agricultural run-off. In developing countries, there are two major, additional problems (Altieri and Pengue, 2005): (i) the practising of intensive monocultures on marginal lands (e.g., in South America the continuous expansion of intensive production of monocultures outside the pampa), which implies the loss of fragile soil; (ii) the rapid expansion of feed monocultures and plantations into non-colonized land, resulting in deforestation and loss of local biodiversity through destruction of habitats.

********************************** end BOX 1***

In relation to any one of the six narratives listed in Box 1, we may assume that we can establish a relation between changes recorded by a given indicator, e.g. the intensity of milk production per cow, and changes recorded by other indicators. For example, if we were to decide whether or not to raise milk productivity from 5 to 50 litre/day per cow, then we should consider in the relative process of deliberation that several other measurable attributes of performance, beside the quality of the milk, will be affected by this change. That is, the final supply of milk and its intrinsic quality are but two of many relevant criteria. Some of these attributes of performance are directly related to the process of milk production (physiology of the cow, structure of the dairy farm), others are related to different dimensions of analysis (social, economic, technological) and refer to characteristics of the context within which the milk production takes place. Some attributes, such as the health of individual cows, can only be observed by adopting a very specific and local scale, while others, such as the implications for South America of feed import policies enforced in Europe, can only be observed when adopting a very large scale. Reductionism makes it possible to measure each one of these attributes of performance within the corresponding chosen narrative. However, what reductionism can *not* do is judging or deciding which mix of narratives is useful and/or relevant for an informed deliberation about sustainability.

The three points to be driven home from this discussion are:

1) When dealing with the sustainability of agriculture, several non-equivalent narratives have to be used simultaneously to study the effects of changes in relation to different dimensions of analysis at different scales;

2) In complex adaptive systems everything is related to everything else and, therefore, it is dangerous to rely on too radical strategies of reductionism. Optimizing strategies based on just one single indicator of performance, corresponding to one given narrative, necessarily tell only a small part of the whole story and are likely to generate failures when overstretched (e.g. maximizing milk productivity according to narrative A).

3) A careful procedure of integrated assessment is the only way to address this predicament.

3. The challenge of simplifying complexity

3.1. Basic Theoretical Concepts

3.1.1 Complexity and Relevance

By definition a complex phenomenon is a phenomenon which can and should be perceived and represented using simultaneously several narratives, dimensions, and scales of analysis (Simon, 1962, 1976; Rosen, 1977, 2000; Salthe, 1985; Ahl and Allen, 1996; O'Connor *et al*., 1996; Funtowicz *et al*., 1999; Allen *et al*., 2001; Giampietro, 2003).

This definition of complexity resonates with the concept of Kolmogorov-Chaitin of complexity applied to mathematical objects. According to this concept, complexity can be associated with the *im*possibility of using an algorithm to compress the information required for a given representation without loosing valuable information. Observe that the definition of the relevance of the information that should not be lost *does not depend on the characteristics of what is observed*, but it *depends on the interest of those that will use the quantification*. The famous example provided by Chaitin (1975) is the following one. To save the cost of a telegram for an alien friend living outside our galaxy, one could compress the information needed to represent the trigonometric tables by writing down just Euler's equation. In fact, from that equation the alien friend can recover all the information contained in even the largest trigonometric table. "*Suppose, on the other hand, your friend is interested not in trigonometry but in baseball. He would like to know the scores of all the major-league games played since he left the earth some thousands of years before. In this case it is most unlikely that a formula could be found for compressing the information into a short message; in such a series of numbers each digit is essentially an independent item of information, and it cannot be predicted from its neighbours or from some underlying rule. There is no alternative to transmitting the entire list of scores*" (Chaitin, 1975, p. 48).

This conceptualization of complexity entails that the definition of *what is complex* for quantitative analysis depends on a preliminary definition of *what is relevant* for those that will use the analysis. In the same line of reasoning, Rosen claims that complexity is not a property of the observed system, but rather of the process of observation. "*A stone can be a simple system for a person kicking it when walking in the road, but at the same time be an extremely complex system for a geologist examining it during an investigation of a mineral site*" (Rosen, 1977, p. 229). That is, the usefulness of a given representation/quantification depends on *why* and *how* one decides to observe the system in the first place. If we overlook this important aspect and randomly apply complicated mathematical models to complex problems without having examined the underlying semantics in the pre-analytical step, we risk employing analytical tools outside their domain of applicability.

3.1.2 Trade-off between compression and representation

When dealing with a complex issue such as the sustainability of farming systems, the power and strength of quantitative analysis entails a potential weakness. This point is beautifully explained by Box (1979), under the heading: "*all models are wrong, but some are useful*". When discussing the usefulness of quantitative models Box says: "*For such a model there is no need to ask the question "is the model true?". If "truth" is to be the "whole truth" the answer must be "No". The only question of interest is "Is the model illuminating and useful?""* (p. 202-203). This consideration points directly at the pre-analytical definition of the purpose of the analysis: Why are we doing this analysis in the first place? Whose problems are addressed by the analysis? Which problems are being ignored because of the choices made when selecting this model? These questions are crucial since different social actors carry different priorities in relation to the problems to be addressed.

The incredible power of quantitative analysis needs to be 'tamed' before being used. This process of taming necessarily requires a quality check on the semantics behind any number used in a quantitative analysis. In spite of the obvious relevance of this observation, the epistemological challenge entailed by multiple scales and multiple dimensions seems to be ignored by the vast majority of people involved in quantitative analysis (Giampietro *et al*., 2006a,b; Munda, 2006), including those working in the field of sustainability of agriculture (Giampietro, 2003). Those making quantitative analyses always seem to rush into their favorite activity of crunching numbers with fancy models, optimizing this and maximizing that, without taking the time to reflect on the soundness of the basic narratives that they implicitly accepted as valid when choosing their models and data.

Perhaps the best metaphor of this fact is given by the use we make of maps. Here we can recall Korzybski's most famous premise "the map is not the territory itself". Maps are handy exactly because they do not contain *all* the information and details that are present in the actual territory. The compression entailed by the making of a map allows us to use a limited amount of information that can be conveniently handled in finite time span. The issue here is that the information carried by the map must match the information required to perform the task for which the map was generated in the first place. What defines the quality and validity of a map (or a mathematical model) is *the choice of how* to compress the virtually infinite information that we can gather from the external world into a finite representation that is *useful* for a given task.

FIGURE 1 – here

To grasp the importance of this challenge we should go back to the roots of the epistemological predicament, eloquently described by great philosophers of the past.

Plato, in his famous allegory of the cave (in *The Republic,* 360 B.C.), describes human observers as prisoners chained in a cave who cannot directly see the external world but can see only distorted shadows of objects projected by a big fire against the wall of the cave (Fig. 1, upper-left quadrant). A similar message is transmitted by the ancient Buddhist and Hindu tales of the blind persons touching an elephant (Fig 1, upper-right quadrant, the Buddhist version). In both tales the message is the same: Blind people touching different parts of the elephant feel different characteristics and therefore will be convinced that the elephant *is* different things. Thus, depending on their particular perception of the external world, non-equivalent observers, even if operating in that same external world, will represent different perceived realities using different narratives. Buddha's wise comment on the tale aptly describes the majority of scientific discussions about sustainability: "*Just so are these preachers and scholars holding various views blind and unseeing.... In their ignorance they are by nature quarrelsome, wrangling, and disputatious, each maintaining reality is thus and thus".* (http://www.cs.princeton.edu/~rywang/berkeley/258/parable.html).

This same epistemological predicament is found over and over again also in modern epistemological discussions. For example, in his introduction of the *Tractatus* of Wittgenstein (1974), Bernard Russell explains one of the ideas of that book by saying that the same solid can generate different projections. Each one of these different linguistic expressions will provide only one aspect of the observed world at the time (Fig. 1, lower-left quadrant). To convey the same concept Bhome (1995, p. 187) uses the example of the non-equivalent perceptions of the movement of a fish obtained by different cameras looking at the same aquarium form different angles (Fig. 1, lower-right quadrant). The message of all these thinkers is the same: whenever we use only one given finite set of observable qualities – a quantitative model associated with a formal representation – we are necessarily focusing only on a subset of the potential set of observable qualities which may be relevant for sustainability.

3.2. Rosen's Theory of Modelling Relation

Robert Rosen (1977, 1986, 2000) developed a theory of modelling relation to address the systemic epistemological problems associated with quantitative modelling. Rosen emphasizes the importance of making the following distinctions:

- The distinction between *the reality*, which cannot be known in substantive terms, and a *given perception of the reality,* which depends on the story-teller's observation generating the perception;
- The distinction between the *perception of the reality* (inside the mind of the story teller) and its *representation* (the formalization used in the model).

Latter distinction is at the heart of the issue of scale, the epistemological implications of which have been explored extensively in the field of complex systems theory, especially by those working on hierarchy theory (Simon, 1962; Koestler, 1968, 1978; Pattee, 1973; Allen and Starr, 1982; Salthe, 1985, 1993; O'Neill *et al.,* 1986; O'Neill, 1989; Allen and Hoekstra, 1992; Giampietro, 1994, 2003; Ahl and Allen, 1996; Giampietro and Mayumi, 2004; Giampietro *et al*., 2005).

The theory of modelling relation proposed by Rosen helps understanding these two conceptual distinctions and for this reason, I will briefly illustrate his theory here. An overview of the operations performed by a modeller, according to Rosen's theory, is given in Figure 2.

FIGURE 2 – here

STEP 1 – choosing a relevant narrative about the system we want to model.

In order to go through this first step, we must understand that any model must have a goal. The goal of the model is needed in order to select a relevant narrative. A narrative is a given perception of causality over a set of finite events leading to the choice of a given representation. The narratives listed in Box 1 may illustrate this connection. For example, narrative B establishes a causal relation between a short life expectancy of the cow and high levels of milk productivity. In accordance with this narrative, we have to adopt a representation dealing with cow physiology. Narrative E, on the other hand, establishes a causal relation between the expansion of soya plantations in South America and the growing import of animal feeds in Europe. In accordance with this narrative we better adopt a representation addressing the effects of international trade. Thus, the storyteller selects a narrative depending on the goal of the model. The narrative has to make it possible to simulate a perceived relation of causality over some relevant attributes of a natural system (the system to be modelled). Again, this ability of explaining and/or simulating is a must for the model. In fact, in order to result valid a model has to make predictions or provide explanations that are relevant (or 'useful' in Box's language) for those that will use the model.

It is important to realize that the first step of choosing a relevant narrative occurs in the *pre-analytical* phase. This step is illustrated in Figure 2 by arrow #1. This step is hardly ever explicitly declared, described and defended by the modeller. In practical terms, selecting a narrative about the observed system translates into deciding 'what the system is' and 'what the system does' within the model in relation to a relevant relation of causality to be explored. This pre-analytical decision is required in order to be able to proceed with step 2, the quantitative representation.

STEP 2 –Formalization: interfacing the narrative and representation with the external world

In order to use an inferential system for the modelling (crunching numbers), modellers have to formalize the chosen narrative into a given representation. They do so by assigning a formal identity to the natural system to be modelled in terms of a series of entities and processes that can be described using proxy variables that, in turn, can be measured. This step of *encoding* (indicated by arrow #2 in Figure 2) is about deciding 'what the system is' and 'what the system does' within the quantitative representation. In other words, in this step the modeller has to decide which variables and equations will represent the processes considered as relevant. The modeller must now confront the narrative (arrow #1) and the representation (arrow #3), both defined in the human mind, with signals coming from the external world. The chosen narrative and the chosen representation have to be compatible with the existence of viable measurement schemes – tracking external referents when choosing proxy variables – thus making it possible to gather the required data in an appropriate time.

STEP 3 – Crunching numbers

A formal inferential system, with axioms, production rules and algorithms (i.e., the chosen mathematical model represented by arrow #3 in Figure 2) has to be selected and operated in order to simulate the perceived pattern of causality (arrow #1) over the numerical values taken by the finite set of proxy variables (output of arrow #2).

STEP 4 – Validation: interfacing the narrative and the representation generated in the model with the external world

To check the usefulness of the model the modellers have to take an additional, crucial step: that of decoding (indicated by arrow #4). Once again, elements operating in the human mind (perceptions and representations) have to be confronted with the signal coming from the external world. The decoding step refers to the process through which the predictions obtained by the mathematical model (arrow #3) – based on the representation by proxy variables (arrow #2) – are used to predict or explain the perceived behaviour (according to the chosen narrative, arrow #1) of the modelled natural system. This step has the goal of validating the model in relation to its usefulness; i.e. checking whether or not the predictions and explanations may be used effectively. This step is crucial because it decides whether or not the model will be used to guide action according to its predictive power in relation to the chosen narrative.

In his original formulation Rosen suggests that a valid model should generate a *commuting diagram* or concordance over the four arrows in the loop (Fig. 2). When this occurs, a given *perception* of causality given by arrow #1 (the semantic of the chosen narrative) is adequately *represented* by the other three arrows. Then, the chosen representation (arrow #2), the syntax of the model (arrow #3), and the process of interfacing this information with the external world through sampling and measuring (arrow #4) makes it possible to make predictions and check their validity. The encoding and decoding steps (choice of a representation and criteria for sampling) are needed to guarantee a quality control on the pre-analytical choices of both narrative (perception of causality) and representation (the variables used in the quantitative model).

The various steps in the cycle of the modelling relation described in Figure 2 refer to different types of tasks and hence imply different types of decisions, each of which requires a different type of quality control.

Step 1 (arrow #1 in Fig. 2) is about choosing the relevant narrative in presence of various possible narratives carried by different social actors (see Box 2). It is about defining values, power relations, and social agreement: what are the issues and the causal relations that matter to deal with existing problems? The quality of the decision regarding this first step of the modelling relation, i.e. the relevance of the story, depends on the purpose of the observation. Why are we studying the system in the first place and why are we defining its performance? Quality control in this pre-analytical phase is purely semantic.

****

Box 2

The various narratives that can be used to describe in a quantitative terms a person.

The very same observed system, i.e. our given person, can be:

- a father, when studied by his children;
- a tax payer, when studied by the Internal Revenue Service (IRS), the US national agency implementing tax laws;
- a potential goalkeeper, when studied by Pep Guardiola looking for a new soccer player;
- a potential dinner, when studied by a starving lion;
- a nice environment where to reproduce, when studied by a tapeworm.

Needless to say that these five different story-tellers when looking for a model useful for describing this very same person, would end up with completely different choices of variables and inferential systems in order to represent what that person is and does.

****

Quality control in step 2 of the modelling relation (arrow #2, encoding) regards the process of interfacing the chosen semantic with a pertinent syntax. How to properly represent the chosen narrative in a useful way? Quality control in this step involves tracking relevant relations of causality through multiple scales and multiple dimensions. For example, using the metaphor of Box 2, a representation useful to a tapeworm story-teller will include in the model variables such as temperature, chemical composition of the environment, and presence/absence of predators. On the contrary, if the representation is to be of any use to the *IRS inspector*, then the model should include variables such as income sources, household structure, and deductable expenses.

Step 3 of the modelling relation is purely analytical and here quality control regards the appropriateness of the choice of the inferential system (the mathematical model) to represent and predict the behaviour of the modelled system. Thus, in this step of the modelling relation we deal with the quality of the *syntax* of the model. How to properly perform the tasks of data collection (sampling and measurement) and data processing (calibration and operation of the chosen model, checking its suitability to handle the chosen representation)? Here, the rigor and the pertinence of the choice made by the scientists can be checked with formal protocols (e.g., sensitivity analysis) for both data gathering and data processing.

Step 4 in itself is about performing an overall quality check on the whole process. That is, after interfacing the proposed coupling of a narrative (arrow #1) considered as relevant with the relative quantitative representation (arrow #3), using external referents (data gathered from the external reality in the step #2) we finally have the opportunity of checking the usefulness of the whole procedure: does the chosen model predict or explain practical situations, when going through step 4? The choices related to step 4 (arrow #4) and step 2 (arrow #2) are needed to link narratives and representations to the information coming from external referents, by interfacing semantic to syntax.

4. The Challenge of Multi-Criteria Analysis

4.1 There Is No Such Thing as the “Right” Set of Sustainability Indicators

An example of a multi-criteria characterization of the performance of a mixed crop farming system in China's province of Hubei is given in Fig. 3 (from Giampietro, 2003). The two radar diagrams shown in this figure – nowadays popular in sustainability analysis – represent in a graphic way an integrated characterization of the performance of two elements, a typology of household (left) and a village (right), both operating in the same farming system.

FIGURE 3 - here

For each one of the four criteria of performance adopted in this analysis, three indicators are selected. The resulting multi-criteria space (12 indicators for 4 criteria) is obtained adopting the flag-model (Nijkamp and Ouwersloot, 1998; Nijkamp and Vreeker, 2000). This is a system of benchmarking indicating the performance for each indicator by assigning different colours to different ranges of values. In this way the value taken on by each indicator is contextualized by the colour of the range it falls into: green is good, yellow is acceptable and red is unsatisfactory. Within this graph, values falling inside the inner circle (= unacceptable) are considered to lie outside the viability domain. A detailed overview of methods of graphic representation to generate an integrated analysis of farming system performance is available in Gomiero and Giampietro (2005).

In Fig. 3, the same type of graph is used to characterize the performance of: (i) a relevant typology of household (a part – on the left); and (ii) the village to which the various household types belong (the whole – on the right). The chosen indicators reflect the choice of a set of narratives about the performance of this farming system typology – i.e. mixed crop farming in the Hubei province of China (Giampietro and Pastore, 1999, 2001; Pastore *et al*., 1999). The chosen characterization of performance makes it possible to characterize the differences among household types, by using the chosen mix of indicators and benchmarks. This characterization reflects the definition of relevance provided by the farmers (upper left quadrants). At the same time the characterization of the performance of village is based on indicators reflecting the definition of relevance given by the local government (lower level quadrants).

Because of the special tailoring of this integrated representation of performance to a specific situation, the radar diagram shown in Fig. 3 – based on the chosen set of criteria, indicators and targets – will result pretty useless to describe the effect of possible changes in a farming system operating in Germany. An analyst willing to apply the same approach to a case study in Germany has to deal with other typologies of crops, farmers, and a different economic and ecological context. This is to say, that scientists can provide useful and relevant information to farmers, but in order to be able to do so, they first have to check what is relevant for the farmers and the other social actors in the given context. This is to say, that the idea of adopting a set of sustainability indicators to measure the performance of different agricultural system of the type “one size fits all” should be considered as a myth.

What happens when we are dealing with two farming systems expressing markedly different characteristics? To deal with this case, let’s consider a simple metaphor. Imagine we have to evaluate the performance of two different makes of cars: a Ferrari and a Trabant. In general terms, we may characterize the performance of a make of car using different criteria (this example is still based on four criteria: economic criterion, safety criterion, driving criterion and status-symbol). Then for each criterion, three indicators are used to define a given performance space – as illustrated in Fig 4. Then, the identity of this performance space does reflect a pre-analytical choice of adopting a set of relevant narratives about the performance of cars associated with the choice of criteria of performance. In this example the two makes of car are markedly different in their performance and this difference can be considered analogous to the systems of milk production discussed earlier: The Ferrari is comparable to the high-tech milk factory with thousands of cows, achieving production levels of more than 50 liters of milk per day, whereas the Trabant is comparable to the back-yard production system with a single cow happily feeding on locally-produced crop residues and pasture and producing a mere 5 liters of milk per day.

FIGURE 4 – here

Using this analogy, we find again the same pattern discussed before: if we want the top performance of Ferrari in terms of speed and status, then we have to expect high input consumption and high economic costs. Analogous, a milk production system with a very high output will require a large economic investment, thousands of cows, and it will imply a large environmental impact. The opposite is true for the Trabant. This confirms the fact, that when dealing with a sound integrated assessment, the various indicators change in a coordinated way across the performance space ("*ceteris* are never *paribus*").

A cursory reflection of the characterization illustrated in Fig. 4 also shows that those interested in buying a Trabant and those interested in buying a Ferrari – different story-tellers living in two different "perceived realities" – will find the type of integrated analysis provided in Fig. 4 useless when actually going out shopping for the right car. Those interested in buying a Trabant are likely to look only at an integrated set of indicators that allows a comparison, with the required accuracy, between the Trabant (one possible option) and other low-cost utilitarian vehicles (the other options), such as the Tata Nano. The same is the case with systems of milk production, a small farmer having only one cow and running a farm of less than two hectares, will look for improvements that can still produce the required feed within the farm and that maintain cow life expectancy at more than ten years.

Thus, we must conclude that criteria and indicators relevant for discussing sustainability in general terms – characterizing the pros and cons of a Ferrari and a Trabant across various narratives – are not particularly useful for studying relevant differences in relation to a specific task in a specific situation – to help the decision of individual farmers operating within a specific farming system.

4.2 The Crucial Importance of the Decisions Made in the Pre-analytical Step

Multi-criteria analysis (MCA) is nowadays a well-developed scientific discipline with an abundant literature (Zeleny, 1982; Bana e Costa, 1990; Nijkamp *et al*., 1990; Vincke, 1992; Roy, 1996; Beinat and Nijkamp, 1998; Jansen and Munda, 1999; Hayashi, 2000; Bell *et el*., 2001; Belton and Stewart, 2002). An overview of the state of the art in this field has been provided by Figueira *et al.* (2005). The final output of a process of MCA generates a final quantitative representation of the problem in the form of an impact matrix used to carry out a MCA (which is the equivalent to the information visualized in the form of a radar diagram of the type illustrated in Figures 3 or 4). However, this formalization of the problem has the effect of limiting the focus on a finite (and small) set of relevant attributes and a finite (and small) set of policy options. That is, it still entails the same problem of simplification of complexity.

FIGURE 5 – here

The key importance of the pre-analytical step within the modeling relation (the simplification required for moving from a practical situation experienced in the external world to a well defined impact matrix) is illustrated in Fig. 5. The graph represents the dramatic compression of the information flow that is potentially available for characterizing a given issue, when defining a given problem structuring of a sustainability issue in the form of an impact matrix (in the middle of the figure). According to the soft-system methodology (Checkland, 1981; Checkland and Scholes, 1990), when dealing with a messy reality that can only be perceived using a virtually unlimited number of perspectives, goals, narratives, and relevant attributes, the most important step of any integrated assessment is the first one: the movement from a virtually infinite information space to a finite amount of information. For this reason rather than multi-criteria analysis one should use the expression societal multi-criteria evaluation (Munda, 2005, 2006, 2008)

The "mission: impossible" of compressing a virtually infinite information space into a finite information space carries ethical implications because of the existence of asymmetries in power relations within human societies. For this reason, implementation of institutional procedures would be warranted to guarantee the quality of the process of information compression (Röling, 1994; Röling and Wagermakers, 1998). Who has the authority to decide *how* to compress the universe of possible stories, narratives, and values found in any practical issue of sustainability into a simple impact matrix? As a matter of fact, the information compression in MCA obtained by applying the strategy of reductionism shares the same ethical problems with cost/benefit analyses (Munda, 2006). Every time we deal with sustainability issues we should first of all answer the following set of questions: What do we want to sustain, for whom, for how long, and at which cost? (Allen *et al*., 2003) When addressing these questions, quantification always results problematic: *"We should reject such a simplification of complexity and exclusion of values, favouring instead the acceptance of a plurality of incommensurable values. In decision-making processes, economics becomes a tool of power. This is the case when applying cost-benefit analysis to individual projects, and also at the level of the macro-economy where increases in GDP trump other dimensions. The question is,* ***who has the power to simplify complexity and impose a particular language of valuation?****"* (Martinez-Alier, 2008, p. 30).

A second complication in the pre-analytical phase is that the perception of the relevant social actors and the representation proposed by the scientists can not always be easily confronted. In many occasions the two groups do not share meaning about the chosen external referents. As discussed earlier, the quality check on the usefulness of a model can only be based on confronting the correspondence between a given perception about the reality and a given representation of the reality. I will illustrate this problem with an example.

Imagine that a team of distinguished scientists goes to a different planet to study a production system adopted by a rural community inhabiting that planet. The production system will almost certainly be based on some set of goals and narratives developed by that unknown civilization within its local system of knowledge. However, this information is unknown or not easily available to the team of experts. The local community has likely organized its activities in the past using a set of external referents which allows *them* to verify the usefulness of their models. The scientists from Earth can only develop models using the set of goals and narratives about the performance of agriculture which are typical of the knowledge system to which they belong. Therefore, after building the model (arrow # 2 and arrow # 3) and performing the step of decoding (arrow # 4 – when deciding how to apply the quantitative results to guide action) they are likely to discover that the usefulness of their model refers to a different definition of "relevant reality" and "eternal referents" different from those used by the local civilization (the choice of arrow # 1 is different from the one used by the locals).

5. Conclusion: Implications for the integrated assessment of farming systems

5.1 The Key Distinction between Assessment and Analysis

According to the definition given by the Millennium Ecosystem Assessment (MEA, 2005) an assessment must go through three distinct quality checks: those related to its scientific credibility, its political legitimacy, and its practical usefulness for guiding action. Starting from this definition, we can appreciate the crucial difference between an *assessment* and an *analysis*; a difference that has important implications for the use of science for governance in the field of sustainability.

An assessment implies the acknowledgment and handling of: (a) the unavoidable co-existence of non-equivalent perceptions of the same problem caused by contrasting but legitimate goals and values held by various social actors; (b) the need of using simultaneously quantitative representations referring to different scales and dimensions, which are therefore not reducible to each other using algorithms; (c) the unavoidable presence of uncertainty when dealing with the future. Thus, an assessment must address not only analytical issues but also normative issues referring to situations which are all 'special'. For this reason it is impossible to have a sound deliberation without addressing explicitly and in an adequate way the specificity of the context in which the policy will be implemented. Recalling the implications of the modelling relation proposed by Rosen, an assessment must refer to the *overall* operation of the model across all four steps. Consequently, quality checks have to be applied to the different typologies of decisions made in relation to both the semantics and the syntax adopted.

An analysis, on the other hand, is related to the production and interpretation of models and indicators associated with a given typology of problems and situations. An analysis deals only with the handling of the descriptive side of an assessment, which is based on the study of typologies of situations.

This distinction entails that the quality of a scientific analysis used in a specific situation for governance (e.g., as an input for normative purposes) depends on the existence of an uncontested agreement among those that will use the quantitative results over the various choices made by the modellers. Thus, the uncontested agreement must refer to:

- The pre-analytical choice of goals and narratives legitimizing the given issue definition (when deciding what is the system under study and what are the relative narratives) and the given problem structuring (when deciding what are the relevant attributes of performance for the system under study);
- The series of quality checks on the decisions referring to the semantics used in operating the modelling relation. These decisions refer to the choices made in the steps of encoding (arrow #2 in Fig. 2), i.e. choice of variables and measurement scheme defining the quality of the data, and decoding (arrow #4 in Fig. 2), i.e. the choice of an external

referent for the validation of the model when deciding *how useful* the quantitative results are for applications to practical situations with normative purposes.

Especially, in case of evolution −a key aspect of sustainability – we are dealing by definition with adaptive, learning systems which become “something else” in time − the issue of uncertainty (and ignorance) can become overwhelming in relation to this last quality check.

5.2 The Epistemological Predicament of Sustainability and Post-Normal Science

When dealing with a quantitative analysis of the sustainability of agriculture it is inevitable to find legitimate but contrasting perspectives (disagreement among story tellers) about the set of narratives that should be used, as well as large doses of uncertainty and ignorance (impossibility to judge the usefulness of the quantitative output of the model). This systemic impasse is at the basis of the paradigm of post-normal science (Funtowicz *et al*., 1998; Funtowicz and Ravetz, 1990, 1991, 1993, 1994). According to this paradigm, decision-making related to sustainability should no longer be based on the traditional praxis: individuation of optimal solutions, through analytical processes carried out by scientists or through discussions carried out by panels of experts. The predicament of reductionism requires the implementation of procedures aimed at guaranteeing a quality control on each one of the four steps required to implement the modelling relation. That is, when dealing with sustainability issues we should “deliberate about how to act wisely”, rather than “finding the optimal thing to do”. Looking for wisdom in face of uncertainty requires reflexivity.

Being autopoietic systems (i.e., systems making themselves), reflexivity is a key characteristic of human societies. Indeed, human societies have to continuously re-discuss how to make themselves. As a consequence, both the solutions to perceived problems and the viability of policies to be implemented cannot be found within the recorded knowledge. They have to be created in the process. This implies that reflexivity (deliberation, negotiation, agreement among social actors) can generate options that were not yet available within the set of perceptions and narratives used by scientists in the building of their models. In that sense, reflexivity creates new knowledge, since it makes it possible for the social system to generate options that were not accessible before. This is a crucial point for sustainability.

Considering the severity of the sustainability predicament that we are facing in agriculture, the creation of something new is exactly what is needed. Humankind does not need to continue seeking and finding optimal solutions based on an old set of problem structuring and an old set of issue definitions; it rather needs the creation of a new option space within which it is possible to develop something completely different.

Acknowledgement

The author gratefully acknowledges the financial support received from EU funded projects: (i) "Synergies in Multi-Scale Interlinkages of Eco-Social Systems" (SMILE, Contract 217213-FP7-2007-SSH-1), and (ii) "Development and Comparison of Sustainability Indicators" (DECOIN, Contract 044428-FP6-2005-SSP-5A) and the Research Council of Norway for the funds of the Project no. 187969/S10, Reflexive Systems Biology. Many thanks are also due to Sandra G.F. Bukkens for editing and polishing the manuscript.

References

Ahl, V. and Allen, T. F. H. 1996. *Hierarchy Theory*. Columbia University Press.

Allen, T.F.H. and Hoekstra, T.W. 1992. *Toward a Unified Ecology*. Columbia University Press, New York.

Allen, T.F.H. and Starr, T.B. 1982. *Hierarchy: Perspectives for Ecological Complexity*. University of Chicago Press, Chicago.

Allen, T.F.H., Tainter, J.A., and Hoekstra T.W. 2003. *Supply-side sustainability.* Columbia University Press, New York.

Allen, T.F.H., Tainter, J.A., Pires, J.C., and Hoekstra T.W. 2001. Dragnet Ecology, "Just the facts Ma'am": the privilege of science in a post-modern world. *Bioscience* **51**: 475-485.

Altieri, M.A. 1995. *Agroecology: The Science of Sustainable Agriculture.* Westview Press, Boulder, CO.

Altieri, M.A. 2004. Linking ecologists and traditional farmers in the search for sustainable agriculture. *Frontiers in Ecology and the Environment* **2**: 35-42

Altieri, M.A. and Koohafkan, P. 2008. *Enduring Farms.* Third World Network.

Altieri, M.A. and Pengue, W. 2005. *GM Soya Disaster in Latin America.* ISIS Report 06/09/05. http://www.i-sis.org.uk/SDILA.php

Bana e Costa, C.A. (editor) 1990. *Readings in Multiple Criteria Decision Aid.* Springer-Verlag, Berlin.

Beinat, E. and Nijkamp, P. (Editors) 1998. *Multicriteria Evaluation in Land-use Management: Methodologies and Case Studies*. Kluwer, Dordrecht.

Bell, D.E., Hobbs, B., Elliot, E.M., Ellis, H., and Robinson, Z. 2001. An evaluation of multi-criteria methods in integrated assessment of climate policy. *Journal of Multi-Criteria Decision Analysis* **10**(5): 229-256.

Belton, V. and Stewart, T.J. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer, Dordrecht.

Bohm, D. 1995. *Wholeness and the Implicate Order.* Routledge, London, third edition, 224pp.

Box, G.E.P.1979. Robustness in the strategy of scientific model building. **In**: *Robustness in Statistics,* pp. 201-236. Launer, R.L. and Wilkinson, G.N., Editors, Academic Press, New York.

Chaitin, G.J. 1975. Randomness and mathematical proof. *Scientific American* **232**(5): 47-52.

Checkland, P. 1981. *Systems Thinking, Systems Practice.* John Wiley, Chicester.

Checkland, P. and Scholes, J. 1990. *Soft-Systems Methodology in Action.* John Wiley, Chicester.

FAO, 2009. The State of Food Insecurity in the World, 2009. Food and Agriculture Organization of the United Nations, Rome. Available at: http://www.fao.org/docrep/012/i0876e/i0876e00.htm

Farm Sanctuary, 2005. *The Welfare of Cattle in Dairy Production.* Available at: http://www.farmsanctuary.org/mediacenter/assets/reports/dairy_report.pdf

Figueira, J., Greco, S., and Ehrgott, M. (Editors) 2005. *Multiple-criteria Decision Analysis: State of the Art Surveys.* Springer International Series in Operations Research and Management Science, New York.

Funtowicz, S.O. and Ravets, J.R. 1990. Post normal science: a new science for new times. *Scientific European* **266**: 20-22.

Funtowicz, S.O. and Ravets, J.R. 1991. A new scientific methodology for global environmental issues. **In:** *Ecological Economics,* pp. 137-152. Costanza, R., Ed., Columbia Press, New York.

Funtowicz, S.O. and Ravets, J.R. 1993. Science for the post-normal age. *Futures* **25**: 735-755.

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

Funtowicz, S.O., Ravets, J.R., and O'Connor, M. 1998. Challenges in the use of science for sustainable development. *International Journal of Sustainable Development* **1**(1): 99-107.

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.

Georgescu-Roegen, N. 1971. *The Entropy Law and the Economic Process.* Harvard University Press, Cambridge, MA.

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

Giampietro, M. 2002. The precautionary principle and ecological hazards of genetically modified organisms. *AMBIO* **31** (6): 466-470.

Giampietro, M. 2003. *Multi-Scale Integrated Analysis of Agro-ecosystems.* CRC Press, Boca Raton, FL, 472 pp.

Giampietro, M. 2008. The future of agriculture: GMOs and the agonizing paradigm of industrial agriculture. **In:** *Science for Policy: Challenges and Opportunities*. Guimaraes Pereira, A. and Funtowicz, S., Eds., Oxford University Press, New Dehli.

Giampietro, M. and Mayumi, K. 2004. Complex systems and energy. In: *Encyclopedia of Energy,* Vol. 1, pp. 617-631. Cleveland, C., Ed., Elsevier, San Diego.

Giampietro, M. and Mayumi, K. 2009. *The Biofuel Delusion: the Fallacy behind Large-scale Agro-biofuel Production.* Earthscan Research Edition, London, 320 pp.

Giampietro, M. and Pastore, G. 1999. Multidimensional reading of the dynamics of rural intensification in China: the AMOEBA approach. *Critical Reviews in Plant Sciences* **18** (3): 299-330.

Giampietro, M. and Pastore, G. 2001. Operationalizing the concept of sustainability in agriculture: characterizing agroecosystems on a multi-criteria, multiple-scale performance space. **In:** *Agroecosystem Sustainability: Developing Practical Strategies*, pp. 177-202. Gliessman, S.R., Ed., CRC Press, Boca Raton.

Giampietro, M., Allen, T.F.H., and Mayumi, K. 2006a. The epistemological predicament associated with purposive quantitative analysis. *Ecological Complexity* **3**(4):307-327.

Giampietro, M., Mayumi, K., and Munda, G. 2006b. Integrated assessment and energy analysis: Quality assurance in multi-criteria analysis of sustainability. *Energy* **31**(1): 59-86.

Gliessman, S.R. 1998. *Agroecology.* Ann Arbor Press, Ann Arbor.

Gomiero, T. and Giampietro, M. 2005. Graphic tools for data representation in integrated analysis of farming systems. *International Journal of Global Environmental Issues* **5** (3/4): 264-301.

Hayashi K. 2000. Multi-criteria analysis for agriculture resource management: a critical survey and future perspectives. *European Journal of Operational Research* **122**: 486-500.

IAASTD (The International Assessment of Agricultural Knowledge, Science, and Technology for Development), 2009. *Agriculture at a Crossroads: Executive Summary of the Synthesis Report.* Island Press, Washington D.C.

Janssen, R. and Munda, G. 1999. Multi-criteria methods for quantitative, qualitative and fuzzy evaluation problems, In: *Handbook of environmental and resource economics,* pp. 837-852. van den Bergh J., Ed., Edward Elgar, Cheltenham, UK.

Joensen, L. 2004. Syngenta emprende campaña publicitaria del Soja RR. *Revista del sur; red del Tercer Mundo* N. 149/150, March-April. http://www.redtercermundo.org.uy/revista_del_sur/texto_completo.php?id=2435

Jordan, C.F. 2001. Genetic engineering, the farm crisis and world hunger. *Bioscience* **52**: 523-529.

Koestler, A. 1968. *The Ghost in the Machine*. MacMillan, New York, 365pp.

Koestler, A. 1978. *Janus: a summing up*. Hutchinson, London.

Krugman, P. 2008. Grains gone wild, *New York Times,* 7 April.

Martinez-Alier, J. 2008. Languages of valuation. *Economic and Political Weekly,* November, 2008: 28-32.

Millennium Ecosystem Assessment 2005. *Synthesis Report – Ecosystems and Human Well-Being*. http://www.maweb.org/documents/document.356.aspx.pdf

Millstone, E., Brunner, E., and White, I. 1994. Plagiarism or protecting public health. *Nature* **371**: 647-648.

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

Munda, G. 2005. Measuring sustainability: a multi-criterion framework. *Environment, Development and Sustainability* **7**(1), 117-134.

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

Munda, G. 2008. *Social Multi-Criteria Evaluation for a Sustainable Economy.* Springer-Verlag, Berlin, 210 pp.

Nijkamp, P. and Ouwersloot, H. 1998. Multidimensional sustainability analysis: the flag model. In: *Theory and implementation of economic models for sustainable development*, pp. 255-277. van den Bergh, C.J.M. and Hofkes, M.W., Eds., Kluwer, the Netherlands.

Nijkamp, P. and Vreeker, R. 2000. Sustainability assessment of development scenarios: Methodology and application to Thailand. *Ecological Economics* **33**:7-27.

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

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

O' Neill, R.V. 1989. Perspectives in hierarchy and scale. **In:** *Perspectives in Ecological Theory*, pp. 140-156. Roughgarden, J., May, R.M. and Levin, S., Eds., Princeton University Press, Princeton, NJ.

O'Neill, R.V., DeAngelis, D.L., Waide, J.B., and Allen, T.F.H. 1986. *A Hierarchical Concept of Ecosystems.* Princeton University Press, Princeton, NJ.

Pastore, G., Giampietro, M., and Li Ji 1999. Conventional and land-time budget analysis of rural villages in Hubei province, China. *Critical Reviews in Plant Sciences* **18**(3): 331-358.

Pattee, H.H. (Editor). 1973. *Hierarchy Theory.* George Braziller, Inc., New York.

Polimeni, J., Mayumi, K., Giampietro, M., and Alcott, B. 2008. *Jevons' Paradox: The Myth of Resource Efficiency Improvements.* Earthscan Research Edition, London, 192 pp.

Poultryhub.org 2009. *Chicken meat (broiler) industry* Available at: http://www.poultryhub.org/index.php/Chicken_meat_(broiler)_industry

Pretty, J., Morrison, J.I.L., and Hine, R.E. 2003. Reducing food poverty by increasing agricultural sustainability in developing countries. *Agriculture, Ecosystems and Environment* **95**: 217-234.

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

Röling, N. 1994. Platforms for decision-making about ecosystems. **In:** *The Future of the Land: Mobilizing and Integrating Knowledge for Land Use Options,* pp. 386-393. Fresco, L.O., Stroosnijder, L., Bouma, J., and van Keulen, H., Eds., John Wiley & Sons Ltd, New York.

Röling, N. 1996. Toward an interactive agricultural science. *European Journal of Agricultural Education and Extension* **2**(4): 35-48.

Röling, N. and Wagemakers, A. (Eds.) 1998. *Facilitating Sustainable Agriculture Participatory Learning and Adaptive Management in Times of Environmental Uncertainty*. Cambridge University Press, Cambridge.

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

Rosen, R. 1986. *Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations*. Pergamon Press, New York.

Rosen, R. 2000. *Essays on Life Itself*. Columbia University Press, New York, 361 pp.

Rosset, P.M. 2006. *Food is different.* Zed Books, New York.

Salthe, S.N. 1985. *Evolving Hierarchical Systems: Their Structure and Representation*. Columbia University Press, New York.

Salthe, S.N. 1993. *Development and Evolution: Complexity and Change in Biology*. The MIT Press, Cambridge, MA.

Simon, H.A. 1962. The architecture of complexity. *Proc. Amer. Philos. Soc*. **106**: 467-482.

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

USDA 2007. *United States 2007 Animal Health Report*, p. 74 – Available at: http://www.aphis.usda.gov/publications/animal_health/content/printable_version/ahr2007.pdf

van Eekeren, N., Maas, A., Saatkamp, H.W., and Verschuur, M. 2006. *Small-Scale Chicken Production* (4th revised edition) Agromisa Foundation and CTA, Wageningen. Available at: http://www.scribd.com/SmallScale-Chicken-Production/d/8422107

Vincke, P. 1992. *Multicriteria Decision Aid*. Wiley, New York.

Waltner-Toews, D. and Lang, T. 2000. The emerging model of links between agriculture, food, health, environment and society. *Global Change & Human Health* **1**(2):116-130.

Wasley, A. 2009. Killing fields: the true cost of Europe's cheap meat. *Ecologist (Film Unit)*

http://www.theecologist.co.uk/trial_investigations/336873/killing_fields_the_true_cost_of_europes_cheap_meat.html
Wittgenstein, L. 1974. *Tractatus Logico-Philosophicus.* Routledge, London.
Zeleny, M. 1982. *Multiple-Criteria Decision-Making.* McGraw Hill, New York.

FIGURE 1

Four metaphors illustrating the epistemological problems generated by the compression of information used to represent a complex perception

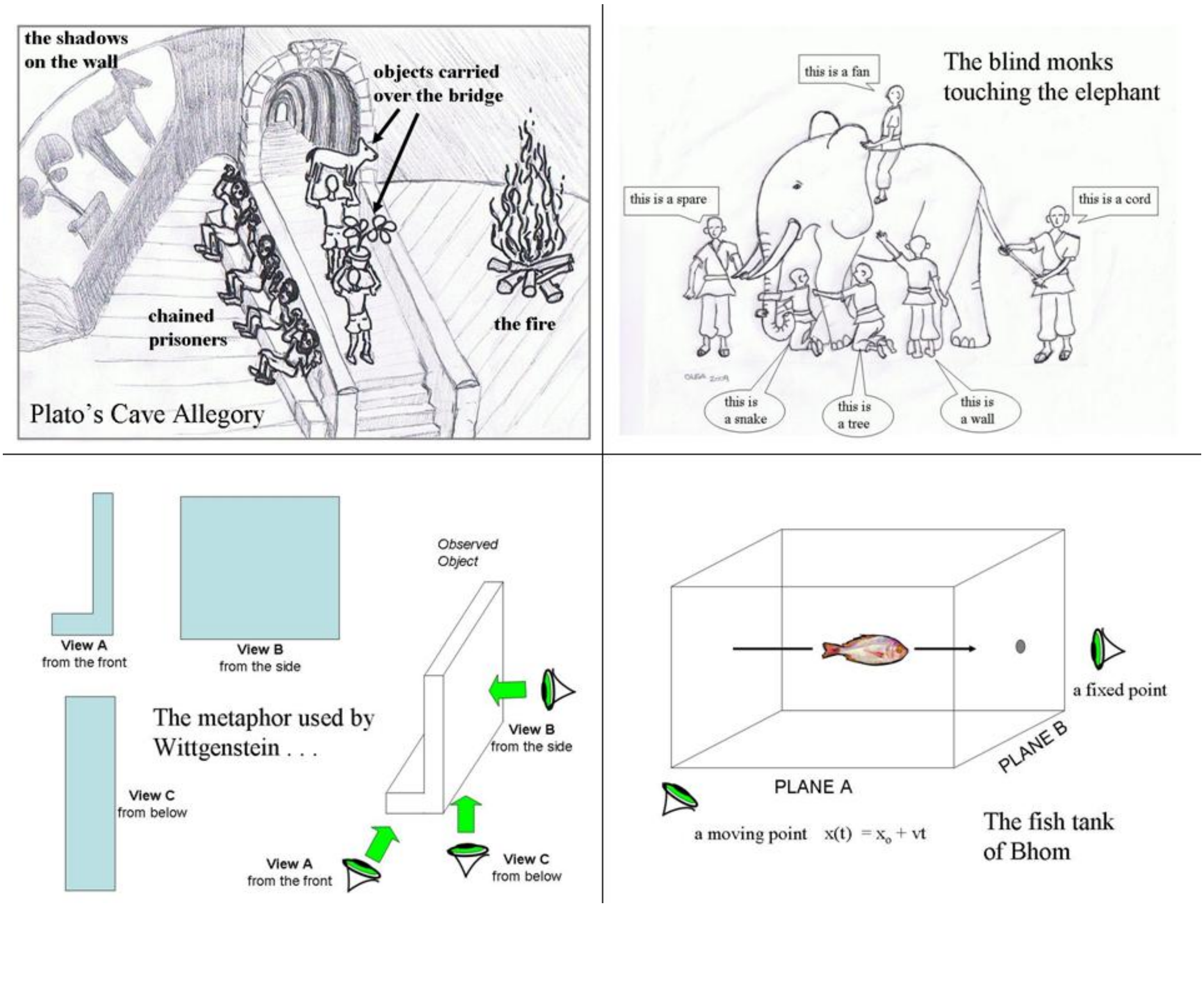
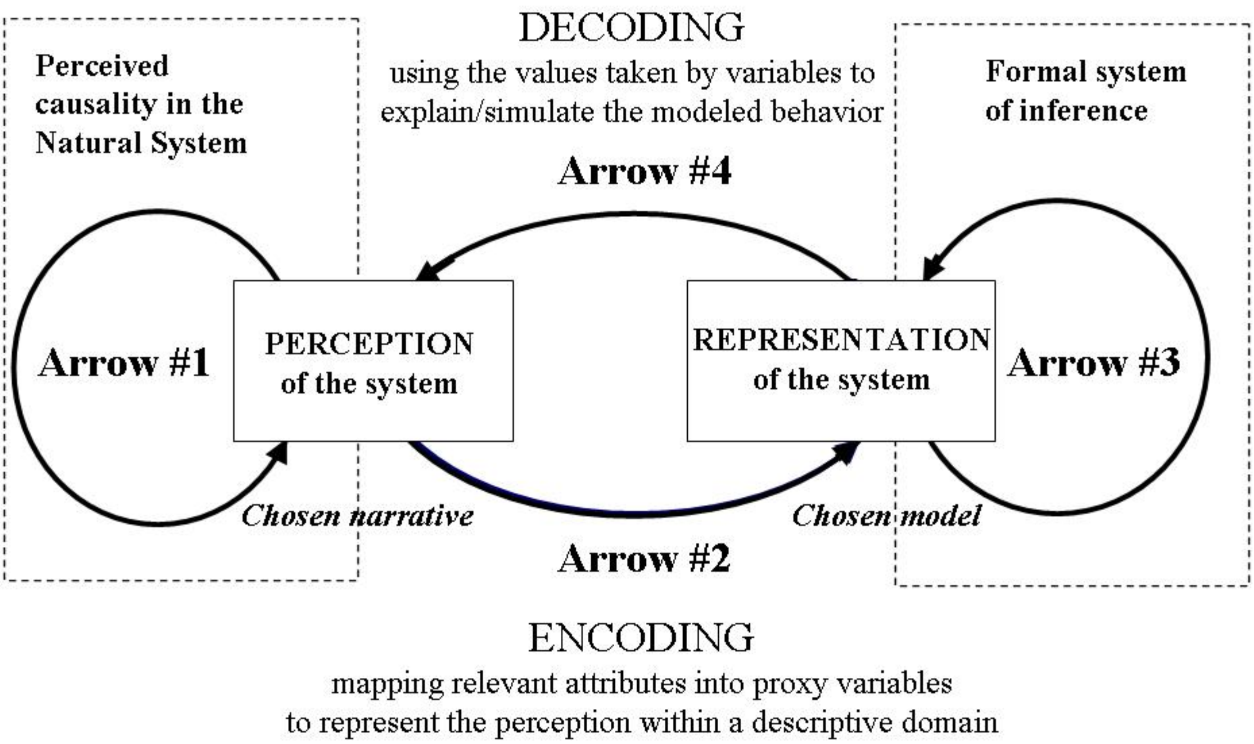


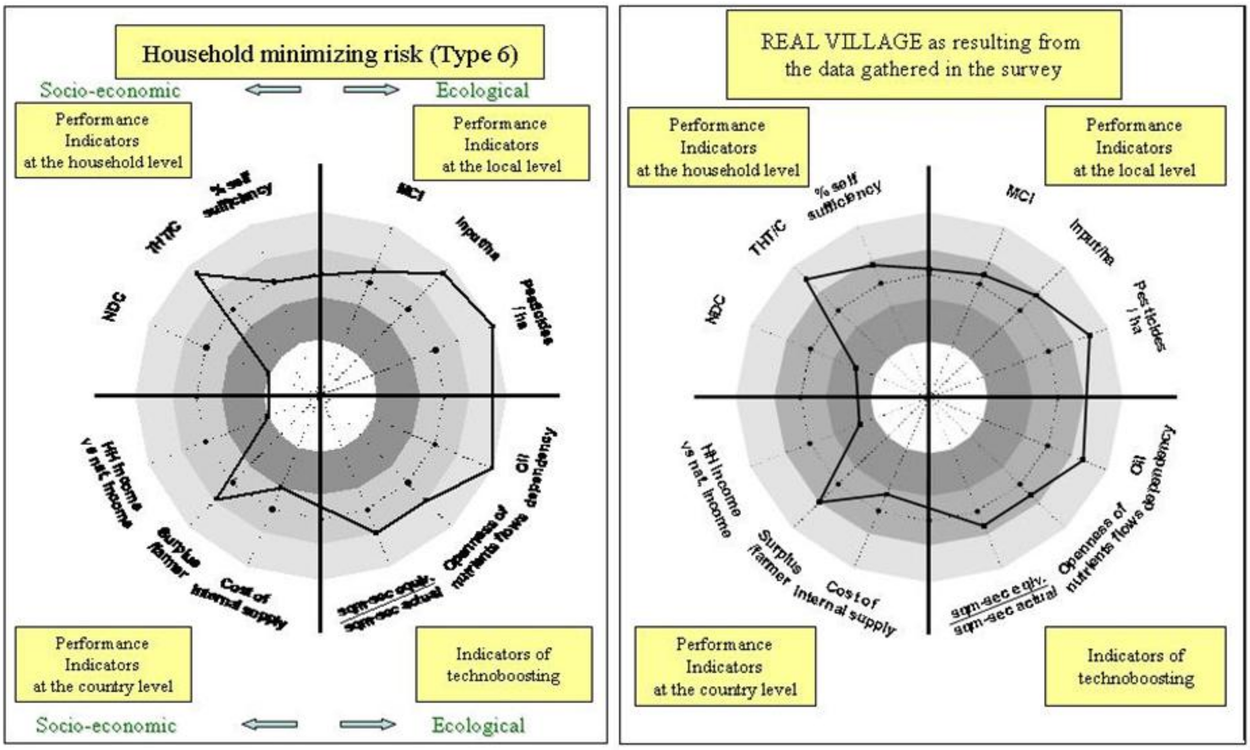
FIGURE 2
An illustration of the scheme of modeling relation according to Rosen



icta Institut de Ciència i Tecnologia Ambientals • UAB

UAB Universitat Autònoma de Barcelona

FIGURE 3
Examples of multicriteria characterization of performance

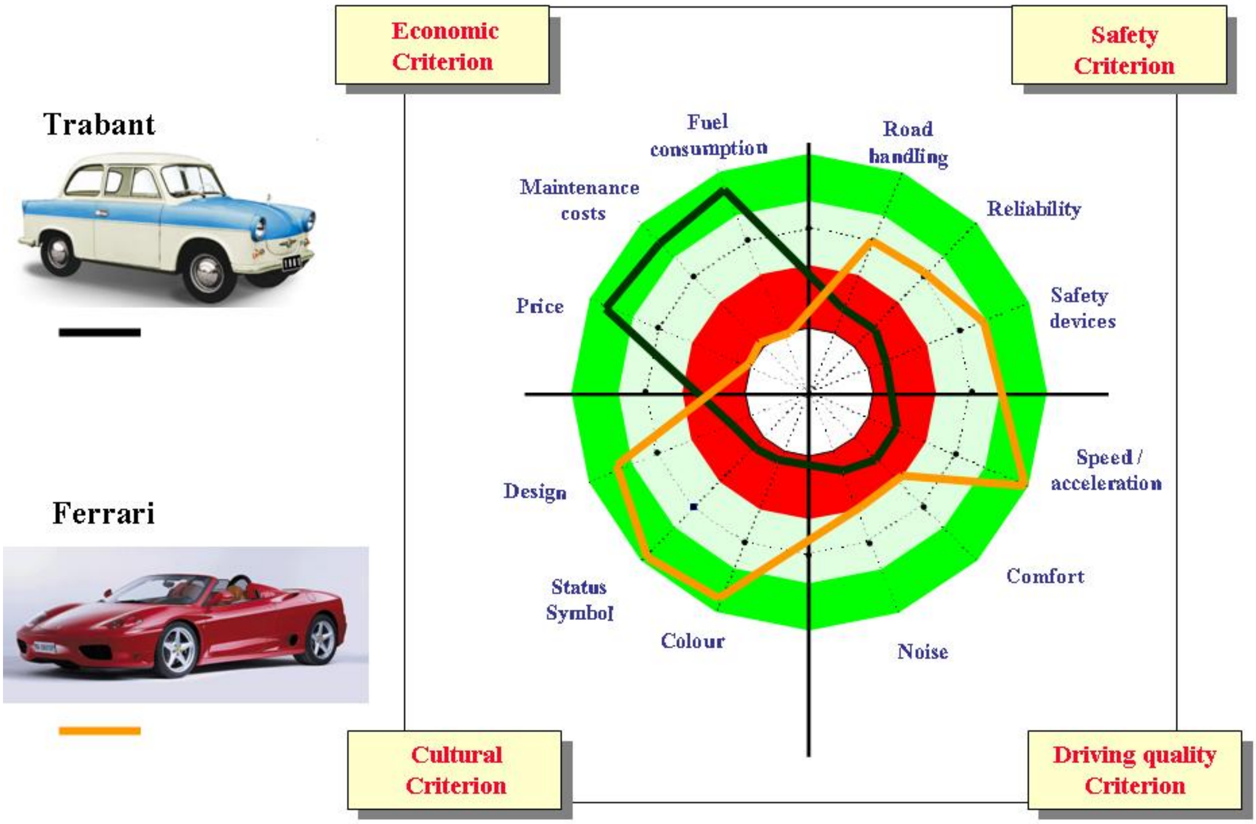


icta Institut de Ciència i Tecnologia Ambientals • UAB

UAB Universitat Autònoma de Barcelona

FIGURE 4

Examples of multicriteria characterization for two systems having a very different definition of performance



Institut de Ciència i Tecnologia Ambientals • UAB

Universitat Autònoma de Barcelona

FIGURE 5
The problem represented by the existence of different story-tellers generating non-equivalent definitions of relevance