The state of the art in indicators construction in the perspective of a comprehensive approach in measuring well-being of societies.

Università degli Studi di Firenze

WORKING PAPER

Filomena Maggino

The state of the art in indicators construction

in the perspective of a comprehensive approach in measuring well-being of societies

Measuring well-being of societies requires a multidimensional and integrated approach aimed at describing the health of societies through a complex multifaceted and compound methodology.

The aim of this work is to unravel some important methodological aspects and issues that should be considered in measuring societal well-being in quantitative perspective and in developing indicators of well-being.

The purpose is to focus on, to examine closely and to investigate the conceptual issues in defining and developing indicators and the methodologies and operative issues in managing the complexity of the obtained observation, integrating objective and subjective aspects of the reality.

This work, or part of it, has been presented in different occasions, among which:

- II Conference of the European Survey Research Association (Prague, Czech Republic, 25-29 June 2007) – paper presented at the session "Models of Analysis for Defining and Constructing Subjective Data and Indicators" (session organizer: F.Maggino).
- XLIV scientific meeting of the Società Italiana di Statistica (Cosenza, Italy, 25-27 June 2008) – invited paper at the specialized session on "Construction of well-being indicators using objective and subjective information" – organized by Stefano Tarantola (European Commission / Joint Research Centre), discussant: M. Attanasio.
- Training Course "Statistics, Knowledge and Policy: Understanding Societal Change" organized by OECD/OCSE (Global Project) (Siena, Italy, September 6-19, 2008) – lecture.
- Scientific Conference ""Birth encouragement: multi-sided practice of Government intervention in population reproduction – the Bulgarian and European experience" (Sofia, Bulgaria, 13-14 March 2009) – invited paper.
- Workshop "L'accreditamento degli Atenei. Metodi e modelli per definire reputazione e posizionamento", organized by Università degli Studi di Milano Bicocca e Società Italiana di Statistica (Milan, Italy, 27-28 March, 2009) – lecture.
- III Conference of the European Survey Research Association (Warsaw, Poland, 29 June – 3 July, 2009) – paper presented at the session "Social Indicators – Integration of objective and subjective indicators: methodological and technical issues" (session chair and organizer F. Maggino).
- Training Course "Statistics, Knowledge and Policy: Understanding Societal Change" organized by OECD/OCSE (Global Project) – Joint Research Centre (European Commission) – ISQOLS (Florence, Italy, 14-17, 2009) – lecture.
- IX Conference of the International Society of Quality of Life Studies "Quality of Life Studies: measures and goals for the Progress of Societies" (Florence, Italy, 19– 23 July, 2009) – paper presented at the session "Objective and subjective approaches in measuring differences in quality of life" (session chairs and organizers K. Land and F. Maggino).
- Koinè Statistica Project "Methodology aimed at constructing and assessing indicators" organized by University of Padua and Region of Veneto (Venice, Italy, September 29, 2009) – lecture.
- Conference "From GDP to Well-Being: Economics on the Road to Sustainability" organized by Università Politecnica delle Marche (Ancona, Italy, December 3-5, 2009).

Table of Contents

Introduction

Any measuring process needs basic principles to be clarified in order to proceed with. Three different measuring processes can be identified.

Measuring by fundamental process: measuring process does not refer to previous measures (operative process) but do reflect natural laws (constitutive process). Assessing a fundamental process requires a theory to be constructed and inspected. Characteristics that can be measured through a fundamental process are length, volume, and so on.

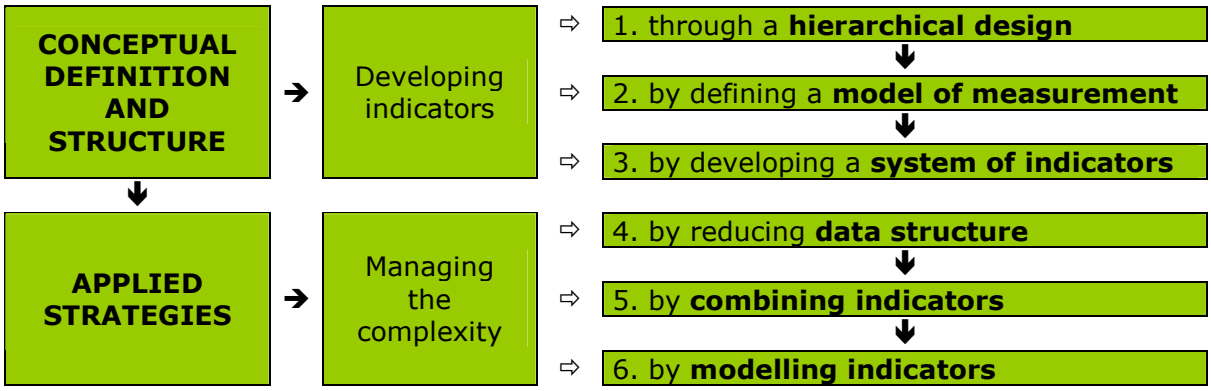
Measuring by deriving process: measuring process is based upon other measures, related to each other through a wider theory allowing algorithms to be defined and applied on fundamental measures; characteristics that can be measured through a fundamental process are density (relationship between mass and volume), velocity (relationship between space and time).

Measuring by defining process: measuring process is carried out in consequence of (and consistently with) a definition confirmed through relationships recorded between observations and defined concepts. All the measures applied in social sciences belong to this category (*socio-economic status*, *capacities*, etc.). This measuring process requires indicators to be defined.
With reference to this, we have to point out that even though "**indicator**" and "**index**" are terms often used in an interchangeable way, they have different origins and meanings: the former comes from the late Latin word "indicator" that means "who or what indicates" and the latter comes from the Latin word "index", which means "any thing that is useful to indicate".
In statistics, on one hand, "index" represents historically a very generic word applied with multiple meanings; on the other, "indicator" represents a more recent term indicating – as seen above – indirect measures of economic or social phenomena not directly measurable.
In this perspective, an indicator is not a simple crude statistical information but represents a measure organically connected to a conceptual model aimed at knowing different aspects of reality. In other words, a generic index value can be converted into an "indicator", when its definition and measurement occur in the ambit of a conceptual model and is connected to a defined aim (indicators as *purposeful statistics*. Horn, 1993).
In particular, a statistical index can be considered as a [social] indicator when (Land, 1971, 1975):

- it represents a component in a model concerning a social system
- it can be measured and analysed in order to compare the situations of different groups and to observe the direction (positive or negative) of the evolution along time (time series analysis)
- it can be aggregated with other indicators or disaggregated in order to specify the model.

The lack of any logical cohesion should not to be hidden by the use and application of sophisticated procedures and methods that can deform reality through distorted results.

In order to develop and manage indicators able to represent different aspects of the reality, to tell meaningful stories and to build interpretable pictures of the reality (in our perspective, well-being of societies), a structured plan should be followed:

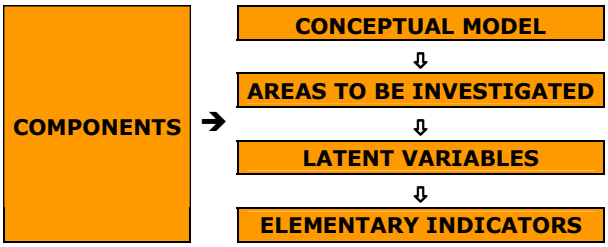


CONCEPTUAL DEFINITION AND STRUCTURE

How conceptually design the picture?

1. HIERARCHICAL DESIGN

Indicators should be developed through a **logical modelling process** conducting from concept to measurement. Given its features, this logical design is defined *hierarchical*, since each component is defined and finds its meaning in the ambit of the preceding one. Conceptually, the hierarchical design is characterized by the following components: (i) the conceptual model, (ii) the areas to be investigated, (iii) the latent variables, and (iv) the elementary (basic) indicators. The hierarchical design is schematically summarized in the following picture:



How conceptually define the indicators?

2. MODEL OF MEASUREMENT

A further component of the hierarchical design definition is represented by the relationships between:

- *Latent variables and the corresponding indicators*: these relations define the **model of measurement**, which will be discussed below. Consistently with the measurement model, also the relationship between the *elementary indicators* should be defined In this perspective, two different states can be identified:
    - indicators are related to each other and relate to the same latent variable (in other words, they contribute to the definition of same variable); in these cases, the indicators are called *constitutive*;
    - indicators are not related to each other and relate to different latent variables; in this case, the indicators are called *concomitant*.
- *Latent variables* for a given area: these relations are defined in the ambit of the conceptual model and identify the structural pattern (**modelling indicators**).

How conceptually organize the picture?

3. SYSTEM OF INDICATORS

A **system of indicators** represents the objectification of the conceptual framework and allows

(i) an effective organizational context to be used, relying on methodological supports and allowing data to be managed;

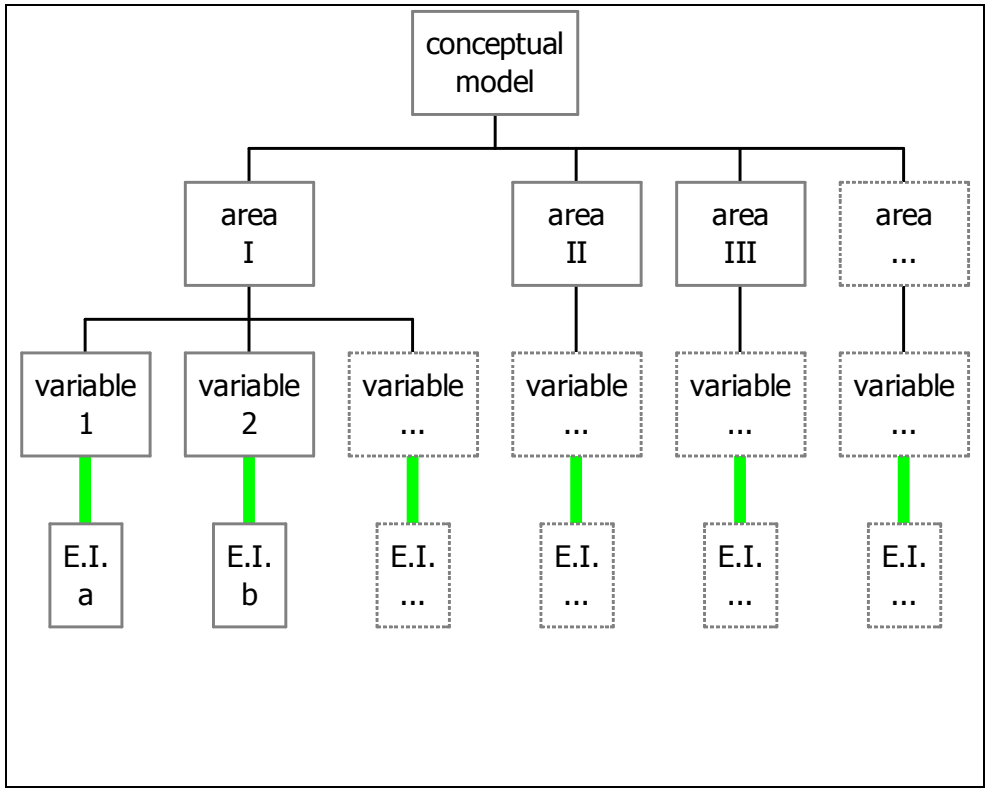
(ii) structured and systematic data to be used, observed in long-term longitudinal perspective[1]. This is particularly demanding with reference to subjective data, which require a great use of resources (beyond a solid survey research methodology).

[1] In fact, if the purpose were to study the phenomenon in predictive perspective, any observed data would need to be collected over sufficiently long periods to successfully capture or model the quality of life and develop an effective knowledge base.

The following table synthesises the conceptual definition and structure:

COMPONENT	QUESTION TO BE ANSWERED	COMPONENT'S DEFINITION
CONCEPTUAL MODEL	Phenomena to be studied	The conceptual model defines the **phenomenon** to be studied, the domains and the general aspects characterizing the phenomenon
AREAS/PILLARS TO BE INVESTIGATED	Aspects defining the phenomenon	Each area represents each **aspect** allowing the phenomenon to be specified consistently with the conceptual model
LATENT VARIABLES	Elements to be observed	Each variable represents each **element** that has to be observed in order to define the corresponding area. The variable is named **latent** since is not observable directly
ELEMENTARY INDICATORS (E.I.)	In which way each element has to be measured	Each indicator (item, in subjective measurement) represents **what is actually measured** in order to investigate each variable. It is defined by appropriate techniques and by a system that allows observed value to be evaluated and interpreted.

GRAPHIC REPRESENTATION



APPLIED STRATEGIES

In order to manage the complexity of the obtained system, different applied strategies can be adopted. The strategies constitute a "composite" **process**, carried out through subsequent/consecutive steps (MULTI-STAGE) and different/alternative analytical approaches (MULTI-TECHNIQUE).

How to simplify the observed picture?

4. REDUCING THE COMPLEXITY OF DATA STRUCTURE

The consistent application of the hierarchical design produces a complex data structure (elementary indicators, cases, variables, areas, etc.). In order to manage the complexity, some dimensions may require a particular treatment, consistently with the conceptual model:

(i) **aggregating elementary indicators** identified for each variable (except those measured by single indicators): the aggregating process aims at re-constructing the conceptual variables consistently with the approach (reflective or formative) adopted at micro level (*construction of synthetic indicators)*

(ii) **aggregating units/cases**: the aggregating process aims at leading information observed at micro-level to the proper and identified macro level of interest (*definition of macro-units*). Identifying the proper aggregation criterion should take into account the nature of measured characteristics (e.g. compositional, contextual, and so on) requiring different analytical approaches.

How to get the whole picture?

5. COMBINING INDICATORS

In some occasion, the complexity of the system of indicators may require the indicators allowing for more comprehensive measurement. This need can emerge in order to (Noll, 2009)

- answer the call by 'policy makers' for condensed information
- improve the chance to get into the media (compared to complex indicator systems)
- allow to make multi-dimensional phenomena uni-dimensional
- allow to compare situations across time more easily
- compare cases (e.g. nations) in a transitive way (ranking)
- allows clear cut answers to questions like the following:
    a. are living conditions getting better or worse across time?
    b. do people living in City A enjoy a better quality of life than those living in City B?
    c. is population subgroup X better off than population subgroup Y?

Dashboards or **composite indicators** can represent useful approaches aimed at summarising indicators.

How to explain the picture?

6. MODELLING INDICATORS

This stage is aimed at analysing different aspects of the defined model (e.g. objective and subjective indicators) in order to find explanation by identifying the proper analytical approaches**.**

The following table synthesises the applied strategies:

Goals	Level of analysis	Stages	Aims	by	Analytical issues
Reducing data structure:	micro	(a-i)	**construction of synthetic indicators**	aggregating elementary indicators	From elementary indicators to synthetic indicators - reflective approach - formative approach
		(a-ii)	**definition of macro-units**	aggregating observed units	From micro units to macro units, by following - homogeneity criterion - functionality criterion
Combining indicators:	macro	(b-i)	**definition of dashboards**	jointly representing indicators	Comparing over time / across units
		(b-ii)	**construction of composite indicators**	merging indicators	Aggregating information very different from each other (e.g. objective and subjective)
Modelling indicators:	macro	(c)	**analysis of indicators**	exploring explanations	Different solutions (consistently with conceptual framework)

This work is aimed at showing the methodological and applied issues related to this structured plan.

1. Defining the hierarchical design

Indicators should be developed, by following the Lazarsfeld's model (1958), through a *hierarchical design* requiring the definition of the following components:

Conceptual model
The definition of the conceptual model represents a process of abstraction, a complex stage that requires the identification and definition of theoretical constructs that have to be given concrete references of applicability. In social sciences, the description of concepts varies according to (i) the researcher's point of view, (ii) the objectives of the study, (iii) the applicability of the concepts, (iv) the socio-cultural, geographical, historical context. Concerning this, we can refer to concepts like health, education, well-being, income, production, trade, etc.
The process of conceptualisation allows us to identify and define:

a. the model aimed at data construction,
b. the spatial and temporal ambit of observation,
c. the aggregation levels (among indicators and/or among observation units),
d. the approach aimed at aggregating the elementary indicators and the techniques to be applied in this perspective (weighting criteria, aggregation techniques, etc.),
e. the interpreting and evaluating models.

Areas
The areas (in some cases named "pillars") define in general terms the different aspects that allow the phenomenon to be clarified and specified consistently with the conceptual model. The process of defining areas can be long and exacting, especially with complex constructs, and requires an analysis of the literacy review.

Latent variables
Each variable represents one of the aspects to be observed and confers an explanatory relevance onto the corresponding defined area. The identification of the latent variable is founded on theoretical assumptions (e.g. homogeneity, dimensionality) and empirical statements so that the defined variable can reflect the nature of the considered phenomenon consistently with the conceptual model.
However, even if we are able to identify a variety of diverse variables, we have to accept the idea that maybe no set of variables can perfectly capture the concept to be measured (e.g. social or economic well-being; Sharpe & Salzman, 2004).

Elementary indicators
Each elementary indicator (item, in subjective measurement) represents what can be actually measured in order to investigate the corresponding variable.[1] This means that each observed element represents not a direct measure of the variable but an **indicator**[2] of the reference variable (DeVellis, 1991). The hierarchical process allows a meaningful and precise position to be attributed to each indicator inside the model. In other words, each indicator takes on and gains its own meaning, and consequently can be properly interpreted because of its position inside the hierarchical structure: each indicator represents a distinct component of the phenomenon within the hierarchical design. The possibility to define and to consider alternative forms for each indicator has to be evaluated.
According to a simple and weak strategy, each latent variable is defined by a single element (**single indicator approach**). This strategy, applied because of its thrifty and functional capacity, requires the adoption of robust assumptions. The adoption of single indicators presents a risk since it is not always possible to define the direct correspondence between one latent variable and one indicator. In other words, the variable is not always directly observable through a single indicator. In fact, defining and adopting the single indicator approach can produce a wide and considerable amount of error that leads to problems concerning:

a. *precision* (*reliability*), since the measurement through one single indicator is strongly affected by random

[1] In specific cases, some variables can be directly measured (e.g. some objective information). In this cases, variable and indicator coincide.
[2] In data analysis, indicators/items are technically defined "variables"; consequently, these are conceptually different from "latent variables".

error;[3]

b. *accuracy* (*validity*), since the chance that one single indicator can describe one latent complex variable is highly dubious and questionable;
c. *relationship* with the other variables;
d. *discriminating* and *differentiating* among observed cases.

That is why, in many cases, the presence of complex latent variables requires the definition of several elementary indicators. This can be done by adopting the **multiple indicators approach**, which considers the multiple indicators as *multiple measures* (Sullivan & Feldman, 1981). Multiple indicators contribute to the measurement of the major aspects of the variable since each elementary indicator corresponds to one particular aspect of the latent variable. This approach allows variability of the defined latent variable to be covered. In addiction, this approach allows the problems produced by the single indicators approach to be avoided, or at least for their significance and weight to be reduced. In technical terms, the complete group of elementary indicators referring to one variable represents a *set of indicators*, while the complete group of indicators defining an area are called *thematic indicators*.

The hierarchical design can be drawn also through sub-designs (e.g. each area could require sub-areas) and its logic can be applied both at micro and macro level.

1.1. Different conceptual frameworks of well-being

"Well-being" is a term largely used and expressing a concept not always clearly defined. Many theoretic models have been developed and try to explain and to operationalise different definitions and concepts. The conceptual frameworks could be distinguished through different criteria.

(A) structure of values

The distinction among all the different definitions can be explained by the different structures of life values adopted. According to Diener and Suh (1997), they can be referred mainly to three philosophical approaches. The different conceptual frameworks and observation strategies are synthesised in the following table, drawing a simplified and reduced picture of the different and numerous concepts defined in order to define and measure societal well-being.

Societal well-being				
Societal well-being is related to	What should be observed	Strategies of observation		measures
		through what	at which level/s	
Functioning and capability to select goods and services that one desires	Income, considered as a mean to achieve an acceptable standard of living	observed or estimated amounts	Macro (national) and/or micro (individual)	economic indices and national accounts
Normative ideals	a set of characteristics inspired by normative aims, grounded in moral values or policy goals	living conditions	macro (e.g. national) and micro (individual)	social indicators
Subjective experiences	Individual's cognitive and affective reactions to his/her whole life (or specific domains) and societies	subjective reactions and perceptions	micro (individual)	subjective indicators

(B) Different perspective of observation

The different concepts that can be used in order to define the well-being of societies can be distinguished with reference to different perspectives, having reference mainly to **processes**, **conditions**, or **goals**.

Seen in terms of **process**, societal well-being finds the concept of "development" (often referring to qualitative dynamic change of an economic system) and of "growth" (referring to quantitative expansion on the scale of physical dimensions of economic system). Both concepts refer to different but interactive components and characteristics (economic, structural and technologic) that should be considered

[3] By using multiple measures, random errors tend to compensate each other. Consequently, the measurement turns out to be more accurate. The greater the error component in one single measure, the larger the number of required measures needs to be.

together (Horn, 1993). A term that can unify those presented above is *progress*, indicating generally "moving forward" (from Latin "*progressus", going forward, advance*). However, as limits or potentialities of the process have been reached, the attention could be turned towards the reverse and opposite process of "de-development" (Horn, 1993).
Seen in terms of **conditions**, societal well-being encounters the concept of
- availability of economic resources (*manpower, equipment, budget*),
- social implications of distribution of income and wealth,
- impacts of economics on national welfare and environment.

This perspective requires testing the improvement by which Individuals identify themselves as a community and acquire collectively the necessary knowledge, power, values and organizational skills to irreversibly share and expand the community's resources for the benefit of all its members without being at the expense of other communities or of the environment (Horn, 1993). In other terms, the conditions should be sustainable.
This perspective moves the attention from the process (development, progress, growth) to the **goal**, which could be represented by sustainability, quality of life, well-being, and so on.

According to Veenhoven's approach (2009), well-being can be defined through two different dimensions, one referring to the chances/outcomes of a system, and the other referring to external or internal of a system. Combination of the two dimensions yields four kinds of well-being:

		Dimension II	
		External	Internal
Dimension I	Chances	Favourable environment (a)	Good functioning (b)
	Outcomes	Positive external effects (c)	Continuance (d)

The patterns can be applied to different systems. In biological organisms, the well-being chances are denoted by (a) *biotope* and (b) *fitness*, while the well-being outcomes are denoted by (c) *ecological functions* and (d) *survival* of organisms/species.
In business organizations, the well-being chances are denoted by (a) *favourable market* and (b) *capital*, while the well-being outcomes are denoted by (c) *public wealth* and (d) *private profit* of firms.
Seen in terms of
- individual well-being, the well-being chances are denoted by (a) *livability* of environment (→ social capital) and (b) *life-ability* (personal capacities → psychological capital), while the well-being outcomes are denoted by (c) *utility of life* (→ meaning of life) and (d) *long and happy life*;
- societal well-being, the well-being chances are denoted by (a) *ecological condition/geo-political position* of environment and (b) *functioning of institutions* (→ social organization), while the well-being outcomes are denoted by (c) *burden to eco-system / contribution to civilization* (nation's impact of the eco-system / innovations in human civilization) and (d) *continuity / morale*.

(C) Different viewpoints

Berger-Schmitt and Noll (2000) well systematized the different conceptual frameworks that can be identified by distinguishing mainly between conceptual frameworks centred on respectively individuals and societies. This distinction allowed them to classify the different conceptual frameworks. The following table synthesises the classification:

Societal well-being concepts	Approaches	Definitions	Main references and inspirations
Quality of life (conceptualised implicitly or explicitly at individual level)	Resources approach (Scandinavian level of living → objective needs)	- Individual's command over resources through which the individual can control and consciously direct his/her living. - External circumstances (arenas) are essential components. - Subjective well-being is not an appropriate criterion for guiding policy.	Eriksson
	Capabilities approach	- Individual capabilities to achieve the high valued functionings and individual ability to live the type of life aspired. - It takes into consideration differences in individual values and preferences.	Sen
	Subjective well-being approach (American quality-of-life concept)	Needs satisfaction assessed by individuals themselves. The **goal is people's subjective well-being**, seen in terms of - satisfaction (cognitive judgment) - happiness (affective dimension). Many authors distinguish between five dimensions of quality of life: physical well-being, material well-being, emotional well-being, cognitive development, productive activities.	Campbell / Converse / Rodgers Andrews / Withey Argyle Diener / Suh
	Basic needs approach	Considered **dimensions** are - **having** (satisfaction of material and impersonal needs) - **loving** (need for social relationships) - **being** (need for integration into society and for harmonious relation to nature) Each category is defined by both objective and subjective dimensions. [4]	Allardt Galtung
	Objective living conditions and subjective well-being approach (German quality-of-life)	**Combination** of - objective living conditions (living circumstances) - subjective well-being (cognitive and affective dimensions)	Zapf

Quality of societies	Liveability and quality of nations		It represents the degree to which its provisions and requirements fit the needs and capacities of its citizens. **Approaches to measurement**: - input → living conditions (matching citizens' needs and capacities) - output → degree to which people flourish in a society (physical and mental health, satisfaction and happiness)	Veenhoven
	Societal integration, solidarity and stability	Social cohesion	Relations among members of society are seen in terms of - shared values, - communities interpretation, - feelings of common identity, - a sense of belonging to community, - trust among members, - reduction of disparities.	Jenson OECD Council of Europe European Commission Club of Rome
			Diifferent approaches to measurement: (i) belonging, (ii) inclusion/exclusion, (iii) participation, (iv) recognition/rejection, (v) legitimacy/illegitimacy	

[4] Different lists of human needs have been proposed by many researchers. Concerning this, we can refer to the Max-Neef's (1992) "Matrix of Human Needs" and the Nussbaum and Glover's (1995) "Basic Human Functional Capabilities". Other research into basic needs including Frisch's (1998) "Quality of Life Inventory", Cummins' (1993) "The ComQuality of life-A5", Maslow's (1954) "Hierarchy of needs," "Need Hierarchy Measure of Life Satisfaction" of Sirgy et al. (1995) and "Quality of Life Questionnaire" of Greenley et al. (1997). However, the exhaustive bibliographic review of all the proposals is not the aim of this work.

			(i) ties (values, identity, culture), (ii) differences (inequalities cultural diversity, geographical divisions), (iii) social glue (associations, networks, infrastructures, values, identity)	
			(i) absence of social exclusion, (ii) interactions and connections based upon social capital, (iii) shared values and group identity.	
		Social exclusion	Analytical concept referring to the **processes** and **causes**. ↓ **Poverty**: concept describing a state or a consequences of social exclusion. It is related to individuals and households. Need to differentiate between - causes of disadvantageous living circumstances - its consequences.	Silver Rodgers Gore Figueiredo De Haan
			Concept related to older concept like disadvantageous social conditions social conditions (poverty, deprivation, unemployment, instability of family, international migration, shortage of welfare benefits). **Different perspectives**:	
			Marginalization lack of solidarity, rupture of relationship between the individual and the society, failure of institutions to integrate individuals into the society. (French approach) Institutional dimensions:	
			(i) democratic and legal system (iii) welfare state system (ii) labour market (iv) family and community system	
			Specialization social differentiation and specialization of individuals' diversity of interests and capabilities. It is caused by individuals' behaviour (Anglo-Saxon approach)	
			Monopoly processes of social closure by which privileged groups protect their monopoly position. The society is characterized by a hierarchy of inclusions and exclusions.	
		Social capital	Different approaches share the concept of SC as a property of a social entity and not of an individual. It is a relational concept and exists only if it is shared by several individuals (public good). **Different concepts**: - Horizontal associations: it is networks of civic engagement, a features of social organization (networks, norms, trust) and facilitate coordination and cooperation for mutual benefit → interpersonal relationships (family, friends, neighbours) - Horizontal and vertical associations: it is a variety of different entities consisting of some aspect of social structure and facilitating actions of actors within the structure → intermediary associations and organizations (clubs, political parties) - Horizontal and vertical associations, formalized relations and macro structures → macro-level of societal institutions	Narayan Putnam Coleman Rossing Feldman/Assaf

	Sustainability	Development meets the needs of the present without compromising the ability of the future generation to meet their own needs. Equal opportunities, equity and solidarity within and between generations. **Dimensions**: - economic → economic growth does not deteriorate natural resources and social conditions of living through a. efficient use of resources b. maintaining people's capacities and supporting social structures providing basic social services (health care, education, equal opportunities) - social → improvement of social conditions enabling people (present and future generations) to pursue their well-being (equal opportunities, equitable wealth distribution) by avoiding environmental damage (equal opportunities and equity of living conditions) - environmental → conservation of natural foundations of life through environmental protection, preservation of biodiversity, limitation of pollution, management of renewable and non-renewable resources. Relations between the three dimensions can be described as a **hierarchy of dependence** (economy is part and depends on society that exists with environment and depends on it) or as a **mutual interdependency**.	OECD, Wiman Becker Hart World Bank Pearce
		OECD's Pressure-State-Response model: (i) *pressure* of human activities on environment, (ii) *state* of the environment, (iii) individual and collective *response* to environmental change. Disadvantages → causal relationships between the three dimensions. Model adopted and implemented by other international institutions and organizations.	
		World Bank's four capital model: (i) natural capital (land, water, wood, minerals, flora, fauna), (ii) produced capital (machineries, factories, buildings, infrastructures), (iii) human capital (people's productive capacities: skills, education, health), (iv) social capital (social networks, associations, institutions with common norms and relationships facilitating co-operation). Weakness → relationship and mutual substitutions between the four capitals.	
	Human development	HU concerns two aspects: (a) formation of human capabilities, (b) way people use their capabilities. Process of enlarging people's choices determined by three **factors**: (i) health, (ii) education, (iii) access to resources. **Components to be measured**: - human freedom - sustainability (equality of opportunities for all people, intergenerational equity) - empowerment (qualifying people for participation) - human security (free and safe choice) - economic growth (mean for human development)	Miles United Nations Development Programme Sen

<table>
<tr><td rowspan="6"></td><td rowspan="6">Social quality</td><td colspan="3">It is the extent to which citizens are able to participate in the social and economic life of their communities under conditions, which enhance their well-being and individual potential.

Components:
- degree of socio-economic security
- extent of social inclusion
- strength of social cohesion and solidarity between and among generations
- level of autonomy and empowerment of citizens</td><td rowspan="6">Beck / van der Maesen / Walker</td></tr>
<tr><td colspan="2" rowspan="2">Model characterized by two **dimensions**:</td><td colspan="2">(B)</td></tr>
<tr><td>institutions / organizations</td><td>Communities /groups/citizens</td></tr>
<tr><td rowspan="2">(A)</td><td>macro-level (social structures)</td><td>socio- economic security</td><td>social cohesion</td></tr>
<tr><td>micro-level (individual)</td><td>social inclusion</td><td>empowerment</td></tr>
<tr><td colspan="4">This original model was criticized and then submitted to different modifications.</td></tr>
</table>

Source of this frame: Berger-Schmitt R. and H.-H. Noll (2000) *Conceptual Framework and Structure of a European System of Social Indicators*, EuReporting Working Paper No. 9, Centre for Survey Research and Methodology (ZUMA) – Social Indicators Department, Mannheim.

1.2 Towards a comprehensive conceptual framework

Each conceptual framework shows strengths and weaknesses, adopts concepts and/or information which can be partially or completely coinciding or overlapping the ones adopted by the others. Consequently, in order to measure societal well-being it is difficult to adopt just one solution and a multidimensional definition and a comprehensive approach need to be assessed.
A possible conceptual framework could be the following: a good and healthy society is that in which each individual has the possibility to participate to the community life, to develop capabilities and independency, to have adequate possibility to choose and control his/her own life, and to be treated with respect in a healthy and safe environment and by respecting the opportunities of future generations.
This is lined up not only with new methodological perspective in measuring the progress but also with a different policy view that looks at the progress in terms of good life. This is not "just a life in which people feel good, no matter how terrible their real life conditions are, but one in which they feel good with the best of all reasons, because the objectively measurable conditions of their lives merit a positive assessment" (Michalos, 2008).
In other words, a **comprehensive approach** is needed allowing objective living information – with reference to micro-individual level and macro-societal level – and subjective well-being to be integrated.

This need is confirmed also by the recent (September 2009) *Report* by the Commission on the Measurement of Economic Performance and Social Progress, chaired and coordinated by J.E. Stiglitz J. E., A. Sen & J.-P. Fitoussi, outlined a similar conceptual framework in order to measure progress of societies. In fact, three wide areas have been identified: (i) classical economical issues, (ii) quality of life and (iii) sustainable development and environment.

Classical GDP Issues		
Existing measurement framework	Improving measurement of	National accounts aggregates
		Services
		government-provided services
	Revisit the concept of "defensive" expenditures	
	Income, wealth and consumption have to be considered together	
Bringing out the household perspective	Adjusting household income measures for government services in kind	
	Medians and means vs. distribution of income, consumption and wealth	
	Broader measures of household economic activity	
	Distribution of full income	
Recommendations:	*1. Look at income and consumption rather than production.* *2. Consider income and consumption jointly with wealth.* *3. Emphasise the household perspective.* *4. Give more prominence to the distribution of income, consumption and wealth.* *5. Broaden income measures to non-market activities.*	
Quality of Life		
Subjective measures of quality of life		
Objective features shaping quality of life	Health	Social connections
	Education	Environmental conditions
	Personal activities	Personal insecurity
	Political voice and governance	Economic insecurity
Recommendations:	*1. Measures of subjective well-being provide key information about people's quality of life. Statistical offices should incorporate questions to capture people's life evaluations, hedonic experiences and priorities in their own surveys.* *2. Quality of life also depends on people's objective conditions and opportunities. Steps should be taken to improve measures of people's health, education, personal activities, political voice, social connections, environmental conditions and insecurity.* *3. Quality-of-life indicators in all the dimensions they cover should assess inequalities in a comprehensive way.* *4. Surveys should be designed to assess the links between various quality-of-life domains for each person, and this information should be used when designing policies in various fields.* *5. Statistical offices should provide the information needed to aggregate across quality-of-life dimensions, allowing the construction of different scalar indexes.*	

Sustainable Development and Environment	
Dashboards or sets of indicators	
Composite indices	
Adjusted GDPs	
Sustainable standard of living,	Adjusted net savings (ANS)
	Footprints
Recommendations	*1. Sustainability assessment requires a well-identified sub-dashboard of the global dashboard to be recommended by the Commission.* *2. The distinctive feature of all components of this sub-dashboard should be to inform about variations of those "stocks" that underpin human well-being.* *3. A monetary index of sustainability has its place in such a dashboard, but under the current state of the art, it should remain essentially focused on economic aspects of sustainability.* *4. The environmental aspects of sustainability deserve a separate follow-up based on a well-chosen set of physical indicators.*

In this perspective, the conceptual framework of the European System of Social Indicators – EUSI – (Berger-Schmitt and Noll, 2000) represents a good example of a comprehensive approach in measuring societal well-being. It tries to avoid the great part of the overlapping concepts and dimensions by respecting the policy goals defined at European level. The concepts considered by EUSI define three pillars, (i) quality of life, (ii) economic and social cohesion and (iii) sustainability.

(i) "Quality of life" concept (micro level)

The adopted approach is that defined by Zapf (1975, 1984), who proposed a model identifying the relationship between two components (objective living conditions and subjective well-being) and two degrees (low and high). The combination produces a category model of individual welfare, as represented in the following table:

Level of → ↓		Subjective well-being	
		high	**low**
Objective living conditions	**high**	*well-being*	*dissonance*
	low	*adaptation*	*deprivation*

In a similar way Michalos (2008) states that the quality of life of a community is a function of actual conditions and what individual (micro level) or the community (macro-level) makes of those conditions. "What individual or the community makes of actual conditions is in turn a function of how the conditions are perceived, what is thought and felt about those conditions, what is done and, finally, what consequences follow from all these inputs." Since an interrelation/interdependency exists between people's perceptions, thoughts, feelings and actions and their own and others' living conditions (Michalos, 2008), four different scenarios can be identified:

Level of → ↓		What people makes of conditions of life	
		Good	**Bad**
Conditions of life	**Good**	*Real Paradise*	*Fool's Hell*
	Bad	*Fool's Paradise*	*Real Hell*

(ii) "Economic and social cohesion" concept

Two goal dimensions has been distinguished:

a) reduction of disparities and inequalities and fighting social exclusion
b) strengthening of connections and social ties including the enhancement of social capital.

(iii) "Sustainability" concept

The sustainable development is referred to the World Bank's four capital approach. In particular, the four goal dimensions are the enhancement and preservation of social, human, produced and natural capital. For each type of capital two aspects have been considered: (i) preservation or enhancement of social capital of present generations and (ii) provision for future generations.

By implementing the EUSI model, we could build the following comprehensive approach:

		Level of observation
QUALITY OF LIFE	subjective well-being	micro
	objective living conditions	micro
SOCIAL COHESION	social exclusion → distribution of welfare (disparities, inequalities of individuals and societies), opportunities	micro
		macro
	social inclusion → social capital (informal networks, associations and organisations and role of societal institutions); integration of individuals and societies	micro
		macro
SUSTAINABILITY PRESERVATION OF	physical capital → behaviour affecting individual health	micro
	social capital → behaviour affecting social relations / networks	micro
		macro
	human capital → processes affecting (in terms of improvement/deterioration) people's skills, education and health	micro
	natural capital → processes affecting (in terms of improvement/deterioration) of natural resources	macro
CONDITIONS (determinant / preconditions)	individual "structure", internal and external (personality traits, …) and individual behaviour	micro
	demographic and socio-economic structures	micro
		macro
	values and interests	micro
	meeting of human needs	micro
	policies	macro

1.3 Objective and subjective components

1.3.1 Their role in measuring well-being of societies

As we have seen, a comprehensive approach is defined by a multidimensional conceptual framework, requiring both objective and subjective information observed at different levels (micro and macro). In other words, a comprehensive approach needs to integrate objective information – observed at micro (e.g. individual) level and macro (e.g. societal) level – and subjective information – observed at individual level. In policy perspective, the need for subjective indicators arises during (i) the assessment of policy results and (ii) the selection of policy objectives (Veenhoven, 2002).
The possibility to integrate objective and subjective information requires

1. a clear and shared definition of the two perspectives (what is objective and what is subjective)[5]
2. a clear conceptualization of the relationships between the two components
3. a solid methodological structure for integration .

1.3.2 Their relationship

Several conceptual frameworks of integration can be identified. Below, some patterns are introduced.

- Objective and subjective dimensions interpreted in terms of descriptive and evaluative dimensions. As previously stated, objective characteristics can be seen in terms of resources and conditions that individuals can use in order to improve their lives and to pursue their life projects. In this sense, the objective approach makes the social indicators model and Sen's capability model very similar. Consequently, the terms "objective" and "subjective" should be respectively replaced, according to Erikson (1993), with the terms "descriptive" and "evaluative."
- Objective living conditions explain subjective well-being. According to "basic needs" approach, subjective appreciation of life depends on the objective living conditions. In other words, objective living conditions is important for the happiness and satisfaction of the individuals. Seen in macro perspective, an improvement in quality of life can occur as a result of social and economic development. It should be taken into account that people's satisfaction with life in socio-economically disadvantage societies is not

[5] With reference to this issue, see the first contribution in "to go deeper" section: *Methodological aspects and technical approaches in measuring subjective well-being*.

necessarily lower than those in advantages communities. In other words, the approach based upon absolute objective standards cannot explain the variances in subjective perceptions. It should be taken account that while objective information can reveal significant discrepancies among places, subjective perceptions and satisfactions differences among individuals can show different variations.

- Subjective well-being explained by comparisons. According to "comparison" approach, subjective well-being is not directly related to objective components or individual living conditions but is based upon the comparison between individual conditions and a series of (actual or ideal) standards (Easterlin, 1974). The comparison can be made at different levels:
  - social level, when comparisons are made between different social entities (social groups, populations, countries, etc.)
  - lifetime level, when comparison are made at individual level and related to individual experiences

		Ambits of comparison				
		Housing	Work	Family	Friends	...
Standards of comparison	previous experiences					
	with other people					
	with aspirations					

The smaller the perceived gap between individuals' aspirations and their reality, the higher their subjective well-being.

This approach – known as "Michigan model" – can be considered as a fundamental step in defining an approach finalized to the evaluation of subjective well-being based upon perceived differences (Andrews & Withey, 1976; Campbell, Converse & Rodgers, 1976), particularly between aspirations and realizations. This approach registered approval but also criticism, since its definition describes the evaluation of subjective well-being exclusively in cognitive terms and excludes the affective component.

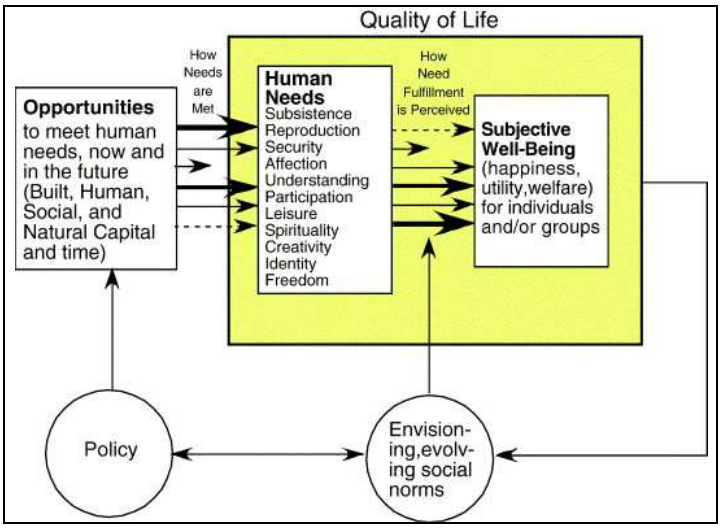
- Multiple discrepancies approach. The previous approach found successive modifications especially thanks to Michalos (1985), who formulates the *Multiple Discrepancies Theory* (MDT). In particular, Michalos introduces the concept of gap (*discrepancy*) between expectations and aspirations (*achievement gap*). According to this theory, subjective well-being represents (is function of) the perceived gap between what one has and wants, and relevant others have, the best one has had in the past, expected to have, expected to deserve, and expected with reference to needs. The gap is observed with reference to different domains (health, finances, family, job, friendships, housing, recreation, religion, transportation, and so on). In this context, happiness is considered a individual trait not dependent on living conditions.
- Disposition approach. According to this approach (Kozma et al., 1990), subjective well-being does not depend on living conditions but depends on stable individual characteristics (personality traits). For this reason, subjective well-being is not produced by the combination of perceptions in different ambits. In other words, the relationships between subjective well-being as a whole and satisfaction in different ambits is definable not in causal terms but in inferential terms (subjective well-being helps in obtaining success in different ambits, c.f. Lyubomirsky et al., 2005). Consequently, the approach pays a special attention on individual traits. Different versions of this approach were defined (*Costa-McCrae* in 1980, *Abbey-Andrews* in 1985). According to the *Kozma-Stones* approach (1990), subjective well-being is composed by two components, one expressed in terms of "reactive state", - acting in short periods (moods) – and the other expressed in terms of trait (disposition). Living conditions act on the reactive state, while the trait can attenuate the effects of that impact. Happiness is considered an additive combination of the two components (and the error). The importance of this approach is mainly in having encouraged interest in personality components of well-being and for having contributed to explanation of well-being in both conceptual and measurement terms.
- Causal approaches: bottom-up approach, top-down approach, and up-down approach. The causal explanation of well-being is at the core of several studies, which found different solutions. They were synthesized as follow by Diener (1984):
  - ***bottom-up*** approach (inductive – Simple Reactivity Model): subjective well-being is explained as a "reactive state" to the environment. The sum of the reactive measures for the defined ambits allows subjective well-being to be quantified.
  - ***top-down*** approach (deductive – Propensity Model): subjective well-being is explained by the presence of individual stable traits, like happiness (individual disposition), which determine satisfaction in single ambits.[6]

  Actually, both approaches are not able to explain completely the relationships between the observed variables. This means that causal effects can emerge in both directions. The subsequent debate[7] did not allow us to identify which of the two approaches is the best explanatory description of well-being, and

[6] The first reports on the empirical evidences concerning the concept of happiness date back to Beiser in 1974 (Stones et al., 1995).
[7] This issue was debated between Veenhoven and Stones on Social Indicators Research in the nineties.

produced the proposal of bi-directional approach (***up-down***). The proposal, which found many supporters, provides for the assessment of causal effects in both directions at the same time. This approach take into account two explanatory components, a long-period component (top-down effect), represented by the personal disposition, and a short-period component (bottom-up effect), represented by satisfaction related to circumstances. The contributions to this approach have been many (Headey et al., 1991; Lance et al., 1995) also from the methodological point of view.[8]

- Needs, opportunities and subjective well-being. A possible model of relationships between objective and subjective components of well-being is that that includes the concepts of (i) human needs, (ii) subjective well-being, and (iii) opportunities, defined in terms of four capital approach (natural capital, produced capital, human capital and social capital) and involving the role of policy, in terms of both input and output. In this perspective, societal well-being is the extent to which objective human needs are fulfilled in relation to personal or group perceptions of subjective well-being. In other words, quality of life can be seen as an interaction of human needs and the subjective perception of their fulfilment, as mediated by the opportunities available to meet the needs. (Costanza et al., 2007)



From Costanza et al. (2007)

The relationships between human needs and perceived satisfaction with each of them can be affected by mental capacity, cultural context, information, education, temperament, and the like. The ability of humans to satisfy their basic needs come from the opportunities and capabilities derived by social, human, built, natural capital (Sen, 1993). For each human need, the corresponding opportunities can be identified, as represented in the following table (from Costanza et al., 2007):

[8] The study conducted by Mallard, Lance & Michalos (1997) is particularly interested regarding the application of the MDT approach, extended with analysis of causal relationships of subjective well-being.

Human Needs	Possible descriptors	Opportunities (types of inputs needed)				
		SC	HC	BC	NC	T
Subsistance	Food, shelter, vital ecological services (clean air and water, etc.) healthcare, rest.	X	X	X	X	X
Reproduction	Nurturing of children, pregnant women. Transmission of the culture. Homemaking.	X	X		X	X
Security	Enforced predictable rules of conduct. Safety from violence at home and in public. Security of subsistence into the future. Maintain safe distance from crossing critical ecological thresholds. Care for the sick and elderly.	X		X	X	X
Affection	Solidarity, respect, tolerance, generosity, passion, receptiveness.	X			X	X
Understanding	Access to information. Intuition and rationality.	X	X	X	X	X
Participation	To act meaningfully in the world. Contribute to and have some control over political, community, and social life. Being heard. Meaningful employment. Citizenship	X	X		X	X
Leisure	Recreation, relaxation, tranquillity, access to nature, travel.	X	X	X	X	X
Spirituality	Engaging in transcendent experiences. Access to nature. Participation in a community of faith.	X	X		X	X
Creativity / emotional expression	Play, imagination, inventiveness, artistic expression.		X		X	X
Identity	Status, recognition, sense of belonging, differentiation, sense of place	X			X	
Freedom	Being able to live one's own life and nobody else's (having certain guarantees of non-interference with certain choices, such as choices regarding marriage, childbearing, sexual expression, speech and employment", mobility)	X				

SC → social capital **HC → human capital** **BC → built capital** **NC → natural capital** **T → time**

Policy and culture help to allocate the four types of capital as a means for providing the opportunities. According to this approach, overall quality of life is a function of

(a) the degree to which each identified human need is met (*fulfilment*)

(b) the *importance* ("weight") of the need to the respondent or to the group in terms of its relative contribution to their subjective well-being.

The subjective *fulfilment* and *importance* with reference to any need may vary within and across time, space contexts and groups of people. Thus, in designing and assessing quality of life, the goal should be to create a tool that will capture the weighting that is being used by a particular person (or group of persons) at a particular time and place.

The *fulfilment* and *importance* scores can be used to create a single overall metric. For example, the product between *fulfilment* and *importance* gives us a single measurement representing the degree to which needs of varying priorities are being met. This would provide an indication to individuals, groups and policy makers of where resources might be allocated (acknowledging that other factors, such as competing needs, perspectives, and resources, must also be considered in final allocation decisions).

This strategy can also provide an index that could allow us to (Costanza et al., 2007):

- compare QOL levels over time and relative to other communities
- determine whether overall QOL is improving because of changes in how well needs are being met (fulfilment) vs. changes in the weights assigned to each need (reprioritisation, possibly as a result of adaptation).
- compare QOL within and between groups of people—defined by population characteristics such as age, residential community, ethnicity, etc.
- uncover potential relationships between the fulfilment and the importance of needs
- identify possible discrepancies between fulfilment and importance grouped by type of resource required to fulfil each need
- observe variation in weights, i.e. the extent to which different components are considered important, by population characteristics
- observe variation in overall QOL (e.g., one community's needs being met over another's).

- Social epidemiology. A different approach looks at integration between objective and subjective indicators by using the logic and the perspective of *social epidemiology*, which can be defined as the systematic and comprehensive study of health, well-being, social conditions or problems, and their determinants.[9] Traditionally, social epidemiology is defined as the combination of epidemiology (the

[9] In this context, we do not refer to the alternative definition of social epidemiology as "the branch of epidemiology that studies the social distribution and social determinants of states of health" (Epidemiological Bullettin, 2002).

study of the distribution and determinants of disease and injury in human populations) with the social and behavioural sciences in order to investigate social determinants of population distributions of health, disease, and well-being, rather than treating such determinants as mere background to biomedical phenomena (Krieger, 2002).

The principal concern of social epidemiology is the study of how society and different forms of social organization influence individuals' and populations' well-being. Social epidemiology goes beyond the analysis of individual risk factors to include the study of the social context in which the well-being/ill-being phenomenon occurs (in Epidemiological Bullettin, 2002).

Even if social epidemiology is strictly related to the definition and identification of "social problems", (e.g. obesity, infectious diseases, violence, child abuse, drug use, and so on), in our viewpoint this approach turns out to be interesting also in the positive perspective of promoting quality of life (by involving not only the concept of "risk" but also the concept of "resource") since it considers both micro (personal behaviour) and macro trends in the social structure (distribution of wealth, social resources, and so on).

This perspective can help in explaining the path between exposure to social characteristics of the environment (with special attention to inequalities) and its effects on well-being by involving concepts and techniques that require the use of multidisciplinary approaches in order to analyse complex social problems.

In the traditional language of social epidemiology, "risk factors" are behaviours, attributes, individual characteristics, and exposures that may increase the probability of a specific outcome (Krieger, 2002). In order to identify risk factors, a central focus is implementing what we know about a particular condition in order to maintain and improve well-being. Inherent in this definition is the equal emphasis that we can give to objective conditions and subjective conditions as determinants of well-being.

For example, the application of this perspective allows the distribution of different levels of living conditions to be analysed in order to understand the relevant factors and their interrelationship between micro and macro trends, and to develop interventions, programs, policies, and institutions that may promote better living conditions and well-being.

The approach of social epidemiology reflects the understanding that social variables or conditions can lie on either side of the equation determining which factors affect well-being. They can be independent variables, which are the characteristics hypothesized to explain the phenomenon. They can also be the social condition or outcome that we are trying to understand, or the dependent variable. For example, depression can be a risk factor for some diseases or social conditions, such as alcohol abuse or child neglect. It can also be the outcome of particular living conditions.

2. Defining the model of measurement

The model of measurement can be conceived through two different conceptual approaches (Blalock, 1964; Diamantopoulos & Siguaw, 2006)

- Models with reflective indicators (referring to the *top-down* explanatory approach). In this case, latent constructs are measured by indicators assumed to be *reflective* in nature. In other words, the indicators are seen as functions of the latent variable, whereby changes in the latent variable are reflected (i.e. manifested) in changes in the observable indicators.[1]
  Structural relationships are identified among latent constructs by statistically relating covariation between the latent constructs and the observed variables or indicators, measuring these latent, unobserved constructs. If variation in an indicator X is associated with variation in a latent construct Y, then exogenous interventions that change Y can be detected in the indicator X. Most commonly this relationship between construct and indicator is assumed to be *reflective.* That is, the change in X is a reflection of (determined by) the change in the latent construct Y. With reflective (or *effect*) measurement models causality flows from the latent construct to the indicators.

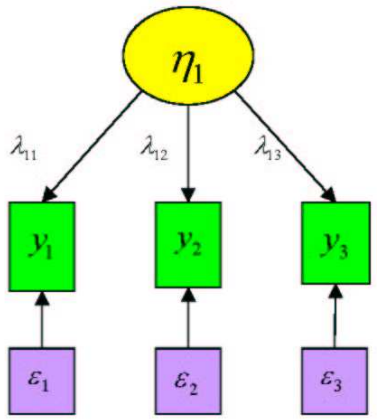
- Models with formative indicators (referring to the *bottom-up* explanatory approach). In this case, indicators are viewed as causing – rather than being caused by – the latent variable. The indicators are assumed to be *formative* (or causal) in nature. Changes in formative indicators, as firstly introduced by Blalock (1964), determine changes in the value of the latent variable. In other words, a construct can be defined as being determined by (or *formed* from) a number of indicators. In this case, causality flows from the indicator to the construct.
  An example is socio-economic status (SES), where indicators such as education, income, and occupational prestige are items that cause or form the latent variable SES. If an individual loses his or her job, the SES would be negatively affected. However, saying that a negative change has occurred in an individual's SES does not imply that there was a job loss. Furthermore, a change in an indicator (say income) does not necessarily imply a similar directional change for the other indicators (say education or occupational prestige).

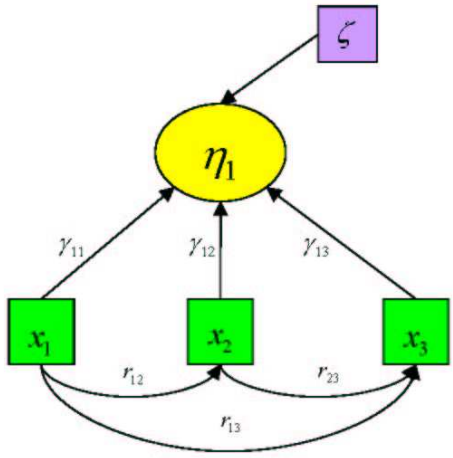
Traditionally, the reflective view is seen related to the development of scaling models applied especially (as we will see) in subjective measurement (*scale construction*), whereas the formative view is commonly seen in the development of *synthetic indicators* based on both objective and subjective measurements.

The figures presented below, summarize both perspectives:

[1] As pointed out, the proposed model is conceptually related to latent structural models that find analytical solutions through the application of the structural equations method (Asher, 1983; Bartholomew & Knott, 1999; Blalock, 1964, 1974; Bohrnstedt and Knocke, 1994; Lazarsfeld & Henry, 1968; Long, 1993a, 1993b; Maggino, 2005a; Netemeyer et al., 2003; Saris & Stronkhorst, 1990; Sullivan & Feldman, 1981; Werts et al., 1974).

Models of measurement[2]





The reflective specification implies that

$$y_i = \lambda_i^{\eta} + \varepsilon_i$$

where

η → a latent variable

$y_1, y_2, \ldots, y_n$ → a set of observable indicators

λ_i → the expected effect of η on y_i

ε_i → the measurement error for the *i*-th indicator (*i*=1, 2, …, *n*).

For $i \neq j$, it is assumed that

$$COV(\eta, \varepsilon_i) = 0 \qquad COV(\varepsilon_i, \varepsilon_j) = 0$$

$$E(\varepsilon_i) = 0$$

Reflective approach

The formative specification implies that

$$\eta = \gamma_1 x_1 + \gamma_2 x_2 + \ldots \gamma_n x_n + \zeta,$$

where

γ_i → the expected effect of x_i on η

ζ → a disturbance term,

with

$$\mathrm{COV}(x_i, \zeta) = 0$$

$$E(\zeta) = 0$$

Formative approach

The distinction between formative and reflective indicators and the necessity of a proper specification are important in order to correctly assign meaning to the relationships implied in the structural model. In this perspective, four different situations can be theoretically identified (Diamantopoulos & Siguaw, 2006) as represented in the following table:

		'Correct' auxiliary theory	
		reflective	formative
Choice of the perspective	reflective	*correct decision*	**Type I error**
	formative	**Type II error**	*correct decision*

Two outcomes are desirable and correspond to the correct adoption of the measurement perspective (operationalisation) following the correct conceptualisation of the construct of interest. The other two outcomes correspond to wrong choices. In particular, two type of error can occur:

- Type I occurs when a reflective approach has been adopted although a formative approach would have been theoretically appropriate for the construct;
- Type II occurs when a formative approach has been adopted even if the nature of the construct requires a reflective operationalisation (an index construction procedure is adopted in place of a scaling model). This error can lead to identification problems.

[2] Generally, the formal representation of models of measurement uses a particular symbology referring to the Greek alphabet:

Greek alphabet

capital	small		capital	small		capital	small		capital	small	
Α	α	alfa	Η	η	eta	Ν	ν	nu	Τ	τ	tau
Β	β	beta	Θ	θ	theta	Ξ	ξ	csi	Υ	υ	upsilon
Γ	γ	gamma	Ι	ι	iota	Ο	o	omicron	Φ	ϕ	phi
Δ	δ	delta	Κ	κ	kappa	Π	π	pi	Χ	χ	chi
Ε	ε	epsilon	Λ	λ	lambda	Ρ	ρ	rho	Ψ	ψ	psi
Ζ	ζ	zeta	Μ	μ	mu	Σ	σ	sigma	Ω	ω	omega

2.1 Reflective approach: statistical rationale

The procedure aimed at aggregating has to take into account the main specific properties of the reflective indicators, which can be synthesized as follows (Diamantopoulos & Winklhofer, 2001):

- indicators are interchangeable (the removal of an indicator does not change the essential nature of the underlying construct),
- correlations between indicators are explained by the measurement model,
- internal consistency is of fundamental importance: two uncorrelated indicators cannot measure the same construct,
- each indicator has error term (ε),
- the measurement model can be estimated only if it is placed within a larger model that incorporates effects of the latent variable.

As a result, assessment of reliability and validity can be accomplished through a statistical approach consistent with the traditional specification used in *factor models*, where an observed measure is presumed to be determined by a latent factor and a unique factor. Reflective measures are presumed to be sampled from the domain of the latent construct. The relationships between latent variables can be inferred only when the significant relationships between indicators and the corresponding latent variables are observed.

The **factor model** was originally defined in the ambit of psychometrics and experimental psychology and presented by Spearman at the beginning of the XX century and then extended by Thurstone several decades later.

The model is based upon the assumption that the total variance of each indicator x_i represents the sum of three uncorrelated components:

- common variance, portion of the total variance that is explained by the presence of the latent variable (ξ) and is measured by the correlation that each indicator registers with each of the other indicators of the same latent variable (*common variance*);
- specific variance, portion of the total variance that is not explained by the latent variable and is not correlated with the other indicators; together with the previous component composes the *reliable variance*;
- error, portion of the total variance that is not correlated with the previous ones and defines the *unreliable variance*.

Each component can be expressed as portion of the unitary total variance. For indicator x_i we can write:

$$\sigma^2_{x_i} = \sigma^2_{x_{ic}} + \sigma^2_{x_{is}} + \sigma^2_{x_{ie}}$$
$$1 = \sigma^2_{x_{ic}} + \sigma^2_{x_{is}} + \sigma^2_{x_{ie}}$$

In the ambit of the factor model, the interested is concentrated on the estimation of common variance. Specific variance and error are not estimated and are jointly considered as unique variance (*uniqueness*, δ^2):

$$\delta^2_{x_i} = \sigma^2_{x_{is}} + \sigma^2_{x_{ie}} \qquad \text{consequently} \qquad \begin{aligned} \sigma^2_{x_i} &= \sigma^2_{x_{ic}} + \sigma^2_{x_{is}} + \sigma^2_{x_{ie}} \\ \sigma^2_{x_i} &= \sigma^2_{x_{ic}} + \delta^2_{x_i} \end{aligned}$$

Generally, in a factor model more latent variables are defined. This means that almost never an indicator is explained by a single latent variable but instead can be described through a linear combination of latent variables (*common factors*). Consequently, the common variance represents the portion of the total variance jointly explained by the latent variables (*communality*, $h^2_{x_i}$).

The goal is to estimate for each indicator not only the total amount of communality but also the portions of communality that can be ascribed to the latent variables. Actually, the consequent procedure leads to evaluate the load of the latent variable ξ_j on indicator x_i. This value is expressed by the *factor loading*, $\lambda_{x_i\xi_j}$, which represents the saturation of each indicator with respect to the corresponding latent variable. *Factor loadings* values, which can be actually interpreted in term of correlation between indicator and latent variable, range

from	to
-1	+1
⇩	⇩
highest negative	*highest positive*

Since a squared correlation represents the proportion of variability that is accounted for by that relationship (coefficient of determination, R^2), a squared *factor loading* represents the amount of variability that is

accounted for by the corresponding latent variable (or factor). Consequently, for indicator x_i, communality $h^2_{x_i}$ represents the sum of the squared *factor loadings* of the latent variables (factors):

$$h^2_{x_i} = \lambda^2_{x_i\xi_1} + \lambda^2_{x_i\xi_2} + ... + \lambda^2_{x_i\xi_m}$$

where *m* = number of latent variables.

The following table tries to synthesize what we have already seen:

total variance	=	common variance	+	specific variance	+	error
$\sigma^2_{x_i}$	=	$\sigma^2_{x_{ic}}$	+	$\sigma^2_{x_{is}}$	+	$\sigma^2_{x_{ie}}$
total variance	=	communality	+	unique variance (uniqueness)		
$\sigma^2_{x_i}$	=	$h^2_{x_i}$	+	$\delta^2_{x_i}$		
total variance		reliable variance			+	error
$\sigma^2_{x_i}$	=	$h^2_{x_i} + \sigma^2_{x_{is}}$			+	$1-\left(h^2_{x_i} + \sigma^2_{x_{is}}\right)$
$\sigma^2_{x_i}$	=	$\lambda^2_{x_i\xi_1} + \lambda^2_{x_i\xi_2} + ... + \lambda^2_{x_i\xi_m}$	+	$1-h^2_{x_i}$		
fundamental equation of common factor model						
$\sigma^2_{x_i} = \sum_{j=1}^{m} \lambda^2_{x_i\xi_j} + \delta^2_{x_i}$						

The basic assumptions of the factor model can be synthesized as follows:

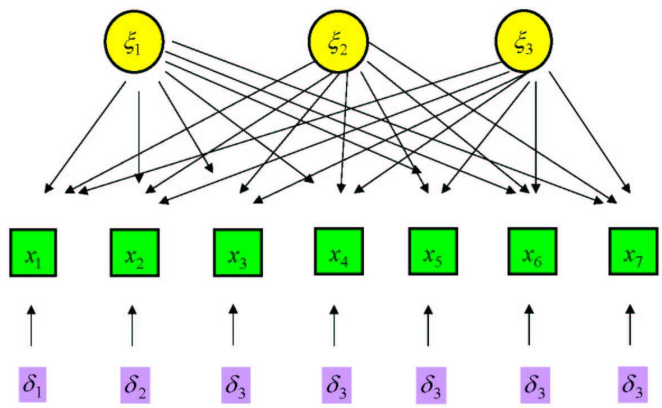
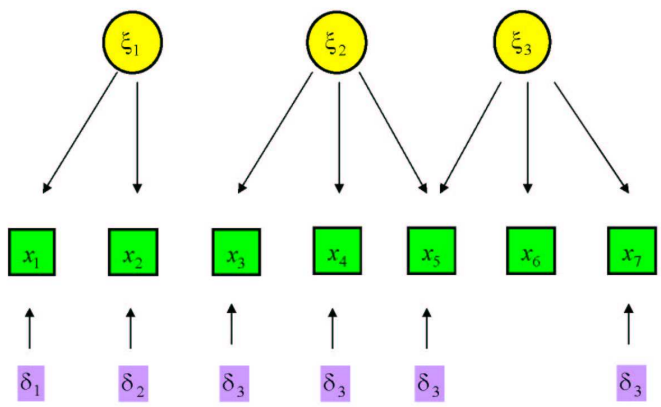
a. indicators are linearly related;
b. correlations between indicators can be interpreted only by the presence of latent variables;
c. total variance of each indicator can be expressed as a function of (i) latent variables or factors (*communality*), and (ii) individual indicator characteristics (*uniqueness*);
d. errors and disturbance factors are not interrelated and are not correlated with latent variables.

Explorative vs. confirmatory model

According to the constraints that are defined in the model, two different approaches can be identified: explorative and confirmatory.
The model is **explorative** (Tucker & MacCallum, 1993) when specifies the number of latent variables and the indicators but not the structure of the relationships between latent variables and indicators. In particular, this model assumes that:

- relationships exist (or do not) between latent variables,
- indicators are directly influenced by the latent variables,
- unique variances are not interrelated,
- latent variables are not correlated with the unique variances.

In order to estimate parameters, other assumptions are required but generally they are arbitrarily adopted. These characteristics make the explorative model limited in its actual application, especially in testing measurement model that are expressed in reflective terms.
Testing this model requires the application of particular techniques (e.g. analysis of reproduced correlations and residuals).
The model is **confirmatory** when specifies (i) the number of latent variables, (ii) the indicators, and the relationships between latent variables and indicators.
This approach allows the assumed model to be statistically verified by determining the statistical significance of the goodness of fit (Bartholomew & Knott, 1999; Bohrnstedt & Knocke, 1994; Lazarsfeld & Henry, 1968; Long, 1993a; Netemeyer et al., 2003).
Actually, testing confirmatory factor model refers directly to *Structural Equation Modelling* (SEM), which, as known, represents a statistical technique for testing and estimating causal relationships using a combination of statistical data and qualitative causal assumptions. SEM starts with a hypothesis, represented as a model, operationalises the constructs of interest with a measurement instrument, and tests the model. The causal assumptions embedded in the model often have falsifiable implications, which can be tested through data evidence. SEM models allow the researcher to explicitly capture the unreliability of measurement in the model, by determining the relationships between latent variables and indicators. The adoption of this model requires a robust theoretical framework in order to avoid any arbitrary adoption of its requisite assumptions (Bartholomew & Knott, 1999).
In the following figure, which allows the difference between the two approaches to be graphically represented, three uncorrelated latent variables (factors) (ξ_1, ξ_2, ξ_3) and seven indicators (x_1, x_2, x_3, x_4, x_5, x_6, x_7) are defined. The latent variables are named *common factors* since they shared a common influence on the indicators. Each indicator can be influenced also by a unique factor (δ_1, δ_2, δ_3, δ_4, δ_5, δ_6, δ_7). In both the examples, latent variables (ξ_1, ξ_2, ξ_3) are not interrelated. In the explorative approach, generic relationships between all the latent variables can be assumed, while in confirmatory approach, differentiated hypotheses can be defined concerning relationships between latent variables (e.g. only between ξ_1 and ξ_3).

explorative approach	confirmatory approach
	
✓ all the indicators (x_1, x_2, x_3, x_4, x_5, x_6, x_7) are influenced by all the latent variables; ✓ each indicator is influenced by a unique factor; ✓ unique factors are not interrelated.	✓ each indicator (but x_5) is influenced by one latent variable; ✓ indicators are not all influenced by a unique factor; ✓ unique factors are not interrelated.

2.2 Formative approach: statistical rationale

In formative perspective, a concept is assumed to be defined by, or to be a function of, its measurements (identified indicators). In other words, the measures are formative when the latent variable is defined as a linear sum of set of measurements. As we have seen, the formative specification implies the following relationship

$$\eta = \gamma_1 x_1 + \gamma_2 x_2 + ...\gamma_n x_n + \zeta$$

where

η	latent variable
x_i	indicator *i*
γ_i	the expected effect of x_i on η
ζ	disturbance term

The procedure aimed at aggregating has to take into account the main specific properties of the formative indicators, which can be synthesized as follows (Diamantopoulos & Winklhofer, 2001):

- the indicators are not interchangeable (omitting an indicator is omitting a part of the construct),
- the correlations between indicators are not explained by the measurement model,
- there is no reason that a specific pattern of signs (i.e. positive vs. negative) or magnitude (i.e. high vs. moderate vs. low); in other words, internal consistency is of minimal importance: two uncorrelated indicators can both serve as meaningful indicators of the construct,
- indicators do not have error terms; error variance is represented only in the disturbance terms (ζ).

Consequently, assessment of reliability and validity can be accomplished through a statistical approach consistent with a principal components specification, where the latent variable is defined as a linear combination of elementary (manifest) indicators.
While fundamental equation of the component model is

$$\sigma_{x_i}^2 = \sum_{j=1}^{m} \lambda_{x_i \xi_j}^2 + \delta_{x_i}^2$$

that is, linear combination of factor weights (λ, *loading*) explains variance of each elementary indicator x_i, principal components approach is based upon a different specification.
In defining the procedure, four critical issues must be considered (Diamantopoulos & Winklhofer, 2001):

- **Content specification**. It refers to the scope of the latent variable, the domain of content the synthetic indicator is intended to capture. In the ambit of formative model, content specification is inextricably linked with indicator specification.
- **Indicator specification**. Ideally, the indicators must cover the entire scope of the latent variable, previously described in terms of content. The exclusion of an indicator is possible but causes the risk of changing latent variable specification. However, an excessive number of indicators is undesirable for difficulties in both data collection and data analysis (number of parameters to be estimated). This issue is particular important especially in aggregative perspective.

- **Indicator collinearity**. Excessive collinearity among indicators makes it difficult to separate the distinct influence of the individual indicator on the latent variable. Multicollinear indicators turn out to be redundant and may cause the exclusion of one of them.
- **External validity**. Since exploring the suitability of indicators can not be performed through the internal consistency perspective (which is typical of reflective approach), in order to assess the wellness measurement, the synthetic indicator can be related to other measures. The basic idea is, in other words, to explore the quality of individual indicators by relating each of them with another variable (external to the synthetic indicator): only the indicators significantly related to the variable of interest would be retained. This process should be supported by a solid theoretical background. Another approach is that including some reflective indicators and estimates a multiple indicators and multiple causes (MIMIC) model (Diamantopoulos & Winklhofer, 2001).

3. Developing a system of indicators

The application of the hierarchical design, strictly connected to the definition of a proper conceptual framework, led to the consistent definition of a set of indicators. Each indicator measures and represents a distinct constituent of the observed phenomenon. Consequently, the set of indicators does not represent a pure and simple collection of indicators but provides researchers with information that is bigger than the simple summation of the elements. Systematizing the structure, also in time perspective, can characterize the set of indicators as a **system of indicators**. The basic requirements defining a system of indicators are synthesized by Noll (2004) as follows:

Key elements:	- conceptual framework requested in order to identify and justify the selection of dimensions to be measured - definition and selection of the dimensions to be measured - system architecture requested in order to support the basic structure and to define measurement procedures - identification of units to be monitored - organization of measuring and monitoring procedures
Characteristics	- objectivity: provided information should turn out to be equal or comparable, independently from who are the users; - quantification: provided values should be quantitative – obtained through standardized procedures and measures; this allows results to be reported with more precision and detail, and data to be analysed through complex methods; - efficiency and fidelity: methods, techniques and instruments that allowed data and results to be obtained have to be communicated and publicized, - economicity: the system has to produce simple, standardized, available and up-to-datable information; - generalization: the system has to allow its generalization to other similar context (exportability); - joint development: the system has to be developed in a shared way by all the "actors".
Formal criteria to be respected:	- comprehensiveness - non-redundancy - consistency - parsimoniousness

Defining a system of indicators can be seen as the realization of a demanding (in terms of resources and skills) study to be conducted through several stages.
Some risks could be faced in developing a system of indicators, like:

- the set of identified indicators is poor or bad-defined and do not fit the conceptual framework, goals and objectives;
- data are not reliable;
- indicators do not allow local realities to be compared (e.g. explanatory variables are not measured);
- system's results are not able to produce effects on the strategic, decision and planning processes.

Systems of indicators can be utilizable for both scientific and operative goals. In particular, they turn out to be useful whenever a process involve a composite evaluation (policy and technique). In this sense, a system of indicators can represent an important and valid support to subjects involved in decision processes. Decision makers need to know and manage a composite mosaic of information in order to define and evaluate priorities to be translated into actions.

3.1 Functions of systems of indicators

Systems of indicators can be distinguished according to the functions (Berger-Schmitt & Noll, 2000; Land, 2000; Noll, 1996) for which they have been created. The different functions, illustrated below, can be seen in cumulative terms since each of them requires the previous one/s.

Description and explanation functions

Monitoring. This basic function concerns and refers to the capacity of the system to:
- identify and clearly define the existing problems,
- draw promptly attention to new problems and to formulate questions,
- control and identify the main critical points of the system,
- measure changes over time if any (economic, social, etc.),
- improve all these capacities.

This function requires timing and frequencies of observation to be defined in order to evaluate any change.

Reporting. In this case the system play an important role of explanation by meeting the need
- *to describe* situation, condition, and dynamics of a certain reality (a country, an institution, etc.); in this perspective, the system answers question like "what is going on?"
- *to analyse* the existing relationships between different components; in this perspective, the system answers questions like "in which way did it happen?"

In this function, description and analysis are strictly related to reporting function, as synthetically represented below (Noll, 1996; Berger-Schmitt & Noll, 2000)

monitoring **+** *analysis* **+** *interpretation* **=** *reporting*

Evaluation functions

Forecasting. The systematic use of indicators allows the effects attributable to change in a series to be documented and consequently trends in observed reality to be forecasted. This function, representing a natural consequences of the reporting function, allows probability to reach some results by allocating resources and planning efficiency procedures *ex-ante*. (Cannavò, 2009)

Program management and performance evaluation. Systems of indicators represent valid supports to *project management* since they allow specific strategic programmes to be evaluated with reference to their realization at the present, their capacity to meet particular and specific purposes, and the prescription of future actions. In the ambit of strategic programmes, indicators must allow the following assessments:
- evaluation of the present state (where are we now?)
- identification of the priorities and the actions to be pursued (where do we want to go?)
- evaluation of adequacy (are we taking the right path to get there?)
- evaluation of progress towards goals and objectives by quantifying the strategic performances (are we there yet? can differences be observed?).

Since these systems are constructed with reference to specific programmes, they can be hardly generalized. In this perspective, this important function can play an important role in policy analysis (policy guidance and directed social change) by allowing problem definition, policy choice and evaluation of alternatives, and program monitoring (Land, 2000).

Accounting. A system can represent a useful mean of *accounting*, by which it is possible to measure and make systematically available data in order to support decision concerning allocation and destination of resources (financial and not only).[1]
In particular, this function allows a system to be able to (Cannavò, 2009)
- control ex post the suitability of the defined standards and of the planned resources flows,
- evaluate efficiency and correctness of the defined procedures,
- test adequacy and actual attainment of results.

Assessment. A system can represent a valid support to assessment procedures (certification and accountability). In this case, the goal may be to certificate or judge subjects (individuals or institutions) by discriminating their performances or to infer functioning of institutions, enterprises, or systems.

[1] The systematic recording and reporting for a given aspect of sectors (economic, social, etc.) of information like uses and resources or changes in assets and the changes in liabilities and/or the stock of assets and liabilities existing at a certain time, transform a system of indicators in a account system. According to OECD definition (http://stats.oecd.org/glossary/index.htm), the set of accounting procedures, internal mechanisms of control, books of account, and plan and chart of accounts that are used for administering, recording, and reporting on financial transactions is defined "accounting system". The transactions accounts include a balancing item which is used to equate the two sides of the accounts (e.g. resources and uses) and which is a meaningful measure of economic performance in itself. In this perspective systems should (i) embody double entry bookkeeping, (ii) record all stages of the payments and receipts process needed to recognize accounting transactions, (iii) integrate asset and liability accounts with operating accounts, and maintain records in a form that can be audited.

3.2 Elements characterising a system of indicators

The main elements defining a system of indicators are (i) aims, (ii) structure, (iii) analytic approaches, (iv) interpretative and evaluating models (Noll, 1996; Berger-Schmitt & Noll, 2000).

i. Aims

One of the main requirements of a system of indicators is the reference to the aims of its construction. Concerning this, we can distinguish between:

- *Conceptual aims* (**goals**) that represent broad statements concerning what has to be achieved or which is the problem to be faced. Usually goals are placed at macro level (national, international, etc.).
- *Operative aims* (**objectives**) that represent the instruments identified in order to attain the conceptual aims. Objectives can have different temporal prospects (monthly, four-monthly, annual, bi-annual, etc.)
- *Planning aims* (**actions**) that represent the specific activities identified to accomplish objective. They can include developments and infrastructural changes in policies, in institutions, in management instruments, etc.

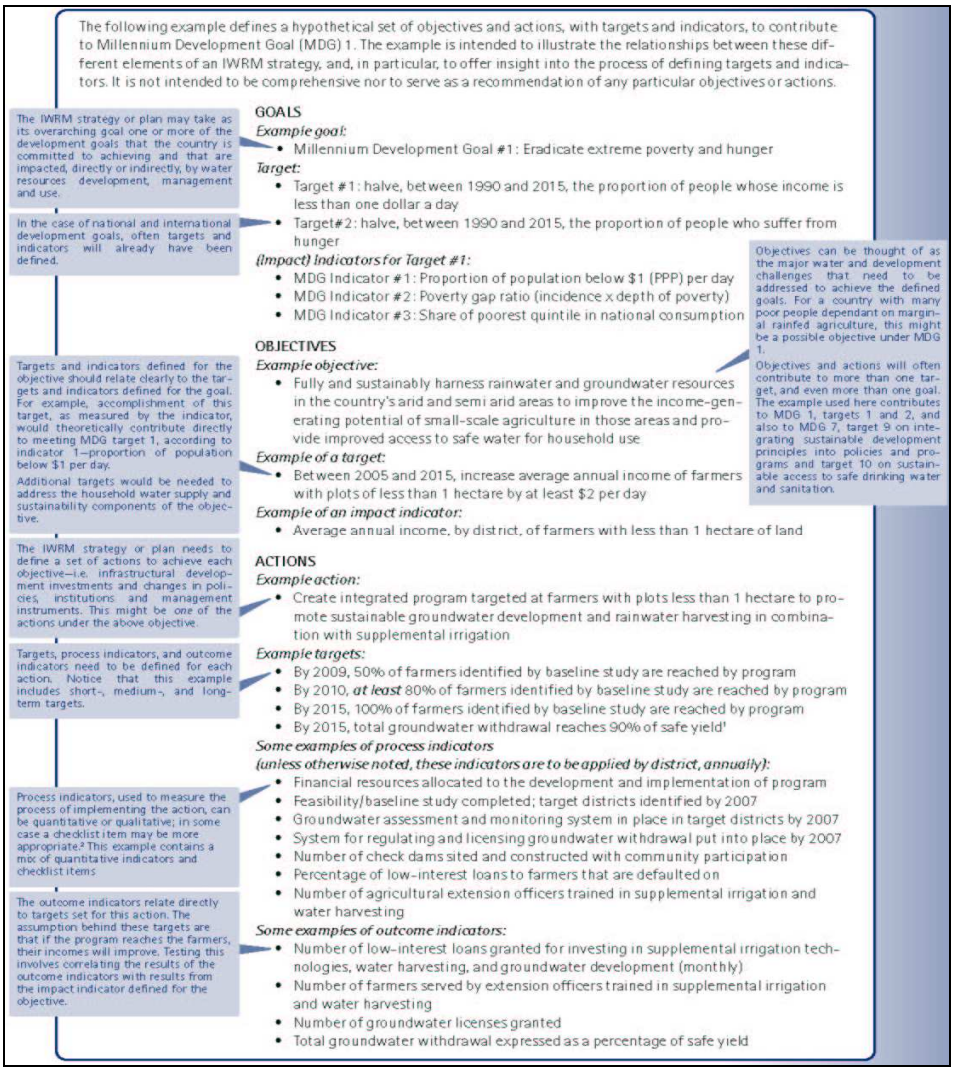
Each goal, objective and action has

- corresponding **targets**, representing those elements allowing each goal, objective and action to find measurable criteria and to define a *timetable*.
- corresponding **indicators** defined in order to assess progress towards the target with goals and objectives and the accomplishment of actions; these indicators can be distinguished in:[2]

indicators		function
- input	→	measuring resources available in the system and indicating some sort of inputs into a process
- process (intermediate output)	→	monitoring the basic progress of implementing the actions defined and outlined at strategic level
- output/outcome	→	monitoring direct results of actions
- impact	→	monitoring progress and improvement towards goals and objectives achievement

These indicators can be combined in order to define composite measures (efficacy or efficiency indicators).
In order to exemplify the relationships between goals, objectives, actions, targets and indicators, an example is presented from a technical report by the Global Water Partnership – Technical Committee *Monitoring and evaluation indicators for Integrated Water Resources Management strategies and plans* (2004):

[2] Another non-alternative classification is that that distinguishes with reference to their polarity, *positive* or *negative* quality of life observations (see the contribution to this by Alex Michalos in Sirgy et al., 2006).



The following example defines a hypothetical set of objectives and actions, with targets and indicators, to contribute to Millennium Development Goal (MDG) 1. The example is intended to illustrate the relationships between these different elements of an IWRM strategy, and, in particular, to offer insight into the process of defining targets and indicators. It is not intended to be comprehensive nor to serve as a recommendation of any particular objectives or actions.

The IWRM strategy or plan may take as its overarching goal one or more of the development goals that the country is committed to achieving and that are impacted, directly or indirectly, by water resources development, management and use.

In the case of national and international development goals, often targets and indicators will already have been defined.

GOALS

Example goal:

- Millennium Development Goal #1: Eradicate extreme poverty and hunger

Target:

- Target #1: halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day
- Target#2: halve, between 1990 and 2015, the proportion of people who suffer from hunger

(Impact) Indicators for Target #1:

- MDG Indicator #1: Proportion of population below $1 (PPP) per day
- MDG Indicator #2: Poverty gap ratio (incidence x depth of poverty)
- MDG Indicator #3: Share of poorest quintile in national consumption

Objectives can be thought of as the major water and development challenges that need to be addressed to achieve the defined goals. For a country with many poor people dependant on marginal rainfed agriculture, this might be a possible objective under MDG 1.

Objectives and actions will often contribute to more than one target, and even more than one goal. The example used here contributes to MDG 1, targets 1 and 2, and also to MDG 7, target 9 on integrating sustainable development principles into policies and programs and target 10 on sustainable access to safe drinking water and sanitation.

OBJECTIVES

Example objective:

- Fully and sustainably harness rainwater and groundwater resources in the country's arid and semi arid areas to improve the income-generating potential of small-scale agriculture in those areas and provide improved access to safe water for household use

Targets and indicators defined for the objective should relate clearly to the targets and indicators defined for the goal. For example, accomplishment of this target, as measured by the indicator, would theoretically contribute directly to meeting MDG target 1, according to indicator 1—proportion of population below $1 per day.

Additional targets would be needed to address the household water supply and sustainability components of the objective.

Example of a target:

- Between 2005 and 2015, increase average annual income of farmers with plots of less than 1 hectare by at least $2 per day

Example of an impact indicator:

- Average annual income, by district, of farmers with less than 1 hectare of land

The IWRM strategy or plan needs to define a set of actions to achieve each objective—i.e. infrastructural development investments and changes in policies, institutions and management instruments. This might be one of the actions under the above objective.

ACTIONS

Example action:

- Create integrated program targeted at farmers with plots less than 1 hectare to promote sustainable groundwater development and rainwater harvesting in combination with supplemental irrigation

Targets, process indicators, and outcome indicators need to be defined for each action. Notice that this example includes short-, medium-, and long-term targets.

Example targets:

- By 2009, 50% of farmers identified by baseline study are reached by program
- By 2010, *at least* 80% of farmers identified by baseline study are reached by program
- By 2015, 100% of farmers identified by baseline study are reached by program
- By 2015, total groundwater withdrawal reaches 90% of safe yield[1]

Some examples of process indicators
(unless otherwise noted, these indicators are to be applied by district, annually):

- Financial resources allocated to the development and implementation of program
- Feasibility/baseline study completed; target districts identified by 2007
- Groundwater assessment and monitoring system in place in target districts by 2007
- System for regulating and licensing groundwater withdrawal put into place by 2007
- Number of check dams sited and constructed with community participation
- Percentage of low-interest loans to farmers that are defaulted on
- Number of agricultural extension officers trained in supplemental irrigation and water harvesting

Process indicators, used to measure the process of implementing the action, can be quantitative or qualitative; in some case a checklist item may be more appropriate.[2] This example contains a mix of quantitative indicators and checklist items

Some examples of outcome indicators:

- Number of low-interest loans granted for investing in supplemental irrigation technologies, water harvesting, and groundwater development (monthly)
- Number of farmers served by extension officers trained in supplemental irrigation and water harvesting
- Number of groundwater licenses granted
- Total groundwater withdrawal expressed as a percentage of safe yield

The outcome indicators relate directly to targets set for this action. The assumption behind these targets are that if the program reaches the farmers, their incomes will improve. Testing this involves correlating the results of the outcome indicators with results from the impact indicator defined for the objective.

ii. Structure

The design through which data are collected and systematised defines the structure of the system. The structure can be:

- Vertical: data are collected from local levels (e.g. regions) in order to be systematized (aggregated) at a higher level (e.g. country). This structure allows policy goals to implemented, according to local information.
- Horizontal: data are collected only at one level (e.g. regional) and allow particular observational ambits (environment, education) to be monitored; usually subjective data are collected at this level.
- Local: this structure is typically designed in order to support local decisional processes. This kind of system is characterized by two levels:
  - internal, when the indicators are aimed at monitoring the internal organization of the level;
  - external, when the indicators refer to parameters existing at higher levels (e.g. transportation).

iii. Analytic approaches

Indicators have to be placed in an analytic context, consistently with aims and structure. In this perspective, as we will see later, different analytic approaches can be distinguished.

iv. Interpretative and evaluating models

The observed results can be interpreted only according to a specific reference frame. This can also include particular *standard-values*, which can be defined a priori, according to the objectives or empirical observations (e.g. surveys).

In certain cases, along with general standards, differential standards can be defined with reference to different groups (e.g. for males and females). Comparisons among groups are possible according to the availability of a unique scale for the observed and standard values.

3.3 Analysis of indicators within a system: conceptual perspectives

Consistently with aims (defined at different levels, from scientific knowledge to policy level) and structure, different analytic approaches aimed at analysing indicators within a system can be distinguished. In general terms, the analysis within a system should (Michalos, 1992):

- allow future trends to be forecasted,
- show and point out social problems,
- help in defining priorities of policies,
- allow territorial comparisons,
- suggest new ambits that need to be study in order to define new theories and a deep knowledge of social structures and functions.

In this perspective, different analytic approaches can be distinguished.

ANALYTIC APPROACH		OBJECTIVE
Monitoring	➔	to monitor developments of a specific condition (e.g. environmental conditions)
Reporting	➔	to report the results as they are obtained in a hierarchical procedure of *decision-making*
Trend	➔	to clarify development trend
Evaluation	➔	to record and evaluate the effects of planned and performed initiatives and actions
Benchmarking	➔	to compare between performances of the considered units (e.g. countries)
Assessment	➔	to clarify the impacts of planned and undertaken initiatives and actions

MONITORING ANALYSIS

The word "monitor" comes from latin *monitor –oris*, from the verb *monere*, which means *to warn, to inform, to advice*.

Originally, the term has been applied in industrial context in order to keep under continuous surveillance a machine, a process, or an operative structure. In the meantime, the concept of warning a context through meaningful information (data) has been spread also in different fields

Monitoring requires programming methods allowing the context to be observed and compared with reference standards. Programming and monitoring represent an interacting cycle with a reciprocal influence. In continuous cycles, the programming stage comes first (reference standards are knows) while, in other approaches (as in social context), it follows the monitoring stage. They can be

			context / process	
			critical	not critical
Continuous cycle	measurement apparatus and tools	simple	earthquakes (seismographs), industrial plants (pressure, temperature, humidity)	earthquakes (seismographs) industrial plants (pressure, temperature, humidity)
		complex	intensive care unit, thermonuclear power station moving systems (airplanes)	telecommunication systems informatics networks
		very complex	financial market	astronomical monitoring
High-frequency cycle	measurement apparatus and tools	simple	air pollution (accidents)	air pollution (norms controls)
		complex	accounting	
		very complex		
Low/mid-frequency cycle	measurement apparatus and tools	simple		opinion polls
		complex		social surveys
		very complex		social surveys

In the case of low/mid frequency approaches, time between observation points allow data to be normalized, ordered, correlated, interpolated, compared and spread through different channels (institutional communications, media-system).

In any approach, the aim is to transform data into meaningful information to be spread through different institutional channel and media. The relevant social impacts and political implications that can be derived, rely on bodies' and organizations' ethics and fairness in managing monitoring systems.

REPORTING ANALYSIS

According to Noll (1996), social reporting, representing the most important function of indicators systems, can be defined as "the presentation of data which enable the evaluation of living conditions of the population and their change over time." This relates with the concept of "social acceptability."
In order to be considered as socially acceptable, reporting should be based on particular basic principles:

- completeness,
- comparability (in terms of areas, years, and so on),
- embedding responsibility,
- verifiable through external bodies and organizations,
- continuously improved (improvable).

Reporting and accountability

Seen in this terms, reporting contains the idea of *accountability*: when A is obliged to inform B about A's (past or future) actions and decisions, to justify them, and to suffer punishment in the case of eventual misconduct", then "A is accountable to B. Accountability is a concept with several meanings. It is often used synonymously with such concepts as responsibility, answerability, enforcement, liability. In governance perspective, it has been central to discussions related to problems in both the public and private worlds.
In general, several types of accountability can be identified: political, administrative, market, constituency relation.[3] Accountability involves either the expectation or assumption of account-giving behaviour. Recently, accountability has become an important topic in the discussion about the legitimacy of international institutions. In fact, global administrative bodies are often criticized as having large accountability gaps. Should institutions – such as the World Bank and the International Monetary Fund who are founded and supported by wealthy nations and provide aid, in the form of grants and loans, to developing nations –be accountable to their founders and investors or to the persons and nations they help?

TREND ANALYSIS

Generally, trend analysis is used to predict future events, it could be used to estimate uncertain events. Data analysis approaches can be mainly distinguished according to the adopted design; particularly, each design allows analysis at different level:

- *macro change level*: analysis of change at group level, allowed by repeated studies;
- *micro change level*: analysis of change at individual level, allowed only by panel studies.

As described below, each level allows different goals to be accomplished (Engel and Reinecke, 1996).

Analysis of macro change: trend analysis. The main goal of analysis of macro change is to detect and to compare trends. The trend is defined as the "change expressed as function of time". Actually, the analysis is aimed at the decomposition of real trends and irregular trends; particularly the time variable can be observed in term of, or, better, the real trend can be decomposed in terms of (Glenn, 1977; Menard, 1991; Firebaugh, 1997):

3

- **Political accountability**. Political accountability is the accountability of the government, civil servants, and politicians to the public and to legislative bodies (congress, parliament). Generally, voters do not have any direct way of holding elected representatives to account during the term for which they have been elected. Moreover, some officials and legislators may be appointed rather than elected. Constitution, or statute, can empower a legislative body to hold their own members, the government, and government bodies to account. This can be through holding an internal or independent inquiry. The powers, procedures, and sanctions vary from country to country. The legislature may have the power to impeach the individual, remove them, or suspend them from office for a period. The accused person might also decide to resign before trial. In parliamentary systems, the government relies on the support or parliament, which gives parliament power to hold the government to account. For example, some parliaments can motion for a vote of no confidence in the government.
- **Administrative accountability**. It refers to internal rules and norms as well as some independent commission are mechanisms to hold civil servant within the administration of government accountable. Within department or ministry, firstly, behaviour is bounded by rules and regulations; secondly, civil servants are subordinates in a hierarchy and accountable to superiors. Apart from internal checks, some supervisory bodies accept complaints from citizens, bridging government and society to hold civil servants accountable to citizens, but not merely governmental departments.
- **Market accountability**. Nowadays, it is "customer-driven" and is aimed at providing convenience and various choices to citizens; ideally, this perspective should improve quality of service. The standard of assessment for accountability requires a neutral body. Government can choose among a shortlist of companies for outsourced service; within the contracting period, government can hold the company by rewriting contracts or by choosing another company.
- **Constituency relations.** A particular agency or the government is accountable if voices from agencies, groups, or institutions, which is outside the public sector and representing citizens' interests in a particular constituency or field, are heard. Moreover, the government is obliged to empower members of agencies with political rights to run for elections and be elected; or, appoint them into the public sector as a way to hold the government representative and ensure voices from all constituencies are included in policy-making process.

- *period* of observation (for instance "year"), interpretable as "changes over time" effect, that refers to change produced by influenced related to historical age under study;
- *age* of observed individuals, interpretable as "life cycle and developmental changes" effect, that refers to change produced by influenced related to age (considered as individual life-cycle status);
- *cohort*: each observed individual is member of cohort, defined with reference to particular conditions (geographical area, event, time of particular event, generation, year of birth, and so on) and interpretable in terms of effects. Observed cohort difference could be interpreted with reference to common experiences or reactions of a cohort (according to the cohort definition).

Each of these may represent an explanatory effect of change to be considered (separately or in combination) in the model. Synthetically:

year of birth = year of measurement – years

cohort = period – age

Four methods can be identified in order to study group trends (Firebaugh, 1997):

1. trend analysis: analysis of average changes in a group over time and comparison between different trends (coincident, parallel, converging, diverging, crossed trends);
2. proximate decomposition of trends (proximate source of change), distinguishing between net change among individuals and gross change due to group turnover; the analysis is based upon linear regression approach;
3. change decomposition of aggregate change in one variable in terms of change in levels and effects of other variables; the analysis is based on decomposition equation regression;
4. changing-parameter method: analysis of change in effects of variables at individual level, in order to determine the time-dependence of individual-level relationships.

Since repeated studies are based on substantially independent samples, trend analysis provides sufficient description of change but fails in providing empirical explanations of change process.

Analysis of micro change: process analysis. It is not simply a descriptive analysis but also explanatory (process analysis or internal analysis). The analysis, allowed by panel studies, is accomplished at individual-level by investigating covariation over time. By taking into account the fundamental constraints aimed at establishing a causal relationship (covariation, temporal precedence and nonspuriousness), we can observe the following kinds of change (Menard, 1991):

- *initiation*, referring to the first time that a case enters a particular state,
- *escalation/reduction*, referring to the entry of, respectively, a higher or a lower state (on an ordinal scale),
- *suspension*, referring to a permanent or temporary exit from all states that indicate involvement a particular state; this kind of change is not always significantly present.

Collecting data in order to study and analyze processes described by this model may be difficult because of possible long terms occurring. The problem, called 'left-hand censoring', indicates the failure to detect a change because it happened before the period of data collection.

Consequently in order to unravel causal relationships, it is important to qualify the model in terms of adequateness of time lag (that is, interval between data collection periods of time) in order to allow

- change to be detected in a variable clearly separated from change in another;[4]
- an effect to be produced by occurrence of cause.

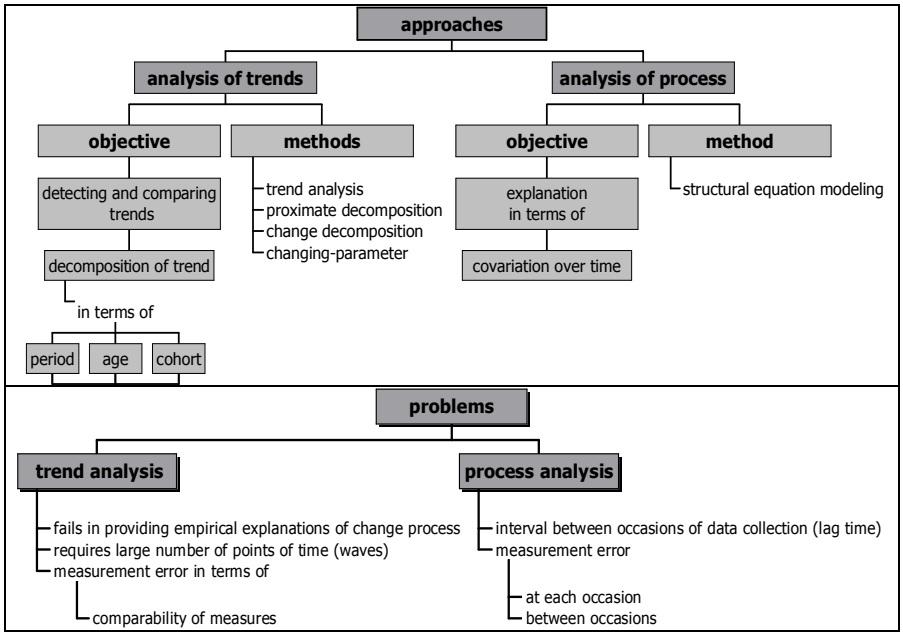
In order to unravel supposed causal relationships by collected data, traditional data analysis methods can be classified as represented in the following table (Menard, 1991).

[4] If the change in both variables occurs in the same period, there are different possible explanations (Menard, 1991):

a. the two variable measure the same thing,
b. the two variables are spuriously related, having a common cause producing changes in both,
c. the length of measurement period does not allow to separate the two changes.

Dependent variable	Independent variable	Methods of analysis
Quantitative/continuous	Quantitative/continuous	Differential equations
		Regression
		Multivariate ARIMA time-series analysis
		Latent variable structural equation models
	Mixed continuous and categorical	ANOVA with ANCOVA
		Regression with dummy variables
	Qualitative/categorical	ANOVA
		Nonparametric ANOVA
		Dummy variable regression
Qualitative/categorical	Quantitative/continuous	Discriminant analysis
		Logit or probit analysis
		Logistic regression
		Survival/event history analysis
	Mixed continuous and categorical	Log-linear analysis
		Logistic regression
		Survival/event history analysis
	Qualitative/categorical	Log-linear analysis
		Multistate life table models
		Survival/event history analysis

Moreover, the following two pictures summarize, respectively, characteristics and problems of the two levels of analysis.



EVALUATION ANALYSIS

Evaluation analysis can be conducted through different approaches, which are classifiable with reference to

(i) **orientation** (role of values / ethical consideration):
- pseudo-evaluation → politically oriented
- quasi-evaluation → question oriented
- true evaluation → value oriented

(ii) **principles** concerning the control of knowledge
- subjectivity
- objectivity

(iii) **perspectives**:
- elite (focusing on interests of particular groups)
- mass (participatory perspective)

(iv) **characteristics**:
- aims (main considerations or cues practitioners use to organize the evaluation analysis/study)
- expected / desired outcomes (for the evaluation study)

When the above concepts are considered simultaneously, fifteen evaluation approaches can be identified.

approaches for conducting evaluations: summary					
ORIENTATION	**PRINCIPLE**	**PERSPECTIVE**	**APPROACHES**	**CHARACTERISTICS**	
				Aims	**Expected outcomes**
POLITICALLY ORIENTED	SUBJECTIVE	ELITE	Politically controlled	Threats	Get, keep or increase influence, power or money.
			Public relations	Propaganda needs	Create positive public image.
QUESTION ORIENTED	OBJECTIVE	ELITE	Experimental research	Causal relationships	Determine causal relationships between variables.
			Management information systems	Scientific efficiency	Continuously supply evidence needed to fund, direct, & control programs.
			Testing programs	Individual differences	Compare test scores of individuals & groups to selected norms.
			Objectives-based	Objectives	Relates outcomes to objectives.
			Content analysis	Content of a communication	Describe & draw conclusion about a communication.
	OBJECTIVE	MASS	Accountability	Performance expectations	Provide constituents with an accurate accounting of results.
			Decision-oriented	Decisions	Provide a knowledge & value base for making & defending decisions.
			Policy studies	Broad issues	Identify and assess potential costs & benefits of competing policies.
			Consumer-oriented	Generalized needs & values, effects	Judge the relative merits of alternative goods & services.
VALUE ORIENTED	SUBJECTIVE	ELITE	Accreditation / certification	Standards & guidelines	Determine if institutions, programs, & personnel should be approved to perform specified functions.
			Connoisseur	Critical guideposts	Critically describe, appraise, & illuminate an object.
			Adversary	"Hot" issues	Present the pro & cons of an issue.
			Client-centered	Specific concerns & issues	Foster understanding of activities & how they are valued in a given setting & from a variety of perspectives.

BENCHMARKING ANALYSIS

Generally speaking, benchmarking is the process of comparing the cost, cycle time, productivity, or quality of a specific process or method to another that is widely considered to be a best practice. Essentially, benchmarking provides a snapshot of the performance and helps in understanding where each case is in relation to a particular standard. The result often stimulates each case in making changes in order to obtain improvements.

This kind of analysis is considered a strategic process, through which various aspects of the monitored system can be evaluated with reference to best practice. This allows plans to be developed aimed at making improvements or adopting best practice, usually with the aim of increasing some aspects of performance.

Different kinds of benchmarking exist:

- **Process benchmarking**: the goal is identifying and observing the best practices from one or more benchmark cases. Activity analysis will be required where the objective is to benchmark cost and efficiency.
- **Financial benchmarking**: performing a financial analysis and comparing the results in an effort to assess – e.g. – the overall competitiveness.
- **Performance benchmarking**: it allows each case to assess its position by comparing results (products, services, and so on) with those of target case.
- **Output benchmarking**: it allows strengths and weaknesses to be found.
- **Strategic benchmarking**: it involves observing how others plan the activities.
- **Functional benchmarking**: it focuses on a single function in order to improve the operation of that particular function. In some cases, more complex functions (Finance, Accounting and Information and Communication Technology) need to be disaggregated into processes in order to make valid comparison in terms of, e.g., cost and efficiency.

From the analytic point of view, the main approaches allowing comparisons to be accomplished are Data Envelopment Analysis (DEA) and regression analysis.[5]

In studying well-being of societies, the benchmark process needs to be considered with great attention and care and needs a strong a shared conceptual framework.

ASSESSMENT ANALYSIS

Three main assessment purposes can be distinguished:

- to certify or qualify cases by discriminating among them
- to assist in the process by providing an understanding of cases' performance
- to make inferences about the functioning of the observed system.

Impact assessment

In policy perspective, assessment is interpreted in terms of *impact assessment*, that is a set of logical steps which structure the preparation of policy proposals. It involves building on and developing the practices that

[5] With regression analysis cases that performed better than average can be rewarded while cases that performed worse than average can be penalized. Such benchmarking studies are used to create yardstick comparisons, allowing outsiders to evaluate the performance of operators in an industry. A variety of advanced statistical techniques, including stochastic frontier analysis, have been utilized to identify high performers and weak performers.

already accompany the process of policy development by deepening the analysis and formalising the results in an autonomous report.
Impact assessment should be considered a support to policy decision-making, not a substitute for it. It involves a number of basic analytical questions: *What is the nature, magnitude and evolution of the problem? What should be the objectives pursued by the system? What are the main policy options for reaching these objectives? What are the likely economic, social and environmental impacts of those options? What are the advantages and disadvantages of the main options? And, last but not least: How could future monitoring and evaluation be organised?* An impact assessment needs not involve a long and detailed study in every case (*proportionate analysis*), but it should allow for an informed debate in all cases.
"Impact assessment follows six key steps in a logical order. However, it is important to understand that it is very much an iterative process, where it is likely that your earlier steps will need to be revisited in the light of work undertaken later in the process. This 'back and forth' process is relevant for all of the major impact assessment steps, but may be of particular importance for setting objectives.

Impact analysis[6]

The analysis of the impacts of each of the options is a crucial element of the impact assessment process and should be conducted for the most relevant policy options, including the no-policy change option. This exercise will help you supply information about likely impacts across the identified policy dimensions (e.g., economic, environmental, and social) and will also help in identifying particular enhancing measures (in terms of effectiveness and efficiency) and/or mitigating measures (such as longer transition periods, exemptions for certain groups or redistributive measures). The analysis of impacts aims also at

- providing sufficient and clear information on the impacts of the various policy options by comparing the options to each other, with the 'no policy change' option
- predicting, across a range of different policy areas, the likely consequences of each option.

Three analysis steps can be identified.

Step 1: *Identifying impacts of a policy, why they occur and who is affected*.
The first step is to identify those (intentional and unintentional) impacts that are likely to occur as a consequence of implementing the policy. The screening can be done by consulting internal/external stakeholders and experts. Impacts identification requires

- stating clearly the links between cause (the action, instrument, etc) and effects (the impacts)
- identifying to what extent the proposed action(s) will contribute to reaching the (operational) objective(s)
- identify systematically who is affected by the identified impacts and over what timescale the impacts will occur.

In this perspective, a useful approach to identifying impacts is to build a causal model in 'bottom-up' terms. A flowchart or map of impacts can then be built that sketches out cause-and-effect linkages between each of the policy options/instruments and their impacts.

Step 2: *Identifying the most important impacts*.
Identifying the most important impacts can be done quickly and cheaply by using the causal model described above. It provides a foundation upon which more sophisticated analyses can be built, taking into account the significance and nature of each proposal (cf. the principle of proportionate analysis).

Step 3: *Advanced analysis of impacts*.
The above steps allow a qualitative analysis of a proposal's impacts to be accomplished.
The in-depth qualitative analysis of selected impacts focuses on selected impacts or chains of impacts about which both qualitative and quantitative data are collected and analysed qualitatively, typically using a case study/scenario approach.
Quantitative analysis of impacts focuses on either a limited and selected impacts. Essentially, the aim is to understand the extent of the impacts of the policy options and to estimate the costs and benefits in monetary form when this is feasible. Combining quantitative and qualitative methodologies is good and fruitful practice in order to ensure an adequate consideration a broader range of direct and indirect, social, environmental and economic impacts.
Regardless of the adopted analytical approach, the obtained results should be:

- transparent: it must be clear to others how you arrived at your estimation of impacts.
- reproducible: others must be able to arrive at the same results, using the same data and approach.

[6] Based upon the European Commission – Joint Research Centre's Handbook, *Impact Assessment Tools. Supporting impact assessment in the European Commission*, http://iatools.jrc.ec.europa.eu/bin/view/IQTool/HandBook.html (June, 2009).

- robust: if using different methods or assumptions to estimate the impacts gives very different results, this may call into question the reliability of accomplished analysis. If results depend on the choice of a specific analytical method, or if the data used are not fully reliable, it is essential to set this out.

4. Reducing the complexity of data structure

4.1 Checking data

The data reduction procedures should be preceded by a data preparation process aimed at checking data completeness (presence of missing data) and data scales homogeneity.

4.1.1 Missing data: imputation strategies and techniques

The lack of **completeness of collected data** can undermine the accuracy and precision of indicators. In this perspective it is important to ascertain the presence of *missing* data and to identify the proper strategy aimed at their management and eventually the proper imputation techniques (Little & Rubin, 1987; Rubin, 1987). In presence of missing data, the strategies used for this can be different:

- *passive* (*complete case approach*), when missing values are ignored (with reference to two different approach, *listwise deletion* or *pairwise deletion*); only the units that present valid values are considered;
- *active*, when missing values are collected in a single new category considered valid for subsequent analysis; this strategy is also used to explain the presence of missing values and their effects;
- *imputation*, when each missing value is substituted by one (single imputation), or two or more (multiple imputation) plausible values. This new value is estimated according to particular techniques. The researcher needs to consider the adoption of this strategy carefully and cautiously because of its potentially strong impact on the creation of the indicator. According to Dempster and Rubin (1983), imputation represents a seductive prospect since it allows the pleasant illusion of completeness of data; however, it is dangerous since it gives a deceptive impression of the dimension of the problem, which seems to be legitimately managed this way.

The imputation techniques can be summarized as follows:

- *identification of a value* corresponding to one that is
  - intermediate or randomly chosen among those in the range of admissible values;
  - the mean/median of the values observed in the particular study or in other analogous and homogeneous studies; this approach produces a distortion in the distribution since it converts the whole group of missing values into a constant;
  - obtained by applying iterative regression techniques to the observed indicators (with the exception of the one presenting the missing value).
- *hot-deck technique*: the complex combined application of *clustering* and regression techniques allows the identification of a unit or a group of units ("donors") showing similar profiles to the one that presents the missing value;
- *cold-deck technique*: this technique is similar to the previous one but selects donors from another dataset;
- *techniques based upon maximum-likelihood estimations*: one of these techniques is based upon an iterative algorithm (*expectation-maximisation*) that first imputes an average value (E-*step*), then (M-*step*) calculated parameters (mean, standard deviation, etc.) until a convergence condition is reached. This approach can produce great distortion especially when statistical different populations are observed;
- *techniques based upon other mathematical approaches (*Markov chains, MonteCarlo approach, and so on).

Each technique presents disadvantages like strengthening pre-existing statistical relationships.
In order to deal with these constraints, it is possible to add a random value (called "residual") to the value that has to be imputed (*stochastic* replacement).
In the attempt to maximize the positive characteristics of imputation techniques, the value to be imputed must be computed only from values belonging to the same statistical population.

4.1.2 Transforming and standardizing data

Before proceeding with any indicators analysis, data should be submitted to a process aimed at checking data scales. This process is aimed at

- giving a new meaning to each indicator

- giving comparability to indicators in view of synthetic indicators creation.

These objectives require proper, respectively, data transformation and standardization techniques.
Since each available technical solution can produce different results and meaning in data, the choice of the procedure is not a trivial matter.
In selecting the more appropriate approach, the following criteria should be taken into account:

- data properties,
- indicator's original meaning,
- values to be emphasized or to be penalized,
- whether or not using absolute value,
- whether or not comparing cases with each other or with a reference unit,
- whether or not evaluating units across time.

The most common approaches (Nardo et al., 2005a, Sharpe & Salzman, 2004) are those represented in the following table:

Technique	The original value is transformed into a new one showing …		Notes
a. Index number	… its position in respect a particular reference value.	$I_{ic}^{t} = \frac{x_{ic}^{t}}{x_{ir}^{t0}}$	The reference can be represented by (i) a unit (a country, a city, and so on), (ii) an averaged value, (iii) a target value The results can be expressed in percentage. It is simple and not affected by outlier. The original information is completely lost. Even if the reference could be defined arbitrarily. the method should be applied whenever a possible and not arbitrary reference can be identified or defined.
		$I_{ic}^{t} = \frac{x_{ic}^{t}}{\overline{x_{i}^{t}}}$	
		$I_{ic}^{t} = \frac{x_{ic}^{t} - x_{ir}^{t0}}{x_{ir}^{t0}}$	
		$I_{ic}^{t} = \frac{x_{ic}^{t} - x_{ic}^{t-1}}{x_{ic}^{t}}$	In time series, the reference value could be defined by taken into account the series (e.g. the value of the previous year).
b. Rescaling	… its position in respect the (actual or theoretical) highest value or the observed range.	$I_{ic}^{t} = \frac{x_{ic}^{t}}{\max\left(x_{i}^{t}\right)}$	The obtained value will fall in the range [0; 1] or [0; 100] if multiplied by 100. Easy to be interpreted. It can produce a distorted effect in presence of outlier.
		$I_{ic}^{t} = \frac{x_{ic}^{t} - \min\left(x_{i}^{t}\right)}{\max\left(x_{i}^{t}\right) - \min\left(x_{i}^{t}\right)}$	
c. Standard score (*z-score*)	… its distance from the mean in terms of standard deviation.	$I_{ic}^{t} = \frac{x_{ic}^{t} - \overline{x_{i}^{t}}}{\sigma_{i}^{t}}$	This approach preserves the distribution shape of the original observed scores. In order to avoid outlier effect, the mean and standard deviation can be calculated by avoiding the extreme values (trimming procedure).
d. Ranking	… its rank position.	$I_{ic}^{t} = rank\left(x_{ic}^{t}\right)$	Not affected by outlier. The original information is completely lost. Variant: original score transformed in the corresponding percentile.
e. Categorization	… its status in respect a particular criterion (eg. "completely achieved", "partially achieved", "not achieved")	See next page (*)	Stable results

where

I_{ic}^{t}	indicator *i*'s value as regards case *c* at moment *t* after transformation	x_{ic}^{t-1}	indicator *i*'s value as regards case *c* at moment *t-1*
x_{ic}^{t}	indicator *i*'s value as regards case *c* at moment *t* before transformation	$\overline{x_{i}^{t}}$	indicator *i*'s mean
x_{ir}^{t0}	indicator *i*'s value as regards reference unit *r*	σ_{i}^{t}	indicator *i*'s standard deviation

(*) A particular approach is that elaborated at the *United Nations Research Institute for Social Development* (Drewnowski, 1970) in order to observe living conditions of countries. This kind of transformation requires strong assumption concerning the categories definition. The approach is widely applied for both objective indicators and subjective indicators; even if it allows problems due to outliers to be avoided, it prevents from absolute comparisons between units. The transformation proceeds through the following stages:

I. **Identifying critical points and categories**:
- Point *O*: *survival point*;
- Point *M*: *minimum requirement point*;
- Point *F*: *full satisfaction point*.

The critical points allow four categories to be identified:

1. < *0* → precarious level
2. ⩾ *0* e < *M* → unsatisfactory level
3. ⩾ *M* e < *F* → satisfactory level
4. ⩾ *F* → optimal level

II. Re-proportioning of indicators on a common scale so that the minimum value coincides with the point *O* and that the maximum (=100) coincides with point *M*:

$$I_{ic} = \frac{x_{ic} - {}_0x_i}{{}_M x_i - {}_0x_i} * d * 100 \text{ where}$$

x_{ic} indicator *i*'s value as regards case *c*

I_{ic} indicator *i*'s value as regards case *c*

${}_0x_i$ indicator *i*'s point *0*

${}_M x_i$ indicator *i*'s point *M*

d value generally based upon Gini coefficient (1-R)

The procedure allows to make differences between cases more evident.

In same cases, the objective is to re-proportion indicators' category codes when the indicators show different numbers of ranked categories. In these cases, the number representing the rank should be harmonized by following the symmetry criterion. The following table allows to establish the new ranking codes harmonized with the scale $1 \div 7$.

Redefinition of the ranking codes to be assigned to each indicator, harmonized with the scale $1 \div 7$

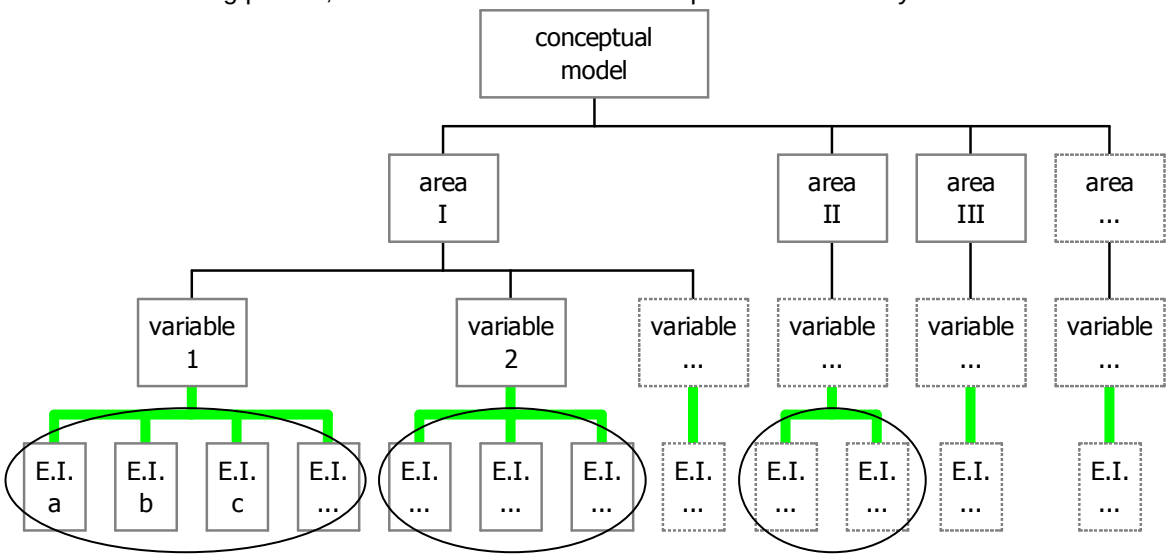
Original values assigned to the categories →	Number of categories ↓						← New values to be assigned to the categories
	2	3	4	5	6	7	
1	**1**	**1**	**1**	**1**	**1**	**1**	
2	**7**	**4**	**3**	**2**	**2**	**2**	
3		**7**	**5**	**4**	**3**	**3**	
4			**7**	**6**	**5**	**4**	
5				**7**	**6**	**5**	
6					**7**	**6**	
7						**7**	

The impact of each technique on the quality of the results can be checked by robustness tests (Nardo et al., 2005a).

4.2 Aggregating indicators and creating synthetic indicators

In order to better manage the complexity of the measured data, analytical models are required providing for significant data aggregations at different levels in order to ensure correct and different comparisons, transversal (between groups, regions) and longitudinal at both micro and macro levels.
In other words, the complexity of this structure can be reduced by defining and applying additional models. The purpose of these models is – through the definition and adoption of particular assumptions – to condense and synthesize the dimension by referring to the *multiple measures*.
The construction of synthetic indicators should be consistent with the adopted measurement model. In this context, the traditional distinction between formative and reflective is particular important since aggregation of indicators has to be consistently accomplished. In other words, indicators can be aggregated into complex

structure through a consistent methodology according to two different criteria: (i) *reflective criterion* and (ii) *formative criterion*. In both cases, the condensation of elementary indicators, considered multiple measures, produces new synthetic values obtained by applying the appropriate aggregating model. Each synthetic indicator tries to re-establish the unity of the described concept described by the corresponding latent variable. In the following picture, the indicators that will make up three different synthetic indicators:



From the operational point of view, the condensation process leads to a simplification of the data matrix in terms of a lower number of columns, illustrating as following:

		elementary indicators					
		1	2	...	j	...	K
individuals	1	V_{11}	V_{12}		V_{1j}		V_{1k}
	2	V_{21}	V_{22}		V_{2j}		V_{2k}
	...	...	...	...	...	...	...
	i	V_{i1}	V_{i2}		V_{ij}		V_{ik}
	...	...	...	...	...	...	...
	n	V_{n1}	V_{n2}		V_{nj}		V_{nk}

➔

		synthetic indicators				
		1	...	j	...	C
individuals	1	I_{11}		I_{1j}		I_{1c}
	2	I_{21}		I_{2j}		I_{2c}
	...	...	...	...	...	...
	i	I_{i1}		I_{ij}		I_{ic}
	...	...	...	...	...	...
	n	I_{n1}		I_{nj}		I_{nc}

4.3 Aggregating observed units and defining macro-units

When the conceptual framework leads with a multidimensional construct emerging from the evaluation of multiple aspects observed at different levels (individual, community, national, and global), the study needs to take into account the different levels at which information is collected and has to be analysed. In fact, some characteristics are observable only at macro level, others can be observed at micro level.
From the operational point of view, this aggregation process leads to a simplification of the data matrix in terms of a lower number of rows, illustrated as following:

		indicators				
		1	...	j	...	C
individuals	1	I_{11}		I_{1j}		I_{1c}
	2	I_{21}		I_{2j}		I_{2c}
	...	...	...	...	...	...
	i	I_{i1}		I_{ij}		I_{ic}
	...	...	...	...	...	...
	n	I_{n1}		I_{nj}		I_{nc}

➔

		indicators				
		1	...	j	...	C
Aggregated units	1	S_{11}		S_{1j}		S_{1c}
	...	...	...	...	...	...
	k	S_{i1}		S_{ij}		S_{ic}
	...	...	...	...	...	...
	G	S_{g1}		S_{gj}		S_{gc}

For example, Costanza and others (2007) propose a list of (illustrative rather than exhaustive) indicators for measuring human needs by distinguishing the level of observation. At the highest level (national)[1] the need of data aggregation is highlighted:

Need	**Individual level**	**National level**
Subsistence	Self reports on: caloric intake access to clean air, water Access to health care	National data on: caloric deficiencies Aggregated data health care
Reproduction and care	Self reports on: maternity leave/child care Family provision for care Household and child care allocation within the household	National data on: existence and scope of family leave laws Aggregated data on family provision and care Aggregated data on household duties
Security	Self reports on: who provides care in case of acute, chronic illness Who provides care for aged parents etc. Interpersonal violence experiences Environmental practices	National data on: nursing homes, shared housing, multigenerational households Aggregated data on who provides care Crime statistics Aggregated data on environmental practices
Affection	Self reports on: level of attachment to significant others	National data on: aggregated data on levels of attachment, suicide, homicide
Understanding	Self reports on: newspaper, radio, tv, internet usage for news information	Aggregated data on: media usage for news
Participation	Self reports on: volunteering, association memberships	National data on: aggregated data on volunteering, association membership
Leisure	Self reports on: time use, activities pursued, money spent	Aggregated data: time use, activities pursued and money spent
Spirituality	Self reports on: spiritual/transcendent experiences spiritual organization membership Time spent on spiritual activities	National data on: religious/spiritual book production/sales number and diversity of religious/spiritual organizations Aggregated data on self-described spirituality
Creativity / emotional expression	Self reports on: free time use Sense of play in work, etc.	National data on "elite culture" organizations, events, participation Aggregated data on free time use
Identity	Self reports on: major statuses, sense of "place"	Aggregated data on: statuses and sense of "place"
Freedom	Self reports on: personal freedoms in various social contexts (family, work, religion, etc.)	National data on: freedom indicators, expression, press, voting policies etc…

In many cases, information needs be aggregated in order to perform analyses on same-level data. In other words, in order to accomplish a correct analysis and obtain a composite picture (e.g. national), the information collected at micro level needs to be in someway aggregated to the proper scales (spatial or temporal).

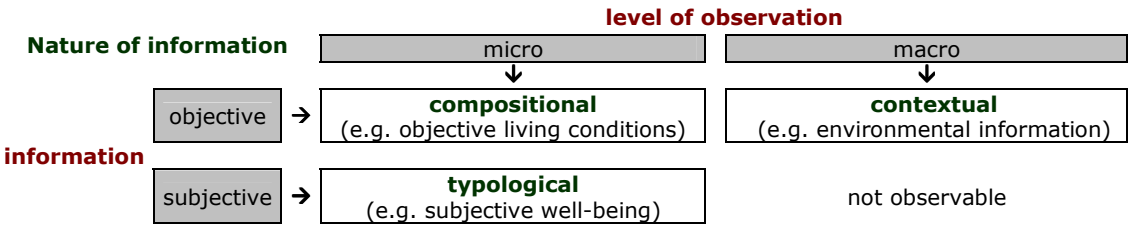
Actually, the aggregation requires values observed at micro/lower levels (usually, individual) to be condensed to higher levels in order to produce new meaningful units identified according to different kind of scales. Generally, the aggregated macro units functionally refer to pre-existent partitions, such as **groups** (social, generational, etc.),[2] **areas** (geographical, administrative, etc.), **time periods** (years, decades, etc.).

This problem involves both objective and subjective indicators, with different solutions, according to the nature of information. [3]

[1] The logic represented in the table is easily applicable to other levels (community, regional, and so on).

[2] If the subjective information is collected from a probabilistic sample, it is possible to take into account the weight that each sampled individual has with reference to the correspondent population by assigning a differential weight. The matter is dealt with statistical approaches related to inference methods and sampling techniques.

[3] Aggregation of scores collected at micro levels is a well-known issue in many scientific fields, like economics and informatics, where particular analytic approaches are applied (like the probabilistic aggregation analysis). In econometric fields, particular empirical methodologies have been developed, allowing the explanation of systematic individual differences (*compositional heterogeneity*) that can have important consequences in interpreting aggregated values (Stoker, 1993).



Objective information can be reduced to the proper scale by considering:

(a) *compositional nature of data*, when information, observable at micro level, can be referred to a given population (e.g., objective living condition),

(b) *contextual nature of data*, when information, not-observable at micro level, refers directly to a given area/territory (e.g., environmental conditions).

Subjective information (like those related to subjective well-being) is not cumulative/compositional in its nature. Consequently, applying the traditional statistical averaging techniques (requiring summing up individuals' values) does not allow us to highlight the distributional characteristics of each aggregated units, which consequently could not be correctly compared in order to avoid the well-known *ecological fallacy*.

Consequently, aggregating subjective information requires particular procedures, adopting techniques allowing the aggregation of individual scores (*aggregating criteria*). Regarding this issue, attempts aimed at weighting average values by different criteria can be identified (Kalmijn & Veenhoven, 2005; Veenhoven, 2005). A solution could be represented by adopting a **homogeneity** criterion: as regards each functional level (area, group, and so on), individuals' values are aggregated (and averaged) only if cases are homogeneous according to the characteristics of interest. The produced aggregated units (**typologies**) allows each functional level to be represented by a certain number of groups and their incidences. Identification of typologies requires particular analytical approaches, allowing homogeneous groups among individual cases to be identified (Aldenderfer & Blashfield, 1984; Bailey, 1994; Corter, 1996; Hair et al., 1998; Lis & Sambin, 1977):

- **segmentation analysis**, which can be conducted through different procedures (*Hierarchical Cluster Analysis, Q Analysis*);
- **partitioning analysis**, which can be conducted through different procedures, like *K Means Methods*, *Iterative Reclassification Methods*, *"Sift and Shift" Methods*, *Convergent Methods*;**tandem analysis**, which is realized by combining Principal Components Analysis and a clustering algorithm; the latter is applied to the synthetic scores obtained through the application of the former. In this perspective *Cluster Analysis* can also be combined with *MultiDimensional Scaling* (MDS) (Nardo et al., 2005a, 2005b).This approach could turn out to be difficulty since the identification of the homogeneous group rely on the quality of the synthetic scores previously obtained.
- **factorial k-means analysis**, which is realized by simultaneously combining a discrete clustering model (*partitioning method* like *K Means method*). and a continuous factorial model (Principal Components Analysis) in order to identify the best partition of the objects. In particular, the partition is described by the best orthogonal linear combinations of the variables (factors) according to the least-squares criterion. This approach has great potentiality since it simultaneously allows two objectives to be reached: data reduction and synthesis, simultaneously in direction of both objects and variables. The factorial k-means analysis applies a fast alternating least-squares algorithm that extends its application to large data sets (Nardo et al., 2005a, 2005b).

Each analytical approach produces results that vary according to the decisions made in terms of (i) selected indicators; (ii) measures used in order to evaluate proximities between individual-points; (iii) method used in order to assign an individual-points to a group; (iv) criterion used in order to determine the number of groups; (v) criterion used in order to check the interpretability of the groups.

Each typology will be considered in the context of the successive higher-level analysis in terms of

- categorical information to which other information can be associated, like the dimension of the group,
- simple descriptive statistics, univariate (mean, median) or multivariate (centroid).

4.4 An example

The particular application illustrated here is aimed at illustrating and exemplifying a possible approach to reduce the complexity of data structure, by using subjective and objective data provided by the European Social Survey project)[4] and the Joint Research Centre (JRC – European Commission), respectively.

First stage: construction of synthetic indicators at individual level

The goal of this stage is to create synthetic subjective indicators through the aggregation of elementary indicators. The aggregation procedure should be consistent to the adopted model of measurement, that is:

- reflective approach: in this case the aggregation procedure requires an approach aimed to confirm the hypothesis concerning the relationship between latent variables and elementary indicators; in case of subjective indicators, scaling models can generally represent valid approaches. The scaling model has to be chosen consistently with the assumed dimensionality, the nature of observed data (preferences, similarities, and so on), the adopted scaling technique (comparative or non-comparative).
- formative approach: in this case the aggregation procedure requires a different approach like the one aimed at composite indicators construction (Nardo et al., 2005a and 2005b).

From the European Social Survey data, some variables have been identified:

European Social Survey – wave 1 (2002)					
Area	Variable	Items	Item number	Scaling technique	Model of measurement
Politics	Trust in	country's parliament	B7	0 (no trust at all) – 10 (complete trust)	reflective
		the legal system	B8		
		the police	B9		
		politicians	B10		
		the European Parliament	B11		
		the United Nations	B12		
	Self-placement	placement on left-right scale	B28	0 (left) – 10 (right)	
	How satisfied with	present state of economy in country	B30	0 (extremely dissatisfied) – 10 (extremely satisfied)	reflective
		the national government	B31		
		the way democracy works in country	B32		
		state of education in country nowadays	B33		
		state of health services in country nowadays	B34		
Subjective aspects	Happiness	how happy are you	C1	0 (extremely unhappy) – 10 (extremely happy)	
	Life satisfaction	how satisfied with life as a whole	B29	0 (extremely dissatisfied) – 10 (extremely satisfied)	
	Values: important in life	family	E13	0 (extremely unimportant) – 10 (extremely important)	formative
		friends	E14		
		leisure time	E15		
		politics	E16		
		work	E17		
		religion	E18		
		voluntary organizations	E19		
Immigration and asylum issues	Acceptance of immigration: allow	many/few immigrants of same race/ethnic group as majority	D4	1. allow many 2. allow some 3. allow a few 4. allow none to come and live here	reflective
		many/few immigrants of different race/ethnic group from majority	D5		
		many/few immigrants from richer countries in Europe	D6		
		many/few immigrants from poorer countries in Europe	D7		
		many/few immigrants from richer countries outside Europe	D8		
		many/few immigrants from poorer countries outside Europe	D9		
Socio-demographic profile	Income	feeling about household's income nowadays	F31	1. living comfortably 2. coping 3. difficult 4. very difficult on present income	

[4] For any further information on European Social Survey project, please refer to http://www.europeansocialsurvey.org/ where data and documentation can be found.

Items referring to each variable were submitted to analysis in order to verify the dimensionality. Afterwards, in case of:

- *uni-dimensional latent variable*, the items aggregation was performed through a simple additive technique,
- *multi-dimensional latent variable*: the items aggregation was performed through principal component analysis that allowed us to obtain scores showing normal-standardized distributions.

Reflective approach: aggregation accomplished by testing multi-dimensional hypothesis

Variable	Items	Item number	Item loading	Factor/dimension	Variance explained (%)	Aggregated score
Trust in	the legal system	B8	0.5	national security	31	TRUST_NS
	the police	B9	1.0			
	the European Parliament	B11	0.8	international institutions	33	TRUST_II
	the United Nations	B12	0.5			
	country's parliament	B7	0.7	national politics	36	TRUST_NP
	politicians	B10	0.7			
How satisfied with	present state of economy in country	B30	0.5	satisfaction for national foundations	41	SAT_NF
	the national government	B31	0.7			
	the way democracy works in country	B32	0.5			
	state of education in country nowadays	B33	0.5	satisfaction for national social services	31	SAT_NSS
	state of health services in country nowadays	B34	0.5			

Reflective approach: aggregation accomplished by testing unidimensional hypothesis

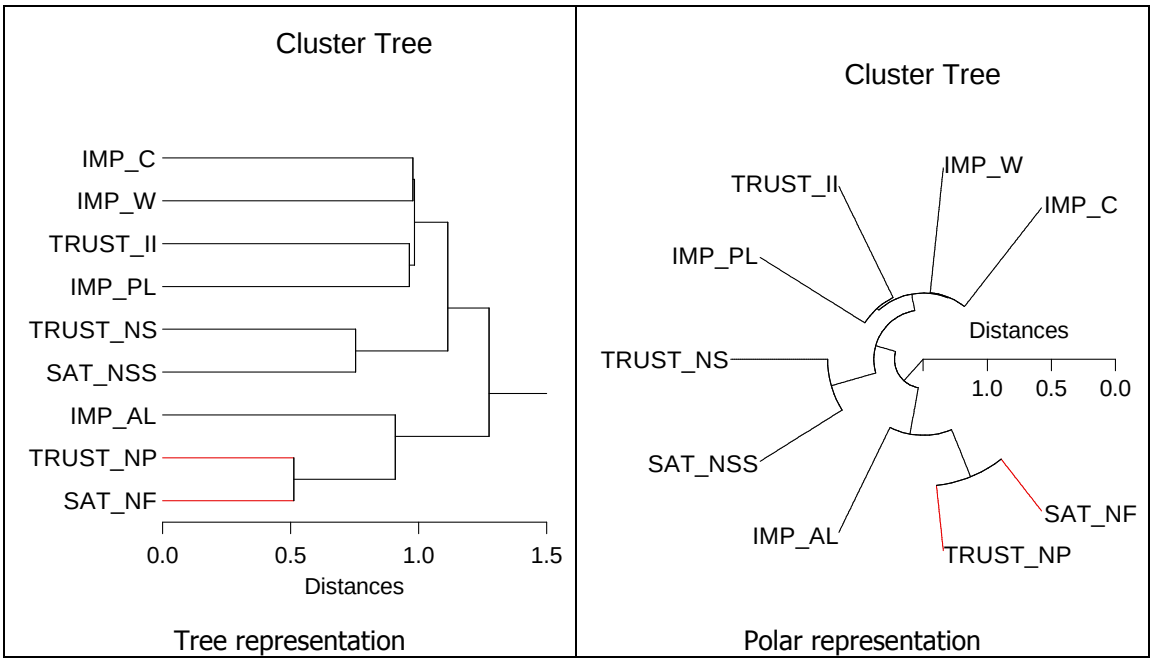
Variable	Items	Item number	Unidimensional model	Aggregated score
Acceptance of immigration: allow	many/few immigrants of same race/ethnic group as majority	D4	aggregation though additive technique	IMMIGR
	many/few immigrants of different race/ethnic group from majority	D5		
	many/few immigrants from richer countries in Europe	D6		
	many/few immigrants from poorer countries in Europe	D7		
	many/few immigrants from richer countries outside Europe	D8		
	many/few immigrants from poorer countries outside Europe	D9		

Formative approach: aggregation accomplished through Principal Component Analysis

Variable	Items	Item number	Item loading	Component	Variance explained (%)	Aggregated score
Values: important in life	family	E13	0.6	Private life dimension	23	IMP_PL
	friends	E14	0.8			
	leisure time	E15	0.7			
	politics	E16	0.8	Active life dimension	18	IMP_AL
	voluntary organizations	E19	0.6			
	family	E13	0.5	Caring dimension	18	IMP_C
	religion	E18	0.9			
	voluntary organizations	E19	0.5			
	work	E17	1.0	Work dimension	15	IMP_W

Ten synthetic indicators were computed and then submitted to a successive level of aggregation, according to the formative approach, in order to obtain a small group of meaningful and interpretable synthetic indicators. This aggregation was obtained through Principal Component Analysis and Hierarchical Cluster Analysis (linkage method: Ward; distance technique: Pearson). The outcomes obtained by the two methods turned out to be identical and show the same four dimensions, each one composed by indicators referring to trust, importance and satisfaction characteristics. A particular result has to be noticed: "importance for private life" indicator obtained significant loadings in two components in Principal Component Analysis.

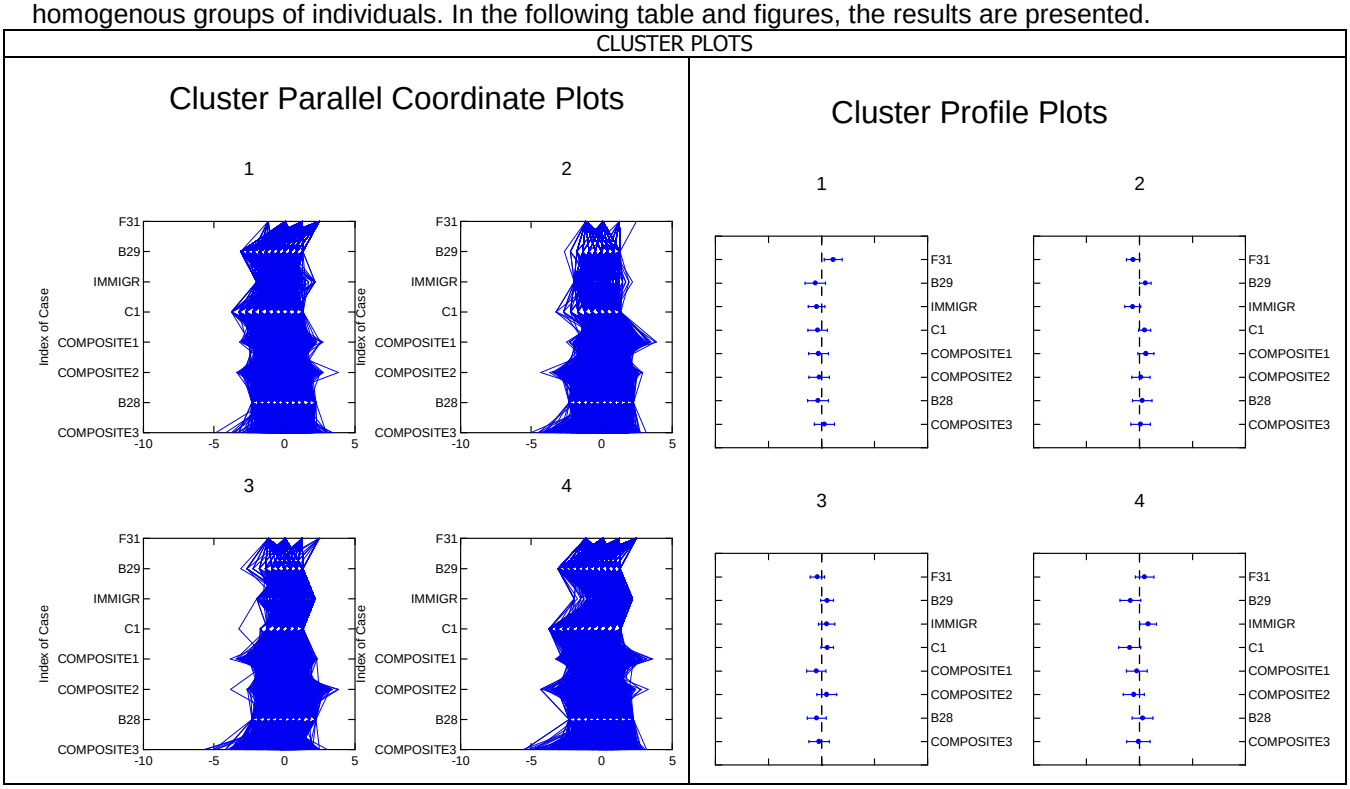
Synthetic indicators		Item loading	Obtained component	Variance explained (%)	Aggregated score
National politics	TRUST_NP	0.8	Public & political life	18	COMPOSITE1
Active life dimension	IMP_AL	0.6			
Satisfaction for national foundations	SAT_NF	0.8			
national security	TRUST_NS	0.8	Welfare dimension	15	COMPOSITE2
Private life dimension	IMP_PL	0.4			
Satisfaction for national social services	SAT_NSS	0.7			
Caring dimension	IMP_C	0.4	Personal life principles	12	COMPOSITE3
International institutions	TRUST_II	0.6			
Private life dimension	IMP_PL	0.4			
Work dimension	IMP_W	0.6			



Composite scores were calculated by means of Principal Component Analysis according to the observed results. At this stage the aggregation process has concerned also objective indicators (construction of composite indicators through formative criterion).

Second stage: definition of macro-units

At this stage, a partitioning analysis were conducted (*K means method*) in order to explore the existence of homogenous groups of individuals. In the following table and figures, the results are presented.

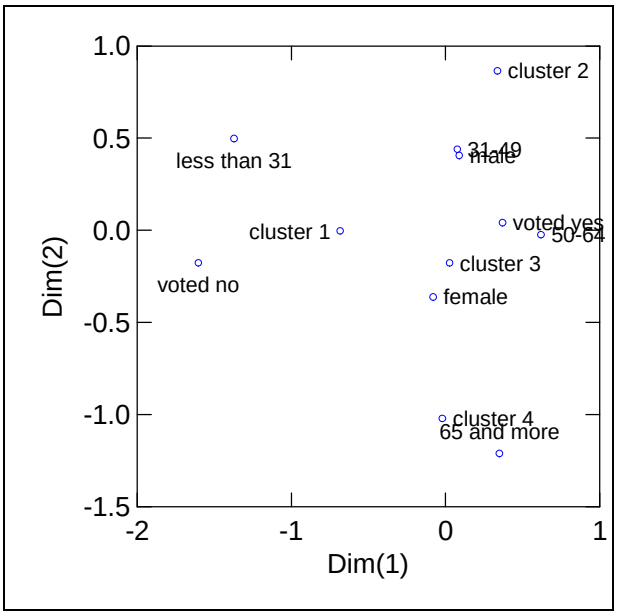


		INDICATOR	min.	mean	max.	SD
CLUSTER 1 (n=7369)	B29	life satisfaction	-3.10	-0.58	1.31	0.97
	C1	happiness	-3.74	-0.37	1.34	0.93
	F31	Feeling about household's income nowadays	-1.14	1.10	2.46	0.85
	B28	self-placement on left-right scale	-2.30	-0.34	2.24	0.98
	IMMIGR	Non-acceptance of immigration	-1.96	-0.47	2.17	0.79
	COMPOSITE1	Public & political life	-3.19	-0.29	3.13	0.95
	COMPOSITE2	Welfare dimension	-3.88	-0.22	3.83	0.98
	COMPOSITE3	Personal life principles	-4.86	0.27	3.44	0.97
CLUSTER 2 (n=14855)	B29	life satisfaction	-3.10	0.54	1.31	0.54
	C1	happiness	-3.74	0.48	1.34	0.59
	F31	Feeling about household's income nowadays	-1.14	-0.61	2.46	0.63
	B28	self-placement on left-right scale	-2.30	0.26	2.24	0.92
	IMMIGR	Non-acceptance of immigration	-1.96	-0.64	2.17	0.76
	COMPOSITE1	Public & political life	-2.50	0.60	4.08	0.76
	COMPOSITE2	Welfare dimension	-4.32	0.12	2.90	0.86
	COMPOSITE3	Personal life principles	-5.03	0.10	3.15	0.91
CLUSTER 3 (n=9703)	B29	life satisfaction	-3.10	0.53	1.31	0.60
	C1	happiness	-3.23	0.54	1.34	0.58
	F31	Feeling about household's income nowadays	-1.14	-0.40	2.46	0.68
	B28	self-placement on left-right scale	-2.30	-0.46	2.24	0.90
	IMMIGR	Non-acceptance of immigration	-1.96	0.48	2.17	0.78
	COMPOSITE1	Public & political life	-3.85	-0.49	2.36	0.90
	COMPOSITE2	Welfare dimension	-3.83	0.48	3.85	0.94
	COMPOSITE3	Personal life principles	-5.71	-0.24	3.07	0.97
CLUSTER 4 (n=10418)	B29	life satisfaction	-3.10	-0.86	1.31	1.00
	C1	happiness	-3.74	-0.93	1.34	1.04
	F31	Feeling about household's income nowadays	-1.14	0.47	2.46	0.89
	B28	self-placement on left-right scale	-2.30	0.30	2.24	0.99
	IMMIGR	Non-acceptance of immigration	-1.96	0.81	2.17	0.79
	COMPOSITE1	Public & political life	-3.47	-0.26	3.61	0.99
	COMPOSITE2	Welfare dimension	-4.34	-0.54	3.29	0.99
	COMPOSITE3	Personal life principles	-5.59	-0.11	3.22	1.11

The obtained clusters have shown quite differentiated profiles. In the following table a possible synthetic description of each cluster is described. Cluster 1 and cluster 4 seem to be the group with problematical profiles. Cluster 1 and cluster 4 seem to be the groups with problematical profiles. In particular, cluster 4 seems to be composed by individual with low level of well-being and trust and importance in society dimensions, high level of non-acceptance of immigration and low, and a clear self-placement on left-right political scale.

		CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4
B29	life satisfaction	Medium-low	Medium-high	Medium-high	low
C1	happiness	Medium-low	Medium-high	High	low
F31	Feeling about household's income nowadays	many difficulties	Very comfortable	comfortable	Some difficulties
B28	self-placement on left-right scale	Centre-left	Centre-right	Left	Right
IMMIGR	Non-acceptance of immigration	Medium-low	Low	Medium-high	High
COMPOSITE1	Public & political life	Medium-low	High	Low	Medium-low
COMPOSITE2	Welfare dimension	Medium-low	Medium-high	High	Low
COMPOSITE3	Personal life principles	High	Medium-high	Low	Medium-low

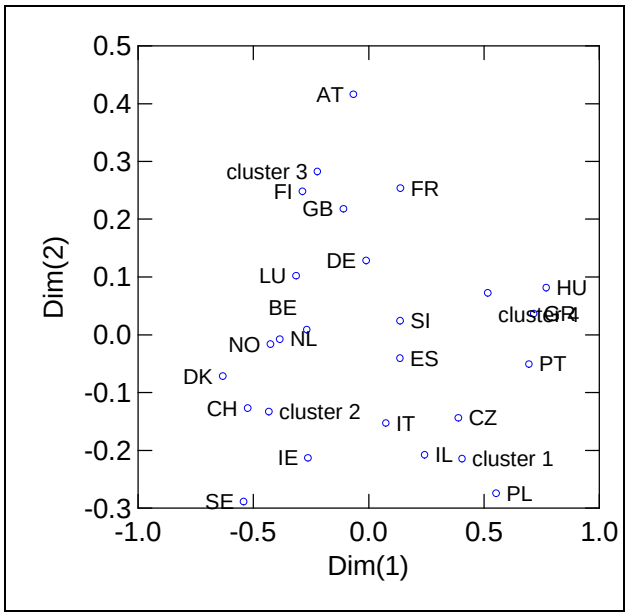
The conceptual framework should point out the individual objective characteristics to be integrated with the subjective ones (synthesized in clusters definition) at micro level. This level of integration is aimed at exploring and understanding subjective responses in terms of individual characteristics.
In this application we have chosen gender, age, and individual position with reference to vote in last political election. These indicators were submitted to correspondence analysis together with the cluster indicator. The analysis, performed on more than 38 thousands respondents with almost 30% of total inertia explained, produced a configuration (see following figure) in which the more frequent profiles can be identified. For example, cluster 1 is more frequent among young individuals who did not vote, while cluster 4 is more frequent among elderly persons.



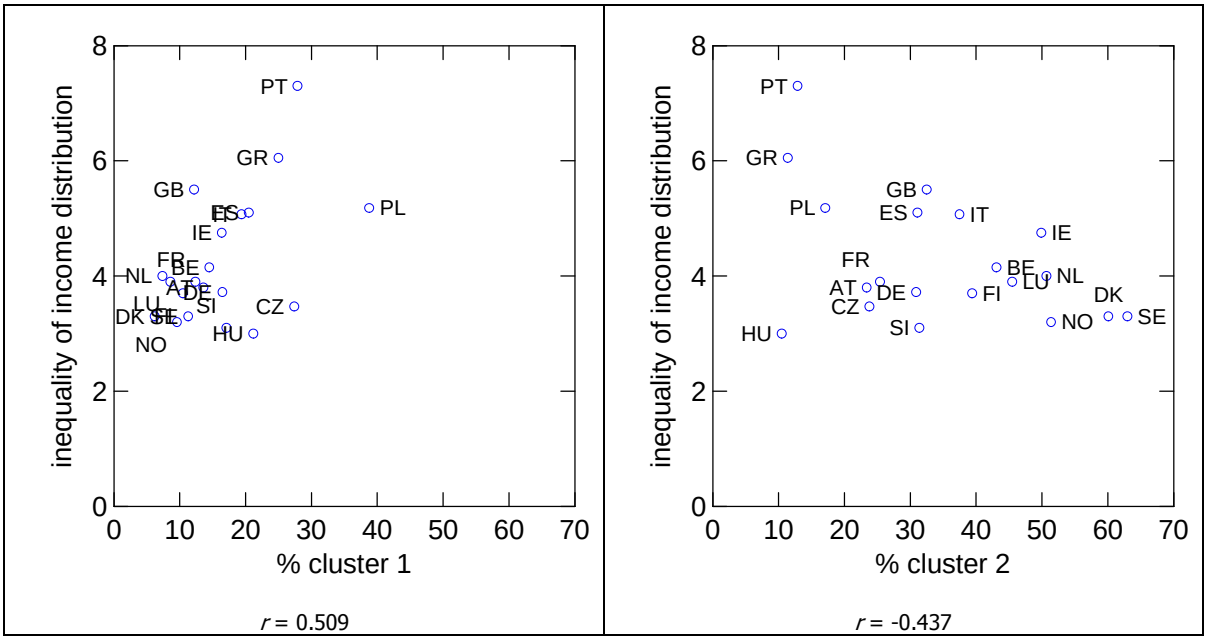
The clusters obtained through the previous stage were considered aggregations of subjective information (homogeneity criterion). In the following table the incidence of each cluster for each country can be observed.

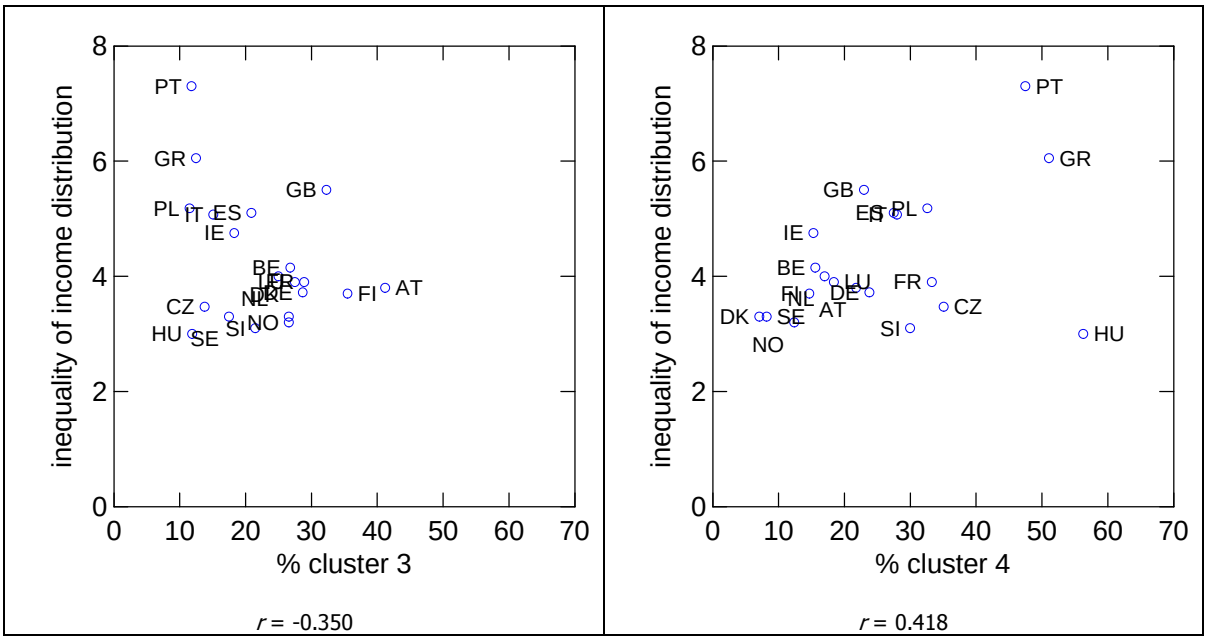
		cluster				Total	N
		1	2	3	4		
AT	Austria	13.6	23.4	41.2	21.8	100.0	2257
BE	Belgium	14.5	43.1	26.8	15.6	100.0	1897
CH	Switzerland	10.9	57.5	22.9	8.8	100.0	2040
CZ	Czech Rep.	27.4	23.8	13.8	35.1	100.0	1360
DE	Germany	16.5	30.9	28.7	23.8	100.0	2919
DK	Denmark	6.2	60.1	26.6	7.1	100.0	1500
ES	Spain	20.5	31.1	20.9	27.5	100.0	1728
FI	Finland	10.5	39.4	35.5	14.7	100.0	2000
FR	France	12.4	25.4	28.9	33.3	100.0	1503
GB	United Kingdom	12.2	32.5	32.3	23.0	100.0	2051
GR	Greece	25.0	11.4	12.5	51.1	100.0	2566
HU	Hungary	21.2	10.5	11.9	56.3	100.0	1685
IE	Ireland	16.4	49.9	18.3	15.3	100.0	2046
IL	Israel	32.6	26.1	19.0	22.3	100.0	2497
IT	Italy	19.4	37.5	15.1	28.0	100.0	1206
LU	Luxembourg	8.6	45.5	27.5	18.4	100.0	1552
NL	Netherlands	7.4	50.7	25.0	17.0	100.0	2364
NO	Norway	9.6	51.4	26.6	12.4	100.0	2036
PL	Poland	38.8	17.1	11.5	32.6	100.0	2109
PT	Portugal	27.9	12.9	11.8	47.5	100.0	1511
SE	Sweden	11.3	63.0	17.5	8.2	100.0	1999
SI	Slovenia	17.1	31.4	21.5	30.0	100.0	1519
Total		17.4	35.1	22.9	24.6	100.0	
N		7369	14855	9703	10418		42345

After that, correspondence analysis was performed by considering different indicators and applying a particular causal model (cluster=country). In the following figure the four clusters turn out to be more frequent with reference to different country. For example, cluster 1 is more frequent in Poland, Israel, and Czech samples.



At this stage the information of the incidence of each cluster for each country was used and related to objective indicators measured at macro level. In the following figures (in which *x* scales show the same range in order to preserve comparability between scatterplots) the national incidences are related to *inequality of income distribution* of each country.





The results show a clear relationship between clusters incidences and the objective indicator measured at country level especially with reference to cluster 1, which represents the more problematic among the four observed clusters.

- . - . - . - . -

The main goal of this example is to illustrate the composite approach through which the complexity of data structure can be managed. The approach is carried out through subsequent stages. In each stage different analytical solutions can be found. The soundness of the approach and of its results relies on the defined and adopted conceptual framework assuming the correct perspective to be identified according to different objectives, (i) the aggregation of indicators and units, (ii) the integration of objective and subjective information, and (iii) the levels at which the previous objectives have to be pursued.
The illustrated application, which was made possible by the contribution of the Econometrics and Applied Statistics Unit (EAS) at the Joint Research Centre of the European Commission, has the restricted goal to illustrate and exemplify the proposed approach.
The intention is that to deeply explore the methodology in order to provide further approaches, especially in longitudinal perspective.

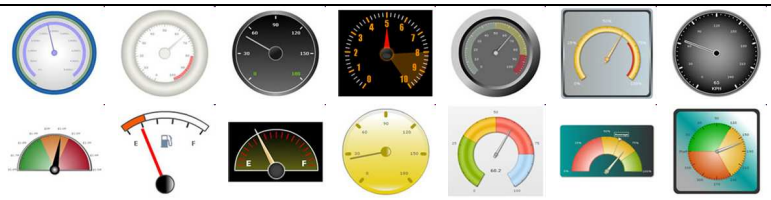
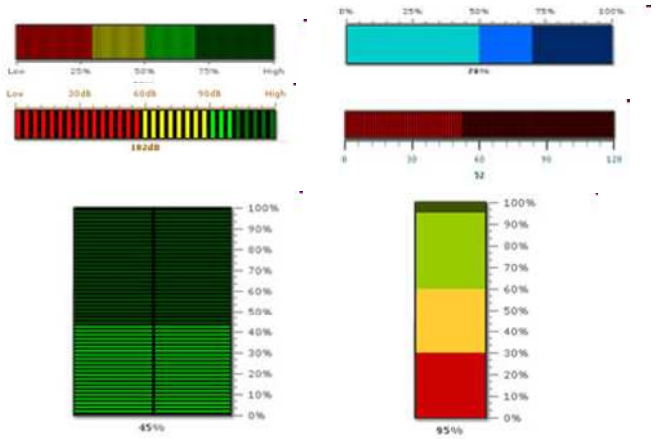
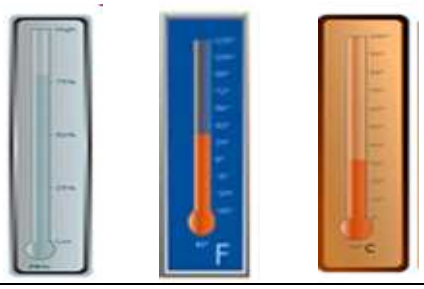
5. Combining indicators

5.1 Dashboards

Dashboards represent useful tools aimed at simultaneously representing, comparing and interpreting indicators' values

- through an analogical perspective
- by setting them on a standardized scale
- by representing them on a colour scale (e.g., a green-to-red colour scale).

Several software programmes (free or not) can be used in order to carry out the graphical representation through different images:

Car dashboards	
Analogical bars	
Digital bars	
Thermometers	

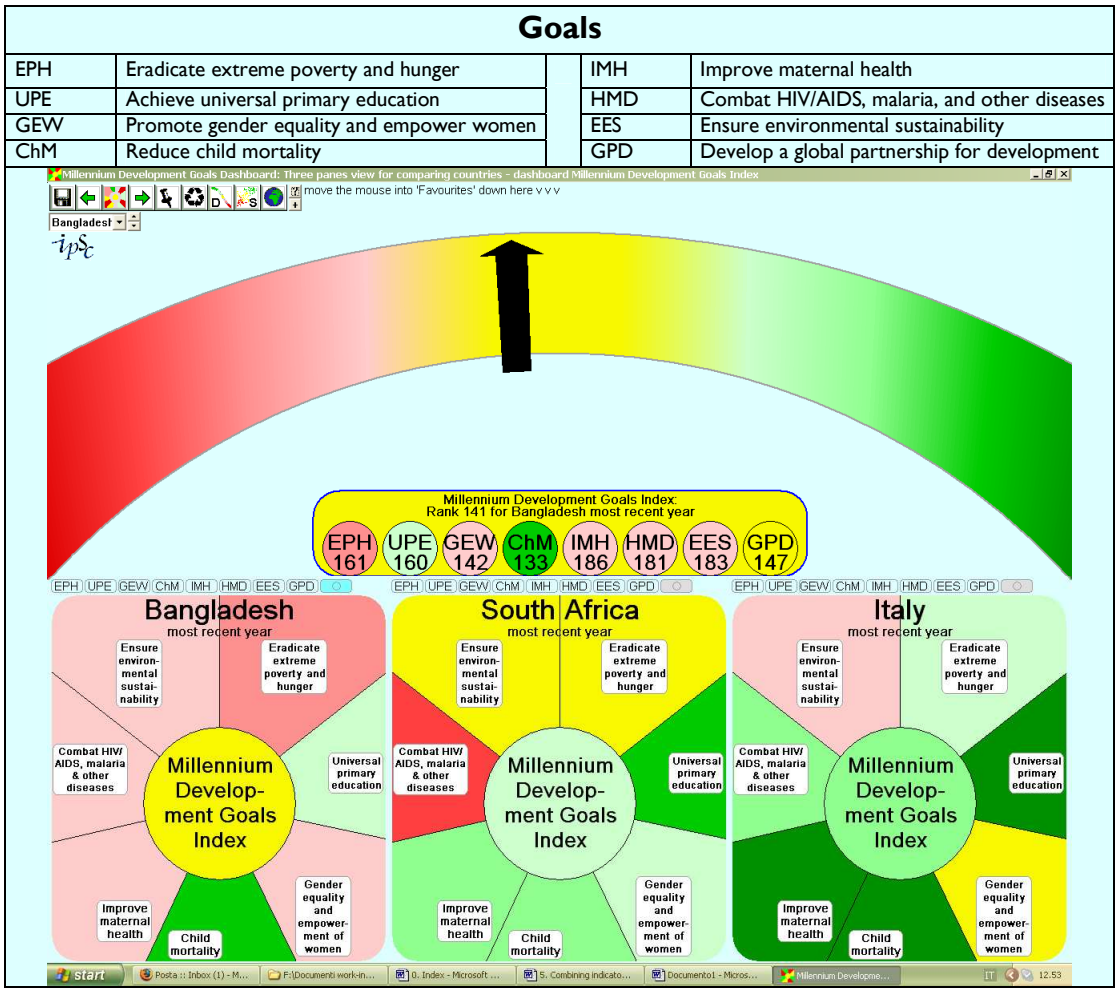
Whichever representation form is adopted, indicators' values are displayed through

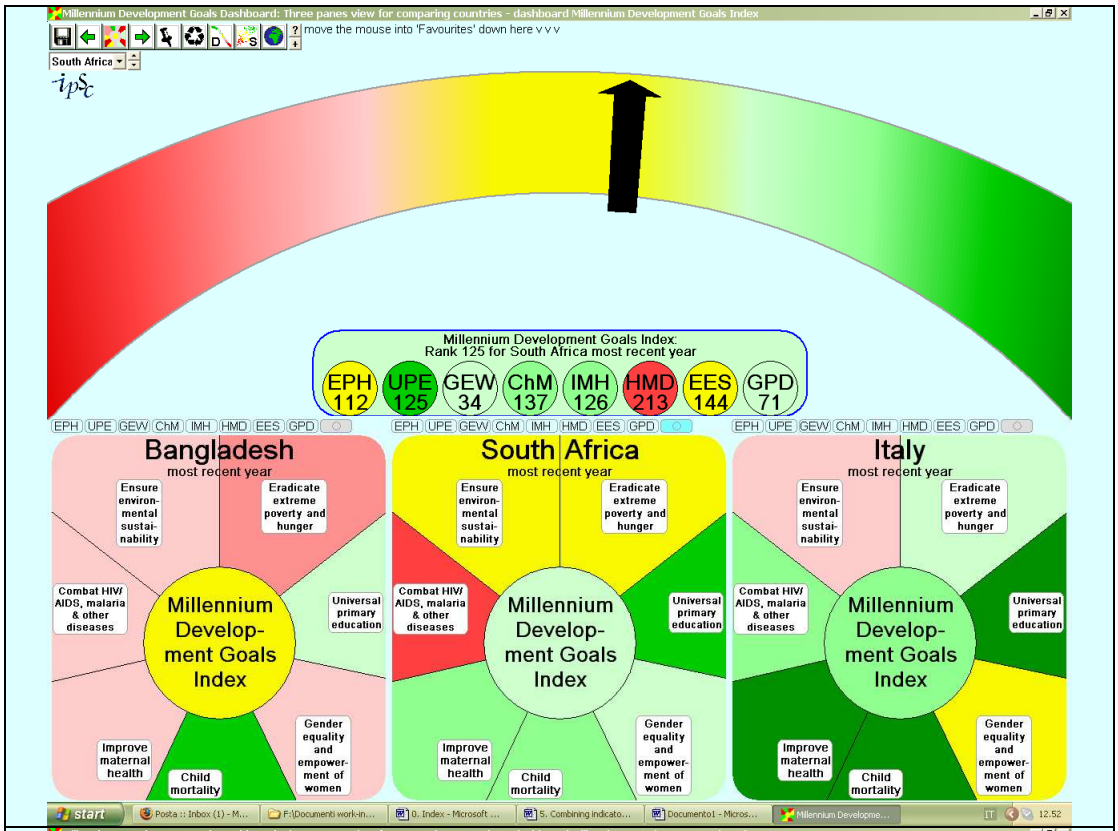
- separated values (values are not aggregated), allowing weak and strong points to be identified,
- colours, allowing the analysis of relative performance (value to be displayed relatively to an expected value or a given level / targets)
- distributions, allowing assessment indicators' meaningfulness, outliers identification, etc.
- scatterplot graph, allowing simple correlation analysis between the indicators to be visualized. This function allows synergies (indicators whose "desirable" values are positively correlated) and potential conflicts (e.g. environment vs. many economic and social variables) to be identified.

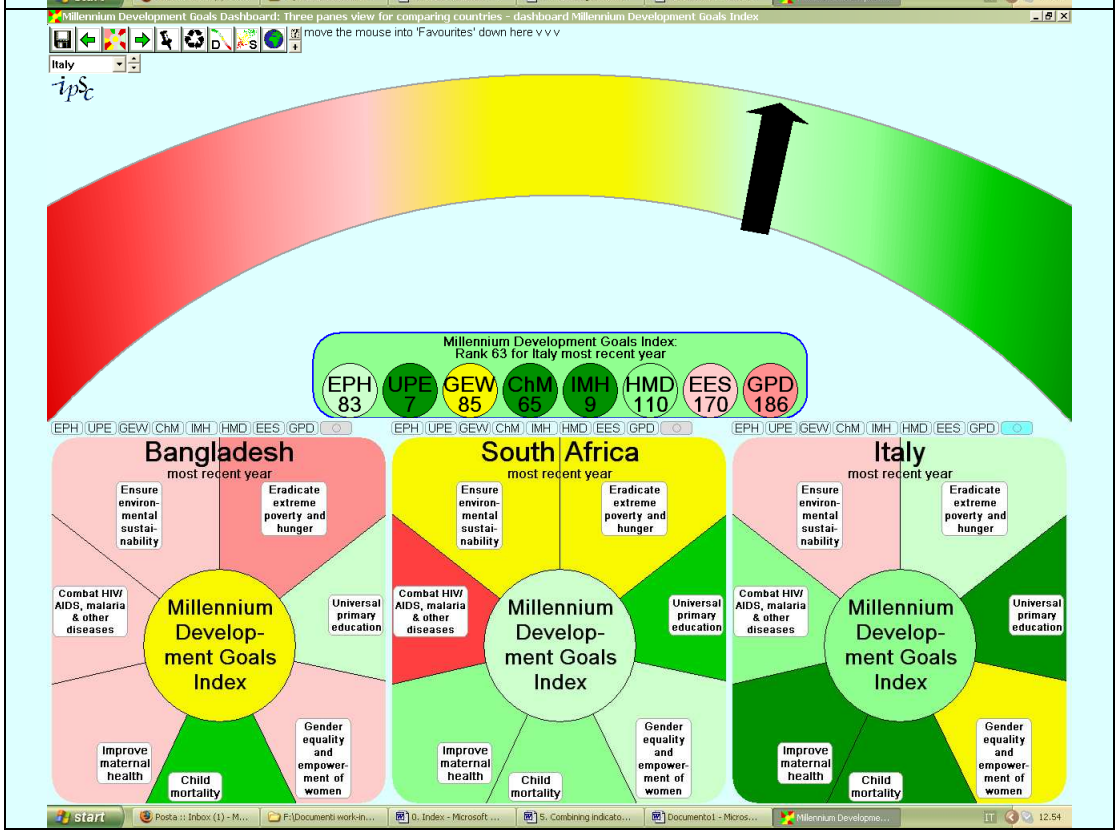
Through the graphical display, dashboards allow comprehensive monitoring and evaluation of programmes, performances or policies, since

- highly complex systems of indicators can be represented by taking into account the hierarchical design,
- easy communications are possible through a catchy and simple graphical representation,
- indicators can be related to weights interpreted in terms of
    a. *importance* (reflected by the size of the segments) and
    b. *performance result* (reflected by the colour, interpretable in terms of "good vs. bad")
- performances of different cases can be compared.

In the perspective of national policy evaluation, the following example (from Joint Research Centre – European Commission – http://esl.jrc.ec.europa.eu/envind/dashbrds.htm) shows and compares three countries' indicators related to the UN Millennium Development Goals.
The colours help in identifying the reached goals (*green*: completely reached – *red*: not reached at all) and the side-by-side arrangement allows different countries' values to be easily compared. Each medium bar refers the rank position for each country rank for each goal (respectively, in the example, Bangladesh, South Africa and Italy):







Below, the hierarchical design of the UN Millennium Development Goals is shown:

Goal (→ area)	**Target** (→ variable)	**Statistical index** (→ elementary indicators)
1 **Eradicate extreme poverty and hunger (EPH)**	A Halve the proportion of people living on less than $1 a day	• Proportion of population below $1 per day (PPP values) • Poverty gap ratio [incidence x depth of poverty] • Share of poorest quintile in national consumption
	B Achieve Employment for Women, Men, and Young People	• GDP Growth per Employed Person • Employment Rate • Proportion of employed population below $1 per day (PPP values) • Proportion of family-based workers in employed population
	C Halve the proportion of people who suffer from hunger	• Prevalence of underweight children under five years of age • Proportion of population below minimum level of dietary energy consumption
2 **Achieve universal primary education (UPE)**	A By 2015, all children can complete a full course of primary schooling, girls and boys	• Enrollment in primary education • Completion of primary education • Literacy of 15-24 year olds, female and male
3 **Promote gender equality and empower women (GEW)**	A Eliminate gender disparity in primary and secondary education preferably by 2005, and at all levels by 2015	• Ratios of girls to boys in primary, secondary and tertiary education • Share of women in wage employment in the non-agricultural sector • Proportion of seats held by women in national parliament
4 **Reduce child mortality (ChM)**	A Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	• Under-five mortality rate • Infant (under 1) mortality rate • Proportion of 1-year-old children immunised against measles
5 **Improve maternal health (IMH)**	A Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio	• Maternal mortality ratio • Proportion of births attended by skilled health personnel
	B Achieve, by 2015, universal access to reproductive health	• Contraceptive prevalence rate • Adolescent birth rate • Antenatal care coverage (at least one visit and at least four visits) • Unmet need for family planning
6 **Combat HIV/AIDS, malaria, and other diseases (HMD)**	A Have halted by 2015 and begun to reverse the spread of HIV/AIDS	• HIV prevalence among population aged 15–24 years • Condom use at last high-risk sex • Proportion of population aged 15–24 years with comprehensive correct knowledge of HIV/AIDS • Ratio of school attendance of orphans to school attendance of non-orphans aged 10–14 years
	B Achieve, by 2010, universal access to treatment for HIV/AIDS for all those who need it	• Proportion of population with advanced HIV infection with access to antiretroviral drugs
	C Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	• Prevalence and death rates associated with malaria • Proportion of children under 5 sleeping under insecticide-treated bednets • Proportion of children under 5 with fever who are treated with appropriate anti-malarial drugs • Prevalence and death rates associated with tuberculosis • Proportion of tuberculosis cases detected and cured under DOTS (Directly Observed Treatment Short Course)

7	**Ensure environmental sustainability (EES)**	A	Integrate the principles of sustainable development into country policies and programmes; reverse loss of environmental resources	
		B	Reduce biodiversity loss, achieving, by 2010, a significant reduction in the rate of loss	• Proportion of land area covered by forest • CO_2 emissions, total, per capita and per $1 GDP (PPP) • Consumption of ozone-depleting substances • Proportion of fish stocks within safe biological limits • Proportion of total water resources used • Proportion of terrestrial and marine areas protected • Proportion of species threatened with extinction
		C	Halve, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation (for more information see the entry on water supply)	• Proportion of population with sustainable access to an improved water source, urban and rural • Proportion of urban population with access to improved sanitation
		D	By 2020, to have achieved a significant improvement in the lives of at least 100 million slum-dwellers	• Proportion of urban population living in slums
8	**Develop a global partnership for development (GPD)**	A	Develop further an open, rule-based, predictable, non-discriminatory trading and financial system	• Includes a commitment to good governance, development, and poverty reduction – both nationally and internationally
		B	Address the Special Needs of the Least Developed Countries (LDC)	• Includes: tariff and quota free access for LDC exports; enhanced programme of debt relief for HIPC and cancellation of official bilateral debt; and more generous ODA (Overseas Development Assistance) for countries committed to poverty reduction
		C	Address the special needs of landlocked developing countries and small island developing States	• Through the Programme of Action for the Sustainable Development of Small Island Developing States and the outcome of the twenty-second special session of the General Assembly
		Official development assistance (ODA)		
		D	Deal comprehensively with the debt problems of developing countries through national and international measures in order to make debt sustainable in the long term	• Net ODA, total and to LDCs, as percentage of OECD/DAC donors' GNI • Proportion of total sector-allocable ODA of OECD/DAC donors to basic social services (basic education, primary health care, nutrition, safe water and sanitation) • Proportion of bilateral ODA of OECD/DAC donors that is untied • ODA received in landlocked countries as proportion of their GNIs • ODA received in small island developing States as proportion of their GNIs
			Market access	• Proportion of total developed country imports (by value and excluding arms) from developing countries and from LDCs, admitted free of duty • Average tariffs imposed by developed countries on agricultural products and textiles and clothing from developing countries • Agricultural support estimate for OECD countries as percentage of their GDP • Proportion of ODA provided to help build trade capacity
			Debt sustainability	• Total number of countries that have reached their HIPC decision points and number that have reached their HIPC completion points (cumulative) • Debt relief committed under HIPC initiative, US$ • Debt service as a percentage of exports of goods and services
		E	In co-operation with pharmaceutical companies, provide access to affordable, essential drugs in developing countries	• Proportion of population with access to affordable essential drugs on a sustainable basis
		F	In co-operation with the private sector, make available the benefits of new technologies, especially information and communications	• Telephone lines and cellular subscribers per 100 population • Personal computers in use per 100 population • Internet users per 100 Population

Of course, dashboard does not allow complex analysis concerning relationships between indicators and comparisons of performance over time (trends) or across units (inter-cases comparisons).
Dashboards can be useful in view of creating composite indicators.

5.2 Composite indicators

A composite indicator synthesizes a number of values expressed by the indicators that compound it (Booysen, 2002; Nardo et al., 2005a; Sharpe & Salzman, 2004) and re-establishing the unity of the concept described in the hierarchical design. The aggregating process allows to obtain not a faithful description of the reality, but an "indication" that will be more or less accurate, meaningful, and interpretable depending on the defined hierarchical design and the applied methodology. In other words, the composite indicators are aimed at describing synthetically a reality, which is and remains complex.

5.2.1 Methodological issues

The methodology aimed to construct composite indicators requires specific techniques aimed at
1. verifying the dimensionality of selected elementary indicators (*dimensional analysis*)
2. defining the importance of each elementary indicator to be aggregated (*weighting criteria*)
3. identifying the technique for aggregating the elementary indicators values into the composite indicator (*aggregating-over-indicators techniques*)
4. assessing the robustness of the composite indicator in terms of capacity to produce correct and stable measures (*uncertainty analysis, sensitivity analysis*)
5. assessing the discriminant capacity of the composite indicator (*ascertainment of selectivity* and *identification of cut-point or cut-off values*)

1. Dimensional analysis

In order to proceed with the construction of complex indicators, the hypothesized **dimensionality underlying the elementary indicators** to be aggregated has to be investigated. The *dimensional analysis* of the indicators, which have to be involved in the aggregation, allows the insight of the dimensional structure underlying the considered elementary indicators, and consequently allows the insight of the level of complexity by which the complex indicator has to be constructed (as previously said).
In particular, the dimensional analysis allows the identification of the number of complex indicators to be calculated by avoiding superimposition and redundancies among elementary indicators. Different approaches can be used for this purpose (Alt, 1990; Anderson, 1958; Bolasco, 1999; Cooley & Lohnes, 1971; Corbetta, 1992, 2003; Cox & Cox, 1994; Hair et al., 1998; Kruskal & Wish, 1978; Maggino, 2004a, 2004b, 2005a; Sadocchi, 1981). The approaches can also be combined. Indirectly the dimensional analysis also allows us to check the validity of the selected elementary indicators in measuring the latent variable.
- *Correlation Analysis*. It allows the possible presence of particular structure underlying the elementary indicators to be evaluated; it is also useful when the aim is to synthesise indicators that are not redundant; in these cases, aggregating indicators that are highly correlated can introduce a *bias* in the composite indicator produced by the multicollinearity (*double counting*). The latent variable could be represented only by one indicator in the extreme case of perfect collinearity between indicators (Nardo et al., 2005a).
- *Principal Component Analysis*. The goal of principal component analysis is essentially to uncover variations in a data set. Principal component analysis can be used to describe the variation of a data set using a number of scores that is smaller than the number of the original elementary indicators. This approach is very often applied to test dimensional structures. This is done following the idea that this approach can be assimilated to Factor Analysis. The two approaches are actually, however, very different from each other. In particular, the main goal of Principal Component Analysis is not to test a (dimensional) model but simply to decompose the correlations among indicators in order to condense the variance among all the indicators as much as possible by calculating new linear variables, defined components. For this reason, the approach can be used also for defining weights).
- *MultiDimensional Scaling*. It allows the underlying dimensionality to be tested and for the creation of a geometrical multidimensional representation (*map*) of the complete group of elementary indicators (Cox & Cox, 1994; Kruskal & Wish, 1978; Torgerson, 1958).

- *Cluster Analysis*. It can in this context be useful to identify meaningful groupings among elementary indicators (Aldenderfer & Blashfield, 1984; Bailey, 1994; Corter, 1996; Hair et al., 1998; Lis & Sambin, 1977; Maggino, 2005a).

In some cases, methods related to reflective model of measurement can be carefully used. Among these we can recall the following:

- *Item Response Theory*. In the case of elementary indicators related to performance variables, a particular analysis, derived directly from the application of the *Item Response Theory* (related to the reflective model of measurement), allows the indicators that better discriminate among units to be selected. In particular, the identified indicators can be distinguished from each other in terms of difficulty and discriminant capacity (Andersen, 1972, 1973; Andrich, 1988; Bock & Aitkin, 1981; Hambleton et al., 1991; Lord, 1974, 1984; Ludlow & Haley, 1995; McDonald, 1989; Rasch, 1960; Sijtsma & Molenaar, 2002; Swaminathan & Gifford, 1982, 1985, 1986).
- *Factor Analysis*. It allows the hypothesized dimensional structure underlying the group of elementary indicators (latent structure analysis) to be tested; the total variance of each indicator is produced by a linear combination of different variance components (additive assumption), that is, common variance (due to the dimensional structure), specific variance (due to the specificity variance of each indicator), and error. The analysis allows the common variance (*communality*) to be estimated (Kim Jae-On, 1989a, 1989b; Marradi, 1981).

In some cases, combined approaches can be used, like ***tandem analysis*** or ***factorial k-means analysis*** (Nardo et al., 2005a).

This analysis aims at selecting the indicators to be included in the composite, showing the best statistical characteristics. However, in selecting the indicators also other criteria should be taken into account:

- *Redundancy*: 2 indicators showing very high correlation can be considered redundant, in building composite indicators, it is recommended to select only one.
- *Political impact*: 2 indicators are highly correlated and convey strong political messages, they can be both included in the final list.
- *Availability*: indicators which prove to be available for a large number of cases are preferable.
- *First comer privilege*: when two indicators are redundant, it is recommended to select the one allowing trend analysis and wide comparisons.

2. Weighting criteria

Since not necessarily all the identified elementary indicators contribute with the same importance to the measurement and evaluation of the latent variable, a weighting system needs to be defined in order to assign a weight to each elementary indicator, before proceeding to the elementary indicators aggregation (Ghiselli, 1964).

When an implicit weighting system can not be identified, a criterion has to be adopted in order to define a weighting system, which can reproduce as accurately as possible the contribution of each elementary indicator to the construction of the composite indicator. In this perspective, the definition of the weighting system can constitute an improvement and refinement of the adopted model of measurement.

From the technical point of view, the weighting procedure consists in defining and assigning a weight to each elementary indicator. The weight will be used in the successive computation of the individual aggregate score; in particular, each weight is multiplied for the corresponding individual value of the elementary indicator.

In order to proceed to the difficult choice among the differential weighting definition approaches, the researcher needs to take into account (Nardo et al., 2005a):

- the defined rationale and theoretical structure on which the latent variable and, consequently, the composite indicator are based
- the meaning and the contribution of each elementary indicator to the aggregation
- the quality of data and the statistical adequacy of the elementary indicators

In this sense, apart from the applied approach, the defined weights represent judgment values.

The researcher has to carefully evaluate and make formally explicit not only the methodology to be adopted but also the results that would have been obtained with other methodologies, also reasonably applicable.

The identification of the procedure for identifying the weights needs to distinguish between **Equal Weighting** (*EW*)[1] and **Differential Weighting** (*DW*). The composite indicator will be strongly influenced by whichever choice is made concerning this. Cases' reciprocal positions can sharply change by simply changing the weights assigned to each indicator.

The adoption of the *differential weighting* procedure do not necessarily corresponds to the identification of different weights but rather to the selection of the most appropriate approach in order to identify the weights among the following (Nardo et al., 2005a).

[1] Equal weighting does not necessarily imply unitary weighting.

1. **statistical methods**:
   a. *Correlation*
   b. *Principal Component Analysis* (PCA)
   c. *Data Envelopment Analysis* (DEA)
   d. *Unobserved Components Models* (UCM).
   The adoption of statistical methods in weighting components of social indices has to be considered carefully since, by removing any control over the weighting procedure from the analysts, it gives a false appearance of mathematical objectivity that is actually difficult to achieve in social measurement (Sharpe & Salzman, 2004).
2. **multi-attribute models**:
   a. *Multi-Attribute Decision Making* (in particular, *Analytic Hierarchy Processes* – AHP) (Yoon & Hwang Ching-Lai, 1995),
   b. *Multi-Attribute Compositional Model* (in particular, *Conjoint Analysis*, CA),[2]
3. **subjective methods**: new perspectives have been introduced recently showing the possibility to involve more individuals (experts or citizens) in the process of defining weighting systems for social indicators. These approaches are defined in the perspective of giving more legitimacy to social indicators by taking into account citizens' importance (values) and not – as usually done in the past – statistical importance.[3]

Assigning differential weights can be just as doubtful, especially when the decision is not supported by:

- theoretical reflections that endow a meaning on each elementary indicator or consider its impact on the synthesis,
- methodological concerns that helps to identify the proper techniques, consistently with the theoretical structure.

In any case, we have to consider that a whole set of weights that is able to express in a perfect way the actual and relative contribution of each elementary indicator to the measurement does not exist.
Independently from the approach adopted in order to define them, the weights can be kept constant or can be changed according to particular considerations concerning each application. In both cases, the researcher needs to rationalize the choice. The former approach can be adopted when the aim is to analyse the evolution of the examined QOL ambit. The latter can be adopted when the aim – for example – concerns the definition of particular priorities.[4]

3. Aggregating-over-indicators techniques

The choice of the aggregating technique must be consistent with the adopted aggregation model. In particular, it has to consider the adopted assumptions concerning the level of complexity of the indicator (*dimensionality*) expressed in terms of homogeneity among indicators to be aggregated, and the relationship between these indicators and the latent variable.
Moreover the choice must take into consideration particular technical characteristics of each technique; in particular, we have to consider if the technique:

a. admits compensability among the elementary indicators to be aggregated,
b. necessitates comparability among elementary indicators,
c. necessitates a homogeneity in the levels of measurement of the elementary indicators.

a. An aggregating technique is <u>*compensatory*</u> when it allows low values in some elementary indicators to be compensated by high values in other elementary indicators. In the typical aggregating table (see below), we can observe all the possible synthetic values, obtainable by aggregating two indicators (A and B) using simple addition (additive technique):

		B		
		1	2	3
A	4	5	6	7
	3	4	5	6
	2	3	4	5
	1	2	3	4

Some of the obtained synthetic values, even if completely identical, are obtained through different original elementary indicators. This means that obtained aggregated values do not allow us to return to original unit's

[2] Hair et al., 1998; Louviere, 1988; Malhotra, 1996. A particular example of Conjoint Analysis application to QOL measurement see Maggino, 2005b.
[3] Hagerty & Land, 2007; Maggino, 2008a, 2008b, 2009; Maggino & Ruviglioni, 2008a, 2008b, 2009). See
[4] With reference to this issue, see the second contribution in "to go deeper" section: *Some technical issues in building composite indicators*.

profile since same synthetic values are obtained through different combinations of scores. In other words, two units with different realities turn out to be identical and not distinguishable from each other.
By using the same data reported in previous table, all the possible synthetic values can be observed, obtainable by aggregating two indicators (A and B) using the multiplicative techniques (following the geometrical approach):

		B		
		1	2	3
A	4	4	8	12
	3	3	6	9
	2	2	4	6
	1	1	2	3

The table suggests that the multiplicative technique is compensatory as well, especially with reference to indicators showing low values.
Generally, in order to make multiplicative functions more manageable, the values of involved indicators are logarithmically transformed (summing up logarithm values corresponds to multiplying the original values). However, this procedure has to be followed with caution since it can also produce problems of interpretation.
If compensability is admitted, a unit showing a low value for one indicator will need higher values on the others in order to obtain a higher synthetic value. In this perspective, a compensatory technique can be useful in some contexts especially when the purpose of applying indicators is to stimulate behaviours aimed at improving the overall performance by investing in those ambits showing lower values.
All this highlights how important the choice of the aggregating technique is in order to avoid incoherencies between the weights previously chosen – in terms of theoretical meaning and importance – and the way these weights are actually used. In other words, in order to continue interpreting the weights as "importance coefficients", a non-compensatory aggregating procedure has to be preferred, such as a non-compensatory multi-criteria approach, like Multi-Criteria Analysis (MCA) (Nardo et al., 2005a).

b. *Comparability* refers to the distributional characteristics of the elementary indicators, in particular to directionality and functional form.

- **Directionality** concerns the direction by which each elementary indicator measures the QOL variable (i.e. positive or negative). In other words, the directionality issue refers to the fact that increases in some variables, such as literacy, correspond to increases in overall well-being, whereas increases in other variables, such as unemployment, correspond to decreases in overall well-being. In some cases it could be necessary to make uniform the directionality of the complete group of elementary indicators before starting with the aggregation. In order to make the directionalities uniform to a chosen direction (generally positive), the reflection procedure can be as follows:

$$\left[\left(higher \cdot value \cdot observed\right)+1\right]-\left(individual \cdot unit's \cdot original \cdot value\right)$$

  In case of quantitative values, the functional form refers to the evaluation that has to be assigned to absolute changes occurring at different levels of the indicators' distribution (marginal changes). In fact, when interpreting the level of a variable, two issues arise: first, are absolute values of a variable proportional in importance for overall-well being? And second, are changes in the value of a variable of equal importance at various levels of the variable? The response to these questions leads us to consider functional forms: linear and non-linear (Sharpe & Salzman, 2004).
- **Functional forms** represent the way changes in a variable are valued at different levels. If changes are valued in the same way, regardless of level, then the functional form should be linear. In some cases same absolute differences between observed values can have different meaning according to different levels of the scale.
  This occurs especially in the presence of indicators that have non-linear scale meaning. For example, a change of 100 euros in terms of income can have a different meaning if it occurs at a high or at a low level of income. In this case it could be necessary to decide the most convenient approach to use in terms of both interpretation and analytical treatment. If changes (Nardo et al., 2005a) are more significant at lower levels of the indicator, the functional form should be concave down (e.g. log or the nth root); on the opposite, if changes are more important at higher levels of the indicator, the functional form should be concave up (e.g. exponential or power).[5] Both the functional form are non-linear by definition.
  Because of the importance of level comparisons in social indices, it is advantageous to apply functional forms to variables so that the marginal changes associated to the value after a functional form has been applied are consistent with the value of a marginal change in society. Sinden writes,

[5] The standard choice is for log as the concave down function and power as the concave up function.

"Constant marginal utility for increases in any variables is a highly unlikely phenomenon". Utility theorists, Anderson, Hardaker, and Dillon, argue that applying a functional form with decreasing marginal returns has more basis than a linear one (Sinden, 1982).
Variables that are commonly taken into account in indices of social and economic well-being, such as per capita GDP, measures of unemployment, poverty gaps and rates, measures of inequality such as ratios of high and low incomes, and environmental depletion, are commonly thought to have significance of marginal changes that varies over the range of the observed values of the variable (Sharpe & Salzman, 2004).
Anand and Sen (1997) state that, in measures of poverty deprivation " the relative impact of the deprivation … would increase as the level of deprivation becomes sharper". According to this motivation, the UNDP develops measures of deprivation and inequality that more heavily penalize countries with higher indicators of deprivation in absolute value terms. For example, a decrease of 5 years of life expectancy from a base level of 40 is more heavily penalized than the same decrease beginning at a level of 80 (Sharpe & Salzman, 2004).

C. *Homogeneity* refers to the presence of common scales (level of measurement) adopted by the complete group of elementary indicators. Almost all the aggregating techniques require homogeneous scales. As we have seen in the previous chapter, some techniques exist allowing the elementary indicators' original scales to be transformed into an interpretable common scale. In order to select the proper approach, the data quality and properties and the objectives of the indicator should be taken into account.

The literature offers several **aggregation techniques** (Nardo et al., 2005a). The linear aggregation approach (additive techniques) are the most widely used. By contrast, multiplicative techniques (following the geometrical approach) and the technique based on multi-criteria analysis (following the non-compensatory approach) allow the difficulties that can be caused by compensation among the elementary indicators to be overcome:

			Aggregating approaches			
			1. Linear aggregation		2. Geometrical aggregation	3. Non-compensatory aggregation
			simple additive	Cumulative		
assumptions	Dimensionality	relationships between elementary indicators	uni	uni	uni	multi
	Model of measurement	relationship between elementary indicators and latent variable	monotonic	differential relationship	monotonic	
	Compensation	among elementary indicators	admitted	not admitted (scalability of indicators)	admitted	not admitted
	Homogeneity	of the level of measurement	requested	requested	requested	not requested

1. Linear aggregation approach

- **simple additive techniques**
  The simplest technique of aggregating indicators is additive. This technique is advantageous because of its methodological transparency (allowing also explicit weighting system to be used) but is also difficult to apply because of its strict prerequisite (the indicators to aggregate should be independent: low collinearity).
  The technique allows elementary indicators' values to be averaged (if elementary indicators' values are expressed by pure numbers, the geometrical mean calculation is to be preferred).
  If the values to be aggregated are rank values, the procedure is simplified.

simple additive	$CI_c = \sum_{i=1}^{n} I_{ic}$	simple additive averaging	$CI_c = \frac{\sum_{i=1}^{n} I_{ic}}{n}$
power additive averaging[6] (Sharpe & Salzman, 2004)	$CI_c = \left(\frac{\sum_{i=1}^{n} I_{ic}^{\alpha}}{n} \right)^{1/\alpha}$	with weights	$CI_c = \sum_{i=1}^{n} w_i I_{ic}$
for ranking data	$CI_c = \sum_{i=1}^{n} rank_{ic}$	CI made relative	$^{*}CI_c = \frac{CI_c - n}{mn - n}$

where

I_{ic}	elementary indicator *i*'s value for for case *c*
n	number of elementary indicators
m	number of cases
CI_c	composite indicator's value for case *c*
w_i	elementary indicator *i*'s weight
$rank_{ic}$	elementary indicator *i*'s rank value for unit *c*

- **cumulative techniques**
  the additive techniques requires indicators able to discriminate cases on the continuum, referring to the observed variable (cumulative in its nature, e.g., capacities, perception of social distance, dispositions, difficulties, and so on), in points that are different from each other. In other words, elementary indicators have to contribute to the description of the measured characteristic in different (cumulative) manners.
  In particular:
  - the selected elementary indicators refer to a single conceptual variable (*unidimensionality*)
  - each indicator has a *differentiated relationship* with the variable (*model of measurement*)
  - the selected indicators should not be compensable. This requirement can be operationalized in terms of *graduality/scalability*.

2. Geometrical aggregation approach (multiplicative techniques)

Developers of an index of social or economic well-being may want to include a variable quantity such as risk, which is a conditional probability and cannot be directly measured by a single variable alone. For example, the Index of Economic Market Well-Being seeks to measure the risk of single parent poverty. The only available variables are poverty incidence of single parent families and the rate of divorce. In order to find the rate of single parent poverty, we need to consider conditional probabilities. That is, the probability of being a single parent in poverty is modelled as the probability of being in poverty if you are divorced, times the probability of being divorced.[7] For this reason, the index measures the rate single parent poverty as the product of the rate of divorce times the rate of poverty among single parents.
This procedure has the advantage of more accurately quantifying the risk of poverty in society, but has some methodological problems. For one thing, if variables are scaled with LST before they are multiplied, the overall risk will not be scaled to the same range as the original variables. For example, the maximum and minimum levels of two variables may not ever be present in one single measurement, causing the multiplicative range to overestimate the actual range of the product of the two variables. Suppose that in 1980 the risk of divorce was .4 and the risk of single parent poverty was .8 and that in 2000 the risk of divorce was .8 and the risk of single parent poverty was .4. Then the total range coming from scaling variables before multiplying them would be .16 to .64, but in reality, both values are at .32. This problem could be overcome if conditional probabilities were first multiplied and then standardized. (Sharpe & Salzman, 2004)

[6] In order to analyze an extensive example, see 1997 Human Development Report for a complete mathematical characterization.
[7] This model assumes that single parents were once married.

4. Assessing robustness

A. Uncertainty and sensitivity analysis

As we have seen, many choices are necessary in order to proceed in aggregating multiple measures; these decisions can influence the robustness, that is, the capacity of the composite indicator to produce correct and stable measures (Edward & Newman, 1982; Nardo et al., 2005a; Saisana et al., 2005; Saltelli et al., 2004; Sharpe & Salzman, 2004; Tarantola et al., 2000). Assessing the robustness allows us to evaluate the role and the consequences of the subjectivity of the choices made as regards:

- the model to estimate the measurement error;
- the procedure for selecting the elementary indicators;
- the procedure for data management (missing data imputation, data standardization and normalization, etc.);
- the criterion for weight assignment;
- the aggregation technique used.

In order to evaluate the robustness of the composite indicator, a specific analysis procedure can be employed dealing with all the choices that can represent possible sources of uncertainty. In other words, the robustness is assessed by testing and comparing all the possible different performances that would have been obtained through different decisions along all the aggregation process. In particular, the procedure allows us to:

- evaluate the applicability of the model of measurement and the factors that contribute to the variability of the composite score,
- detect the choices producing values as stable as possible,
- understand the performance of the adopted model,
- ascertain the quality of the adopted model.

This procedure, which can be included in the wider field of the *what-if analysis*, is conducted through two stages; each stage corresponds to a different methodology of analysis (Nardo et al., 2005a):

1. ***uncertainty analysis***: the aim of this method is to analyze to what extent the composite indicator depends on the information composing it. In order to evaluate how the uncertainty sources influence the synthetic score obtained, the procedure identifies different scenarios for each individual case; each scenario corresponds to a certain combination of choices that produces a certain synthetic value;
2. ***sensitivity analysis***: the aim of this method is to evaluate the contribution of each identified source of uncertainty by decomposing the total variance of the synthetic score obtained; to this end, the procedure tests how much the synthetic score is sensitive to the different choices (small differences reveal low sensitivity).

The two approaches, generally treated in separate contexts, are very popular in any scientific field that requires the development and assessment of models (financial applications, risk analyses, neural networks); in addiction, the *uncertainty analysis* is adopted and applied more frequently than the *sensitivity analysis* (Jamison & Sandbu, 2001). The iterative and synergic application of both the procedures have been revealed to be useful and powerful (Saisana et al., 2005; Saltelli et al., 2004; Tarantola et al., 2000) in developing aggregated measures.[8]

B. Assessing discriminant capacity

Assessing the discriminant capacity (Maggino, 2007) of the composite indicator requires exploring its capacity in:

- discriminating between cases and/or groups; this can be accomplished by applying the traditional approaches of statistical hypothesis testing,
- distributing all the cases without any concentration of individual scores in few segments of the continuum; to this end, some coefficients were defined (Guilford, 1954; Maggino, 2003, 2007),
- showing values that are interpretable in terms of selectivity through the identification of particular values or reference scores. It allows the interpretation of the individual scores and eventually the selection of individual cases according to particular criteria; the reference scores are called **cut-point** and **cut-off**, with reference respectively to continuous and discrete data. The selection of these reference scores is particular useful when the composite indicator is applied for diagnostic and screening purposes.

Receiver Operating Characteristic or *Relative Operating Characteristic analysis* represents a valid method to be applied in order to test the discriminant capacity of a composite indicator. This analysis, connected directly to cost/benefit analysis in the ambit of *diagnostic decision making*, allows the relationship between sensitivity and specificity to be studied and analyzed in order to identify discriminant *cut-point*, *cut-off*, or *operating-point*.

ROC analysis is realized by studying the function that relates

- the probability of obtaining a "true alarm" among cases that needs an action (→ sensitivity → *hit rate* → **HR**)
- the probability of obtaining a "false alarm" among cases that do not need an action (→ 1-specificity → *false alarm rate* → **FAR**).

[8] The possibility of applying techniques such as *cluster analysis* should not be ignored since these techniques allow different and alternative typologies to be evaluated among the observed cases.

In order to study this relationship two rates are computed for each *cut-point.* An optimal curve can be obtained by defining many *cut-points* along the supposed continuum of the composite indicator. [9]

5.2.2 Functions of composite indicators

As previously pointed out, the utility in using composite indicators can arise in order to (Noll, 2009)
- answer the call by 'policy makers' for condensed information
- improve the chance to get into the media (compared to complex indicator systems)
- allow to make multi-dimensional phenomena uni-dimensional
- allow to compare situations across time more easily
- compare cases (e.g. nations) in a transitive way (ranking and benchmarking)
- allows clear cut answers to defined questions related to change across time, difference between groups of population or comparison between cities, countries, and so on.

In this perspective, each composite indicator can be classified according to several criteria.

Purposes

The indicators can be distinguished according to their **purpose**, which can be:
- *descriptive*, when the indicators are aimed at describing and knowing a particular reality (for example, quality of life). These indicators are said to be informative and baseline-oriented; in other terms, they allow changes along time, differences between geographical areas, and connections between social processes to pointed out;
- *explicative*, when the indicators are aimed at interpreting reality;
- *predictive*, when the indicators help to delineate plausible evolutional trends that is possible to describe in terms of development or decrement; these indicators require strong prediction models and continuous observations along time;
- *normative*, when the indicators are aimed at supporting, guiding, and directing decisions and possible interventions (policies) concerning problems to be solved. The normative function needs the definition of particular referenced standards defined in terms of time, territory, etc.; the reference values allow to evaluate the attainment of defined goals;
- *problem-oriented*, when the indicators are defined as a function of a specific hypothesis of research and analysis aimed at identifying contexts, kinds, severities of specific problems (for example the lack of quality of life conditions among immigrants);
- *evaluating*, that can be distinguished in
  - *practical*: indicators interfacing with observed process (e.g. in an organization),
  - *directional*: indicators testing if the observed condition is getting better or not,
  - *actionable*: indicators allowing change effects to be controlled.

Governance contexts

The indicators can be distinguished according to the **context** in which they are created, used, and interpreted. In this perspective, we can identify different contexts. For example:
- *public debates*: in this case the indicator/s have the function of informing, stimulating, forming and developing particular sensitiveness;
- *policy guidance*: in this case the indicators/s can support particular policy decisions;
- *administrative guidance*: in this case the indicator/s can support the evaluation of the different impacts of different alternatives.

[9] The procedure was conceived during the Second World War in order to study and improve the reception of radars and sonars. (Peterson, W. W., Birdsall, T. G., and Fox, W. C. (1954). *The theory of signal detectability*. Institute of Radio Engineers Transactions, PGIT-4, 171–212.). Main references:

Bamber, D. (1975) “The area above the ordinal dominance graph and the area below the receiver operating characteristic graph”, *Journal of Mathematical Psychology,* 12,387–415.

Bottarelli E., S. Parodi (2003) “Un approccio per la valutazione della validità dei test diagnostici: le curve R.O.C. (Receiver Operating Characteristic)”, *Ann. Fac. Medic. Vet. di Parma*, Vol. XXIII, 2003, pp. 49-68 (www.unipr.it/arpa/facvet/annali/2003/49.pdf)

Cummins, Robert A. 1995 “On the Trail of the Gold Standard for Subjective Well-Being”, *Social Indicators Research*, 35, 170-200.

Egan, J. P. (1975) *Signal detection theory and ROC analysis*, New York: Academic Press.

Coombs, C. H., Dawes, R. M., and Tversky, A. (1970) *Mathematical psychology: An elementary introduction*, Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Green, D. M. and Swets, J. A. (1966) *Signal detection theory and psychophysics*, New York: John Wiley & Sons, Inc.

Hanley, J. A. and McNeil, B. J. (1982) “The meaning and use of the area under a receiver operating characteristic (ROC) curve”, *Radiology,* 143, 29–36.

Sakitt, B. (1973) “Indices of discriminability”, *Nature,* 241, 133–134.

Simpson, A. J. and Fitter, M. J. (1973) “What is the best index of detectability?”, *Psychological Bulletin,* 80, 481–488.

Perspectives of observation

The indicators can be distinguished according to the different **perspectives of observation.** For instance, in the ambit of quality of life, a complex indicator that measures through :

- a *conglomerative* approach measures overall well-being, where increases in well-being of the best-off can offset decreases in well-being of the worst-off;
- a *deprivational* approach measures only the welfare of the worst-off (Anand & Sen, 1997).

Anand and Sen (1997) arguing that the conglomerative and deprivational perspectives are not substitutes for each other, proposed a *complementary* approach. “We need both, for an adequate understanding of the process of development. The plurality of our concerns and commitment forces us take an interest in each”. The adoption of complementary approach allows us to construct indices of social and economic well-being that should reflect the aggregated and disaggregated approaches. According to this methodology, conglomerative and deprivational indices should be constructed separately side-by-side along the lines of the United Nations Development Programme indicators (Sharpe & Salzman, 2004).

Forms of observation

The indicators can be distinguished according to the different **forms of observation.** In this perspective we can distinguish between:

- *status indicators*, which measure the reality in a particular moment; they allow for cross-comparisons between different realities. These indicators can produce cross data that need to be carefully managed since not the different realities can not always be directly compared; this is particularly true in the case of subjective characteristics observed in different geographical, social, cultural, political, environmental, and administrative conditions;
- *trend indicators*, which measure reality along time; they require a defined longitudinal observational design (for example, repeated surveys on particular populations). These indicators can produce *time series* that need to be carefully managed since the observed moments could reveal themselves to be incomparable and/or the defined indicators could reveal themselves as non applicable after some time.

Levels of communication

The indicators can be distinguished according to the different **levels of communication.** It regards the target group to which the final indicator will be communicated. In this perspective, indicators can be classified in:

- *cold indicators*: in this case, the indicators have a high level of scientific quality and show a high level of complexity and difficulty;
- *hot indicators*: in this case, the indicators are constructed at a low level of difficult and show a high level of understanding. It is unusual for these indicators to be used in a policy context;
- *warm indicators*: in this case, the indicators show a good balance between quality, comprehensibility, and resonance.

6. Modelling indicators

Dealing with a comprehensive conceptual framework requires exploring possible explanations of the relationships among the indicators, which conceptually model and hierarchically design the variables.
In this perspective, a proper analytical approach should be identified according to the defined conceptual framework. The feasibility of the different statistical approaches needs to be considered by taking into account their specific assumptions. The goal is to identify a procedure able to yield results, not only statistically valid and consistent with reference to the defined conceptual framework, but also easy to be read and interpreted at policy level.

Structural models approach

With reference to the causal explanatory perspective, we can refer to *Structural Equation Modelling* (SEM), which, as known, represents a statistical technique for testing and estimating causal relationships using a combination of statistical data and qualitative causal assumptions.
SEM is considered a confirmatory rather than exploratory approach. It usually starts with a hypothesis, represented as a model, operationalises the constructs of interest with a measurement instrument, and tests the model.
The causal assumptions embedded in the model often have falsifiable implications, which can be tested through data evidence. SEM can also be used inductively by specifying the model and using data to estimate the values of free parameters. Often the initial hypothesis requires to be adjusted in light of model evidence, but SEM is rarely used purely for exploration.
SEM models allow unreliability of measurement in the model to be explicitly captured and, consequently, structural relations between latent variables to be accurately estimated.
In the ambit of its specific assumptions, this approach can be adopted only in presence of a strong and indubitable conceptual interpretative framework concerning the causal relationships between objective and subjective indicators. In other words, it requires a strong acceptance of the direction of the relation between objective and subjective indicators.
Moreover, as shown above, two possible directions can be defined in casual explanation of well-being, *bottom-up* and *top-down*, which, however, are not separately able to explain completely the relationships between the observed variables. This means that causal effects can emerge in both directions. Diener (1984) suggested using both *bottom-up* and *top-down* approaches in order to examine the causal directions of well-being. Consequently, the application of the model allowing bi-directional effects to be estimated, has to be carried on with extreme caution (Scherpenzeel & Saris, 1996) and requires longitudinal data and analyses. The caution should increase especially in presence of both objective and subjective indicators.
Because of these difficulties, any application of this approach requires a strong conceptualisation of an explanatory model. Otherwise, any result can turn out to be misleading.

Multi-level approach

Multi-level analysis refers to statistical methodologies, first developed in the social sciences, which analyse outcomes simultaneously in relation to determinants measured at different levels (for example, individual, workplace, neighbourhood, nation, or geographical region existing within or across geopolitical boundaries) (Goldstein, 1999; Hox, 1995; Krieger, 2002).
This approach can be applied in the perspective of integrating objective and subjective indicators by assuming that people living in the same territory (e.g. city or region) share the same macro-level living conditions (objective quality of life) that contributes together with the micro-level living conditions (objective quality of live) to the subjective well-being. If the conceptual model is clearly specifiable and acceptable with reference to which variables are to be included in the study and at which level, these analyses can potentially assess whether individuals' well-being is influenced by not only "individual" or "household" characteristics but also "population" or "area" characteristics (Krieger, 2002). In fact, this approach assumes that structural characteristics of territories come before individual living conditions and that both precede subjective well-being. The goal is to describe the relationships between subjective well-being ("outcome" variable), territorial characteristics (macro-level living conditions: socio-economic conditions, demographic trend, and so on) and individual objective characteristics (micro-level living conditions: sex, religion, family composition, level of education, and so on).
The general analytical framework could be multiple regression: the subjective well-being is regressed on territorial and individual characteristics. If the goal is to evaluate the importance of territorial characteristics on subjective well-being, we could aggregate individual data at territorial level, but – as we know – this could result in the well-known *ecological fallacy*. In fact, the correlation between the observations resulting from the multilevel structure (the individuals on the same territory present the same values concerning the territory

characteristics) of data make the outcomes of the same territory more homogeneous than those yielded by a random sample of individuals drawn from the whole population. This higher homogeneity is naturally modelled by a positive within-territory correlation among individual level of subjective well-being in the same territory. This problem can be avoided by applying a variance component model.

In statistics, a *variance components model*, also called *random effect/s model***,** is a kind of *hierarchical linear model*. These models (along with generalized linear mixed models, nested models, mixed models, random coefficient, random parameter models, split-plot designs) are part of *multilevel models* (Bryk & Raudenbush, 2002), which are statistical models of parameters that vary at more than one level. These models can be seen as generalizations of linear models (also extendible to non-linear models)[1] and represent more advanced forms of simple linear regression and multiple linear regression. They are appropriate for use with nested data. In particular, they assume that the data describe a hierarchy of different populations whose differences are constrained by the hierarchy.

In other words, multilevel analysis allows variance in outcome variables to be analysed at multiple hierarchical levels, whereas in simple linear and multiple linear regression all effects are modelled to occur at a single level.

For example, in educational research, where data is often considered as pupils nested within classrooms nested within schools, it may be necessary to assess the performance of schools teaching by one method against schools teaching by a different method. It would be a mistake to analyse this kind of data as though the pupils were simple random samples from the population of pupils taught by a particular method. Pupils are taught in classes, which are in schools. The performance of pupils within the same class will be correlated, as will the performance of pupils within the same school.

Conceptually the model is often viewed as a hierarchical system of regression equations. For example, assume we have data in J groups or contexts and a different number of individuals N_j in each group. On the individual (lowest) level we have the dependent variable Y_{ij} and the explanatory variable X_{ij}, and on the group level, we have the explanatory variable Z_j. Thus, we have a separate regression equation in each group:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \tag{1}$$

The β_j are modelled by explanatory variables at the group level:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \tag{2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \tag{3}$$

Substitution of (2) and (3) in (1) gives:

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij} + u_{1j} X_{ij} + u_{0j} + e_{ij} \tag{4}$$

in general there will be more than one explanatory variable at the lowest level and also more than one explanatory variable at the highest level. Assume that we have *P* explanatory variable *X* at the lowest level, indicated by the subscript *p* (*p*=1,…, *P*), and *Q* explanatory variables *Z* at the highest level, indicated by the subscript *q* (*q*=1, …, *Q*). Then, equation (4) becomes the more general equation:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + u_{pj} X_{pij} + u_{0j} + e_{ij} \tag{5}$$

Multilevel analysis generally uses Maximum Likelihood (ML) estimators, with standards errors estimated from the inverse of the information matrix. Computing the ML estimates requires an iterative procedure. (Bryk and Raudenbush, 1992; Goldstein, 1999; Hox, 1995)

Even if the multilevel approach presents logic and analytic solutions acceptable from the statistical point of view, this method should be considered carefully in the context of quality of life. For instance, when the territorial characteristics do not affect individuals in the same manner and with the same degree (territorial heterogeneity), some authors (Rampichini & Schifini, 1998) suggest introducing a new level in the hierarchy, represented by individuals within each territory. For example, different clusters of individuals could be identified sharing same living conditions at micro-level. This could lead to results in which similar clusters are in different territories.

Life-course perspective

Life-course perspective refers to a conceptual model that considers well-being status at any given individual state (age, sex, marital status) not only reflecting contemporary conditions but also embodying prior living circumstances. This means that we could try to study people's developmental trajectories (environmental and social) over time, by considering also the historical period in which they live, in reference to their society's social, economic, political, and ecological context. This approach assumes that some components can exist which can determine an effect, at a sensitive or "critical" period of individual life, lasting, or having a

[1] Multilevel analysis has been extended to include multilevel structural equation modelling, multilevel latent class modelling, and other more general models.

lifelong significance. The interest could be oriented to analysing which of these processes are reversible and which could be the role of objective micro or macro level characteristics in this.
This perspective deserves particular attention and consideration. Its limit is mainly represented by the difficulty to obtain detailed and consistent individual longitudinal data and by the complexity of managing, analysing, and modelling this kind of data. According to its characteristics, this approach turns out to be useful in order to study phenomena circumscribable through a clinical logic.

Bayesian networks approach

A Bayesian network is a graphical model representing a certain reality described by variables. The goal is to explore the relationships among the variables of interest through probabilities.[2]
When used in conjunction with statistical techniques, the Bayesian network model has several advantages for data analysis because:

1. the model encodes dependencies among all variables and handles situations where some data entries are missing
2. it can be used to learn causal relationships, and hence can be used to gain understanding about a problem and to predict the consequences of intervention
3. it has both a causal and probabilistic semantics, it is an ideal representation for combining prior knowledge (which often comes in causal form) and data[3]
4. Bayesian statistical methods in conjunction with Bayesian networks offer an efficient and principled approach aimed at data overfitting.

This is called diagnostic, or "bottom up", reasoning, since it goes from effects to causes; it is a common task in expert systems. Bayes nets can also be used for causal, or "top down", reasoning. For example, we can compute the probability that the grass will be wet given that it is cloudy. Hence Bayes nets are often called "generative" models, because they specify how causes generate effects

A Bayes net represents a model, reflecting the states of some part of a world that is being modelled and describing how those states are related by probabilities. All the possible states of the model represent all the possible worlds. The direction of the link arrows roughly corresponds to "causality". That is the nodes higher up in the diagram tend to influence those below rather than, or, at least, more so than the other way around.
In a Bayes net, the links may form loops, but they may not form cycles.
In the past, when scientists, engineers, and economists wanted to build probabilistic models of worlds, so that they could attempt to predict what was likely to happen when something else happened, they would typically try to represent what is called the "joint distribution".
This model has several **advantages** for data analysis:

1. the model encodes dependencies among all variables, it readily handles situations where some data entries are missing.
2. it is adaptable since it can be used to learn causal relationships, and hence can be used to gain understanding about a problem domain and to predict the consequences of intervention.
3. it has both a causal and probabilistic semantics, it is an ideal representation for combining prior knowledge (which often comes in causal form) and data.
4. it offers an efficient and principled approach aimed at data overfitting.
5. Since a Bayes net only relates nodes that are probabilistically related by some sort of causal dependency, an enormous saving of computation can result. There is no need to store all possible configurations of states. All that is needed to store and work with is all possible combinations of states between sets of related parent and child nodes (families of nodes).
6. it can be useful in assisting decision making. If some states lead to “positive” results (e.g. pleasure), while others to negative outcome (e.g. pain), it is possible to implement the model in order to maximize the former and minimize the latter. There is a science of decision making that mixes probability with measurements of value. It is called *Decision Theory* or *Utility Theory*. Bayes nets are easily extended to computing utility, given the degree of knowledge we have on a situation, and so

[2] Bayesian networks are based upon the concept of conditional probability. **Conditional probability** is the probability of some event *A*, given the occurrence of some other event *B*. Conditional probability is written $P(A|B)$, and is read "the probability of *A*, given *B*". The conditional and marginal probabilities of two random events are related in probability theory by **Bayes' theorem** (often called **Bayes' law** after Rev Thomas Bayes). It is often used to compute posterior probabilities given observations. For example, a patient may be observed to have certain symptoms. Bayes' theorem can be used to compute the probability that a proposed diagnosis is correct, given that observation.
As a formal theorem, Bayes' theorem is valid in all common interpretations of probability. However, it plays a central role in the debate around the foundations of statistics: frequentist and Bayesian interpretations disagree about the ways in which probabilities should be assigned in applications. According to frequentist approach, probabilities are assigned to random events according to their frequencies of occurrence or to subsets of populations as proportions of the whole. In Bayesian perspective, probabilities are described in terms of beliefs and degrees of uncertainty.

[3] Classical inferential models do not permit the introduction of prior knowledge into the calculations. This prevents the introduction of extraneous data that might skew the experimental results. However, there are times when the use of prior knowledge would be a useful contribution to the evaluation process.

they have become very popular in business and civic decision making as much as in scientific and economic modeling.

Some **limitations** can be identified.

1. the remote possibility that a system's user might wish to violate the distribution of probabilities upon which the system is built.
2. the computational difficulty of exploring a previously unknown network.
3. the quality and extent of the prior beliefs used in Bayesian inference processing. A Bayesian network is only as useful as this prior knowledge is reliable. Either an excessively optimistic or pessimistic expectation of the quality of these prior beliefs will distort the entire network and invalidate the results. Related to this concern is the selection of the statistical distribution induced in modelling the data. Selecting the proper distribution model to describe the data has a notable effect on the quality of the resulting network.

- . - . - .

Traditional explorative approaches, such as clustering and mapping approaches, multidimensional analysis, correspondences analysis (Aldenderfer & Blashfield, 1984; Bailey, 1994; Corter, 1996; Hair, 1998; Lis and Sambin, 1977), should be added to the approaches presented above. The approaches are all practicable but in view of their application, their capability to meet assumptions and to fit the needs of the conceptual framework need to be explored.

Composite indicators

Composite indicators could represent one of the possible technical approaches to modelling indicators, which would turn out to be aggregated in a unique value referring to each unit of interest (city, country, and so on). This proposal can appear attractive at a first glance but does not reveal to be easy and creates conceptual, interpretative and analytical problems when the aggregation involves measures very different, conceptually and metrically (e.g. objective and subjective indicators).

For example, we can consider the standardization issue: in order to create composite indicators, data need to be reduced to a common reference-metric. That is particularly significant when data are measured with reference to different methodologies; for example, individual data do not always meet the requirement of metric measurement (like some objective individual information, for example, family typology); the problem is how to face the issue without adopting sophisticated approaches. In our opinion, this approach could be carefully considered as one of the possible solutions for integration.

7. Closing remarks

7.1 Methodological challenges in indicators construction for the measurement of societal well-being

Actually, even a quick check of the academic literature allows us to notice a long tradition and intense research work existing in the field of measuring societal well-being through complex approaches. Sometimes, this tradition has been set against the hard economic perspective that accounts the economic indicators as the main and unique approach allowing progress to be measured.
The recent debates on different perspectives in measuring societal well-being led to different scenarios also in academic research. Some challenges can be drawn:

1) concerning the conceptual model:
    a. More attention and efforts are needed in order to:
        - better define **sustainability**, in particular on its relationship with quality of life
        - join the concept of sustainability (more related to the future generations dimension) with the concept of **vulnerability** (more related to the future of present generations dimension)[1]
    b. **Subjective indicators** should not be seen as antagonist of objective indicators but as an important tool allowing information to be added, which cannot be provided by objective measures. In both perspective, the measurement process needs
        - an agreement on what and how to measure
        - a clear conceptual framework clarifying the relationship between objective and subjective measures and their integration
2) concerning methodological issues:
    a. It is impossible to assess complex phenomena with a single indicator (even using composite indicator) and it is necessary to define and deal with **sets of indicators**.
    b. As regards **subjective indicators**, it is important to
        - define accurate measures (e.g., notable academic research exists in the field of scaling techniques)
        - improve and enhance existing data sources
    c. More work should be done on **reliability** of indicators and **their comparative capacity** among countries, across time to deal with different level of analysis.
3) concerning strategic issues:
    a. More attention should be paid in order to improve
        - **quality** of indicators
        - **legitimacy**, trust, authority and credibility of indicators of well-being of societies

There is a great need of exchanging information and dialogue on these issues between different actors and within different research contexts.

7.1.1 Some key issues

7.1.1.1 Selecting indicators

Different issues need to be addressed in order to selecting and managing indicators, especially when this is carried out into a complex system allowing the accomplishment of functions like monitoring, reporting and accounting. Michalos (in Sirgy et al., 2006) identified 15 different issues related to the combination of social, economic and environmental indicators. As Michalos asserts, the issues collectively yield over 200,000 possible combinations representing at least that many different kinds of systems (Sirgy et al., 2006):

- Settlement/aggregation area sizes: e.g., the best size to understand air pollution may be different from the best size to understand crime.

[1] The issue has been pointed out by Enrico Giovannini in different occasions (e.g., the oral communication at the conference “From GDP to Well-being”, December 3-5, 2009 – Ancona, Italy).

- Time frames: e.g., the optimal duration to understand resource depletion may be different from the optimal duration to understand the impact of sanitation changes.
- Population composition: e.g., analyses by language, sex, age, education, ethnic background, income, etc. may reveal or conceal different things.
- Domains of life composition: e.g., different domains like health, job, family life, housing, etc. give different views and suggest different agendas for action.
- Objective versus subjective indicators: e.g., relatively subjective appraisals of housing and neighbourhoods by actual dwellers may be very different from relatively objective appraisals by "experts".
- Positive versus negative indicators: negative indicators seem to be easier to craft for some domains, which may create a biased assessment, e.g., in the health domain measures of morbidity and mortality may crowd out positive measures of well-being.
- Input versus output indicators: e.g., expenditures on teachers and school facilities may give a very different view of the quality of an education system from that based on student performance on standardized tests.
- Benefits and costs: different measures of value or worth yield different overall evaluations as well as different evaluations for different people, e.g., the market value of child care is far below the personal, social or human value of having children well cared for.
- Measurement scales: e.g., different measures of well-being provide different views of people's well-being and relate differently to other measures.
- Report writers: e.g., different stakeholders often have very different views about what is important to monitor and how to evaluate whatever is monitored.
- Report readers: e.g., different target audiences need different reporting media and/or formats.
- Conceptual model: e.g., once indicators are selected, they must be combined or aggregated somehow in order to get a coherent story or view.
- Distributions: e.g., because average figures can conceal extraordinary and perhaps unacceptable variation, choices must be made about appropriate representations of distributions.
- Distance impacts: e.g., people living in one place may access facilities (hospitals, schools, theatres, museums, libraries) in many other places at varying distances from their place of residence.
- Causal relations: before intervention, one must know what causes what, which requires relatively mainstream scientific research, which may not be available yet.

Choices and options selected for each issue have implications for the other issues. The issues are not mutually exclusive and are not expected to be exhaustive as other can be identified

Dealing with these issues is merely a technical problem to be solved by statisticians or information scientists. On the other side, the construction of indicators of well-being and quality of life is essentially a political and philosophical exercise, and its ultimate success or failure depends on the negotiations involved in creating and disseminating the indicators, or the reports or accounts that use those indicators. (Michalos, in Sirgy et al., 2006)

Within a system, we consider also the difficulties related to the availability of indicators (across time and space) and in harmonizing different data sources and levels of observation.

7.1.1.2 Quality of indicators

Many international institutions, like World Bank & Unesco (Patel et al., 2003) and Eurostat (2000) tried to identify the attributes of **quality** that indicators (and approaches aimed at their management) should possess and need to be considered in the process of developing of new indicators or of selecting available indicators:

(I) Methodological soundness

This characteristic refers to the idea that the methodological basis for the production of indicators should be attained by following internationally accepted standards, guidelines, or good practices. This dimension is necessarily dataset-specific, reflecting different methodologies for different datasets. The elements referring to this characteristic are (i) concepts and definitions, (ii) scope, (iii) classification / sectorization, and (iv) basis for recording. Particularly important is the characteristic of **accuracy and reliability**, referring to the idea that indicators should be based upon data sources and statistical techniques that are regularly assessed and validated, inclusive of revision studies. This allows accuracy of estimates to be assessed. In this case accuracy is defined as the closeness between the estimated value and the unknown true population value but also between the observed individual value and the "true" individual value. This means that assessing the accuracy of an estimate involves analyzing the total error associated with the estimate: sampling error and measurement error.

(II) Integrity

Integrity refers to the notion that indicator systems should be based on adherence to the principle of objectivity in the collection, compilation, and dissemination of data, statistics, and results. The characteristic includes institutional arrangements that ensure
(i) professionalism in statistical policies and practices,
(ii) transparency, and
(iii) ethical standards.

(III) Serviceability

Comparability is a particular dimension of serviceability. It aims at measuring the impact of differences in applied concepts and measurement tools/procedures

- over time, referring to comparison of results, derived normally from the same statistical operation, at different times,
- between geographical areas, emphasizing the comparison between countries and/or regions in order to ascertain, for instance, the meaning of aggregated indicators at the chosen level,
- between domains. This is particularly delicate when involving subjective measurement (e.g. cultural dimensions).

(IV) Accessibility

Accessibility relates to the need to ensure
(i) clarity of presentations and documentations concerning data and metadata (with reference to information environment: data accompanied with appropriate illustrations, graphs, maps, and so on, with information on their quality, availability and – eventual – usage limitations)
(ii) impartiality of access
(iii) pertinence of data
(iv) prompt and knowledgeable support service and assistance to users
In other words, it refers also to the physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro or macro data, various formats (paper, files, CD-ROM, Internet…), etc.

AN INDICATOR SHOULD BE	clear appropriate exhaustive	meaningful accurate well-designed	**WITH REFERENCE TO ITS CAPACITY AND POSSIBILITY TO**	**define and describe** (*concepts, definitions and scopes*)	**(I) METHODOLOGICAL SOUNDNESS**
	measurable	stable		**observe** unequivocally and stably (in terms of space and time)	
	reliable valid repeatable robust	rigorous precise exact faithful		record by a **degree of distortion as low as possible** (explored through statistical and methodological approaches)	
	transparent	with ethical standards		**adhere** to the principle of objectivity in the collection, compilation, and dissemination	**(II) INTEGRITY**
	consistent	pertinent coherent		**reflect adequately the conceptual model** in terms of aims, objectives and requirements underlying its construction (knowing, monitoring, evaluation, accounting, …)	**(III) SERVICEABILITY**
	relevant			meet current and potential users' needs. It refers to whether all indicators that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflects user needs. The identification of users and their expectations is therefore necessary.	
	practicable revisionable	up-to-datable		**be observed** through realistic efforts and costs in terms of development and data collection (for example, short time between observation and data availability)	
	well-timed	timely periodic		reflect the length of time between its availability and the event or phenomenon it describes	
	regular	punctual		reflect the time lag between the release date of data and the target date when it should have been delivered	
	comparable discriminant	disagregable thrifty		**be analyzed** in order to record differences and disparities between units, groups, geographical areas and so on, by employing the available information as much as possible	
	believable accessible interpretable	comprehensible simple manageable		**be spread** that is, it has to be easily findable, accessible, useable, analyzable, and interpretable in order to gain also users' confidence (*brand image*)	**(IV) ACCESSIBILITY**

Prerequisites of quality

Although it does not represent a dimension of quality in itself, prerequisites of quality refers to all those (institutional or not) preconditions and background conditions for quality of statistics allowing.
In other words, indicators construction is not simply a technical problem but should become part of a larger debate concerning how to construct indicators obtaining a larger legitimacy to be promoted. These prerequisites cover the following elements:
(i) legal and institutional environment, allowing

a. conceptual framework to be defined
b. coordination power within and across different institutions to be framed
c. data and resources to be available for statistical work

(ii) quality awareness informing statistical work.

7.1.1.3 Technical issues

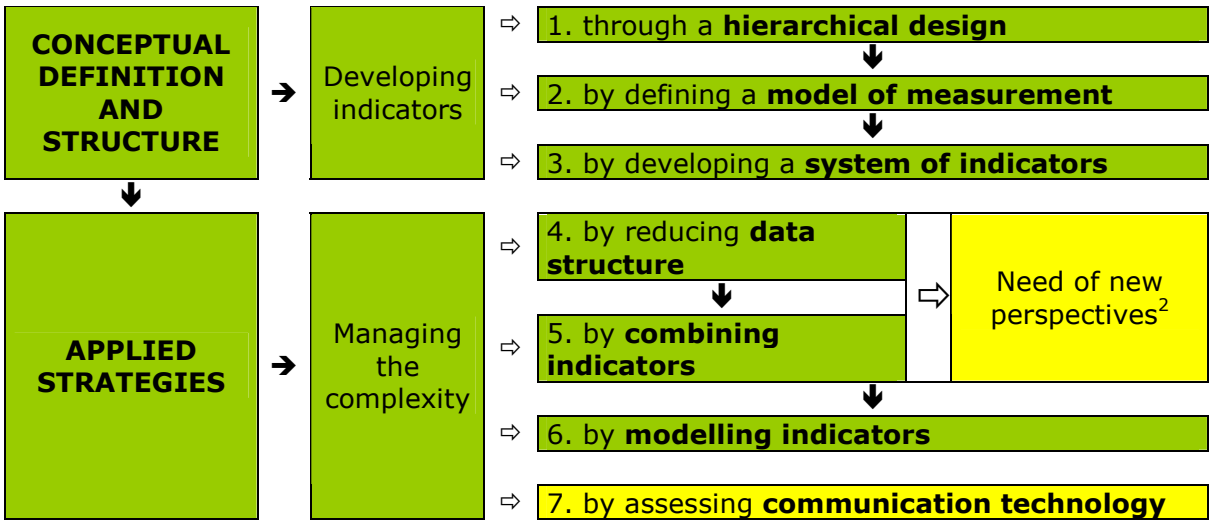
The application of hierarchical design produces a complex data structure. In order to manage the complexity it should be take into account that data

- refer to a complex reality
- are ambiguous and softened
- are multidimensional
- are dynamic and evolutionary
- are qualitative also when quantitatively measured
- contain errors and approximations
- are sensitive

By considering all this, new challenges and perspectives can be identified in order to improved the technical strategies allowing the complexity to be managed. In particular, we can identify the following critical issues:

(i) reducing data structure (e.g., aggregating observational units – micro-units – in order to define analytical units – macro-units) and combining indicators

(ii) correctly and significantly communicating the "picture" obtained through the indicators (correctly presenting results).



7.2 Institutional challenges: national statistical offices and the measurement of societal well-being

As we have seen, measuring and monitoring well-being of societies require a complex and comprehensive framework and integrated approaches at conceptual and methodological level. This perspective is urged not only by researchers belonging to academics but also by other organizations and institutions.
Also the awareness aroused by many personalities directs towards more comprehensive approach in measuring societal well-being. The Report of Commission on the Measurement of Economic Performance and Social Progress (Stiglitz, Sen & Fitoussi, 2009) – chaired by Joseph E. Stiglitz – represents a further evidence of that and proposes the following twelve recommendations:

[2] A particular and promising approach is that based upon the study of structured relationships based upon logical rather then statistical approaches. We are going to present a promising application on this.

1 When evaluating material well-being, look at income and consumption rather than production.

2 Emphasise the household perspective.

3 Consider income and consumption jointly with wealth.

4 Give more prominence to the distribution of income, consumption and wealth.

5 Broaden income measures to non-market activities.

6 Quality of life depends on people's objective conditions and capabilities. Steps should be taken to improve measures of people's health, education, personal activities and environmental conditions. In particular, substantial effort should be devoted to developing and implementing robust, reliable measures of social connections, political voice, and insecurity that can be shown to predict life satisfaction.

7 Quality-of-life indicators in all the dimensions covered should assess inequalities in a comprehensive way.

8 Surveys should be designed to assess the links between various quality-of-life domains for each person, and this information should be used when designing policies in various fields.

9 Statistical offices should provide the information needed to aggregate across quality-of-life dimensions, allowing the construction of different indexes.

10 Measures of both objective and subjective well-being provide key information about people's quality of life. Statistical offices should incorporate questions to capture people's life evaluations, hedonic experiences and priorities in their own survey.

11 Sustainability assessment requires a well-identified dashboard of indicators. The distinctive feature of the components of this dashboard should be that they are interpretable as variations of some underlying "stocks". A monetary index of sustainability has its place in such a dashboard but, under the current state of the art, it should remain essentially focused on economic aspects of sustainability.

12 The environmental aspects of sustainability deserve a separate follow up based on a well-chosen set of physical indicators. In particular there is a need for a clear indicator of our proximity to dangerous levels of environmental damage (such as associated with climate change or the depletion of fishing stocks.)

As a consequence, measuring and monitoring societal well-being arouse a great need of statistics but statistics needs to find to elaborate new and shared working models.
Moreover, require huge investments in order to accomplish survey projects (systematic or finalized) and systematic control on data quality.
Managing this complexity requires the involvement of different governance levels, which represents a new challenge for statistics and for the statistical offices.
Following the OECD Istanbul Declaration – signed by representatives of the European Commission, the Organisation for Economic Cooperation and Development, the Organisation of the Islamic Conference, the United Nations, the United Nations Development Programme and the World Bank, during the II OECD World Forum on "Statistics, Knowledge and Policy" (2007) – societies urge statistical offices, public and private organisations, and academic experts to work alongside representatives of their communities to produce high-quality, facts-based information that can be used by all of society to form a shared view of societal well-being and its evolution over time. [3]
A possible model could be that aimed at involving different public corporations operating in statistical ambits and interacting between them in order to define an organic system, operating as coordinated network organization (*statistical offices network*). Network's activities should be structured in nodes and needs to be

- aimed at defining clear statistical goals and programs
- organized at different level (national, regional or local)
- planned with special reference to data production, in order to avoid redundancies, to rationalize the network and to qualify the nodes
- harmonized with reference to the statistical function, by overcoming fragmentations, diversities, superimpositions at different network levels
- adjusting forms of communication and involvement for the different actors.

These actions could be conceived at

- general level, since they should define norms concerning the statistical functions to be considered as a transversal service and a common and multifunctional wealth. Statistics should be considered in terms of knowledge and assessment;

[3] This concept is presented and broadly examined by Enrico Giovannini (OECD – chief statistician) with reference to the spreading of statistical information – in the paper *The role of statistics in a globalised world: risks and challenges* presented at the DGINS (Directors-General of the National Statistical Institutes) Conference, 20–21 September 2007, Budapest, Hungary

- specific level, since they should promote i) increasing the production of data and indicators at local level; ii) interacting and integrating different data bases and data sources; iii) developing appropriate analytical methods.

Some risks could arouse, related to the lack of coordination (the activities could turn out to be dispersed, fragmented, marginalized and excessively differentiated) and reciprocal knowledge of each node's activities.
In order to avoid that, the network requires

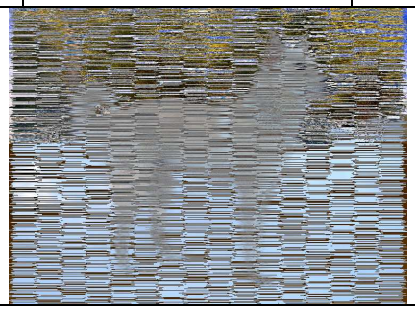
- new professionals to be defined
- new competences to be developed
- a system of statistical data certification to be implemented
- a strong support from administrative sectors to be assured.

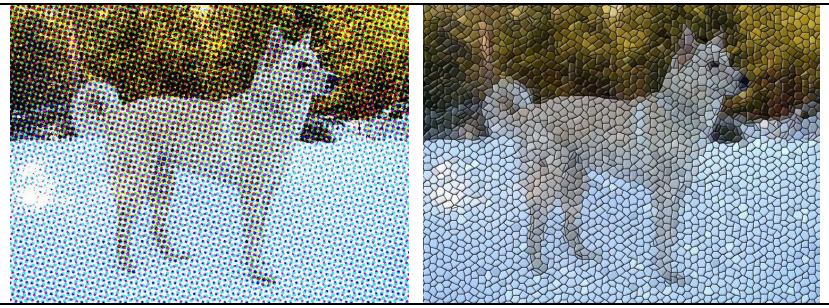
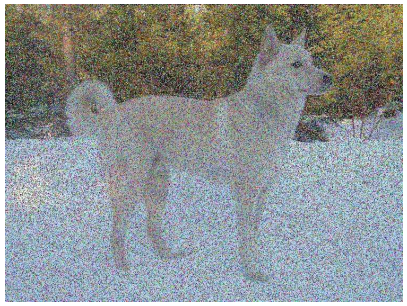
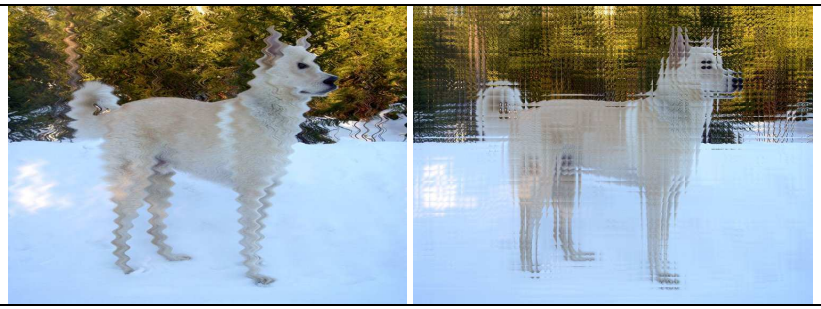
All these efforts should aim at splitting the role of official statisticians from "information providers" to "knowledge builders". [4]

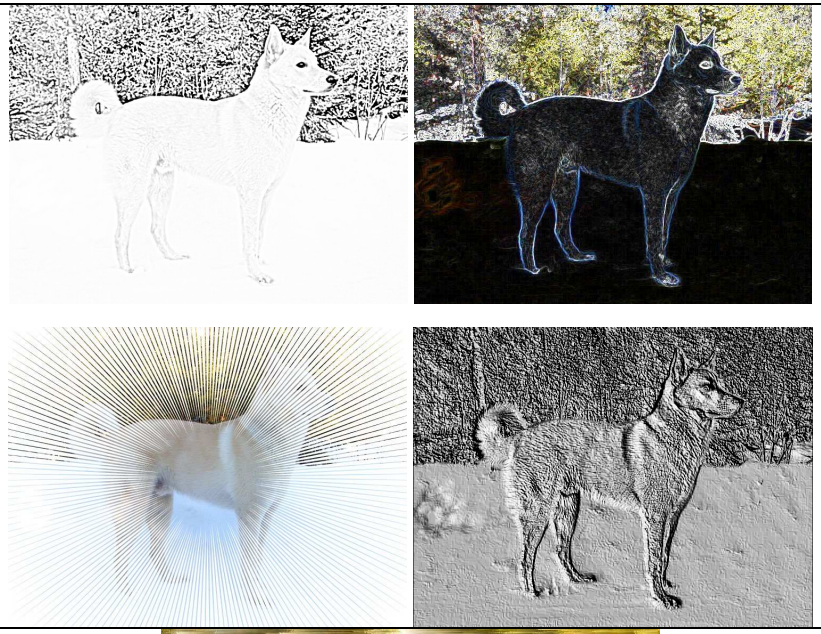
7.3 Observing and monitoring the dog

At the end of this work, we could try to simplify the process of indicators construction and to identify the indicators' role in measuring and monitoring societal well-being, in the following way.
The example is unpretentious but is just aimed at showing in a simple way the characteristics of indicators construction.

[4] The issue has been pointed out by Enrico Giovannini in different occasions (e.g., seminar on "*New Techniques and Technologies for Statistics (NTTS)"* – EUROSTAT, February 18-20, 2009 – Brussels, Belgium).

			A complex reality exists ➔				
How can we reach a reasonable picture of that reality?	Definition of the hierarchical design	(i)	Conceptual framework ➔	The dog (*Canis lupus familiaris*) is a domesticated form of the Wolf, a member of the Canidae family of the order Carnivora. The term is used for both feral and pet varieties. ……. [from Wikipedia]			
		(ii)	Areas ➔	Biology	Intelligence	Behaviour	…
		(iii)	Variables ➔	Sense Physical characteristics …	Cognitive development …	Social structure Social cognition …	
		(iv)	Indicators and their modelling hypothesis ➔	sight, hearing, smell, coat, tail, health, mortality, predation, diet, reproduction, …	Reinforcement …	Dominance hierarchy, …	
	Definition of the relationship between the indicators and the pictures.		In order to obtain a comprehensive picture, we need a high numbers of indicators (points). Collecting all the indicators does not mean that the whole picture is recognizable. ➔				

		The possibility to reconstruct the whole picture depends by the number of adopted indicators (picture's resolution). ➔	
		In order to all the indicators the capacity to reconstruct the whole picture we need a proper analytical approach. A proper analytical model allows the picture to be reconstructed by all the indicators. However, the high numbers of indicators are difficult to be managed. ➔	
		In order to reduce the difficulties in managing all these indicators, composite indicators may be built (cards). However the picture can turn out to be not completely trusty (lost of detailed information) ➔	
	Errors	In any case the reconstruction may contain errors – of different origins – that may deform the whole picture. ➔	

	Relationship between the picture, as reconstructed by indicators, and the reality.	The obtained picture of course is not the reality but can help in describing it. In order to give a meaning to the picture, actually related to the reality, we need to look at the picture by going back to the conceptual framework. This means that the meaning of the picture will be different according to different conceptual framework. ➔	
	But the reality is "on the move" (longitudinal perspective)	Of course, the picture does not refers to a motionless reality and if the observation would like to take into consideration the "on the move" perspective the conceptual model should be more complex. ➔	

7.4 The "flight desk"

As asserted, a complex approach is needed in order to measure and monitor well-being of societies. Consequently, the complexity requires multiple indicators conceived and organized in a conceptual structure. Of course, the system needs a guidance at different levels (e.g., policy). By attempting to depict this situation, we could image the policy maker as a pilot sitting at the flight desk:

actually, the present work aimed at presenting and discussing the process allowing the "flight desk" to be defined, constructed, and assessed. But, in order to let the system taking off, the "flight desk" needs to

- share the decision on the destination (→ democracy)
- know the destination (→ goals)
- know pre-conditions (→ resources, …)
- know flight conditions (→ monitoring)communicate flight conditions (→ communication system)
- know how to manage emergency conditions

At the same time, it requires a wide sharing and support, allowing a research environment urging investigations aimed at implementing and improving the measurement and monitoring of well-being of societies.

References

Abbey A., F.M. Andrews (1985) “Modeling the psychological determinants of life quality”, *Social Indicators Research*, Volume 16.

Aldenderfer M.S., Blashfield R.K. (1984) *Cluster Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-044, Beverly Hills, CA:Sage.

Alt M. (1990) *Exploring Hyperspace. A Non-Mathematical Explanation of Multivariate Analysis*, McGraw-Hill, New York - London.

Anand, S., Sen A. (1997) “Concepts of Human Development and Poverty: A Multidimensional Perspective” *Human Development Papers 1997*, UNDP.

Andersen E.B. (1972) "The Numerical Solution of a Set of Conditional Estimation Equations", in *Journal of the Royal Statistical Society*, Serie B, 34.

Andersen E.B. (1973) "A Goodness of Fit Test for the Rasch Model" in *Psychometrika*, 38.

Anderson T.W. (1958) *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, Inc., New York, London, Sidney.

Andrews, F. M. & Withey, S. B. (1976) *Social Indicators of Well-Being: Americans’ Perceptions of Life Quality*, New York, Plenum Press.

Andrich D. (1988) *Rasch Models for Measurement*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-068, Newbury Park, CA:Sage.

Asher H.B. (1983) *Causal modelling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-003, Newbury Park, CA:Sage.

Bailey K.D. (1994) *Typologies and Taxonomies. An Introduction to Classification Techniques*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-102, Thousand Oaks, CA: Sage.

Bartholomew, D J, and Knott, M (1999) *Latent Variable Models and Factor Analysis*, Kendall's Library of Statistics, vol. 7. Arnold Publishers, London.

Berger-Schmitt R. and H.-H. Noll (2000) *Conceptual Framework and Structure of a European System of Social Indicators*, EuReporting Working Paper No. 9, Centre for Survey Research and Methodology (ZUMA) – Social Indicators Department, Mannheim.

Bezzi C., L. Cannavò, M. Palumbo (2009 in printing) *Gli indicatori sociali e il loro uso valutativo*, Franco Angeli Editore, Milano.

Blalock, H.M. (1964) *Causal Inferences in Nonexperimental Research*, Chapel Hill, University of North Carolina Press, NC.

Blalock H.M. jr., ed., (1974) *Measurement in the Social Sciences. Theories and Strategies*, Aldine Publishing Company, Chicago.

Bock R.D., M. Aitkin (1981) "Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM algorithm" in *Psycometrika*, 46.

Bohrnstedt G.W. and Knocke D. (1994) *Statistics for Social Data Analysis*, Peacock P.

Bolasco S. (1999) *Analisi multidimensionali dei dati. Metodi, strategie e criteri di interpretazione*, Carocci, Roma.

Booysen, F. (2002) “An overview and evaluation of composite indices of development”, *Social Indicators Research*, 59(2), 115-151.

Bryk A.S., Raudenbush S.W. (1992, second edition 2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*, SAGE Publications, Thousand Oaks.

Campbell A., P.E. Converse, W.L. Rodgers (1976) *The Quality of American life*, Russell Sage Foundation, New York.

Cannavò L. (2009 in printing) “Dell’incertezza e della complessità: gli indicatori tra ricerca valutazione” in Bezzi C., L. Cannavò, M. Palumbo (eds) *Gli indicatori sociali e il loro uso valutativo*, Franco Angeli Editore, Milano.

Cooley W.W., Lohnes P.R. (1971) *Multivariate data analysis*, John Wiley & Sons, Inc., New York, London, Sidney.

Corbetta P. (1992) *Metodi di analisi multivariata per le scienze sociali*, Il Mulino, Bologna.

Corbetta P. (2003) *Metodologia e tecniche della ricerca sociale*, Il Mulino, Bologna.

Corter J.E. (1996) *Tree models of similarity and association*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-112, Thousand Oaks, CA:Sage.

Costa, P.T. & McCrae, R.R. (1980) “Still stable after all these years: personality as a key to some issues in adulthood and old age”, in P. B. Baltes & O. G. Brim (eds.) *Life span development and behaviour*, New York, NY: Academic Press, 65-102.

Costanza R., B. Fisher, S. Ali, C. Beer, L. Bond, R. Boumans, N.L. Danigelis, J. Dickinson, C. Elliott, J. Farley, D. Elliott Gayer, L. MacDonald Glenn, T. Hudspeth, D. Mahoney, L. McCahill, B. McIntosh, B. Reed, S. A. Turab Rizvi, D.M. Rizzo, T. Simpatico and R. Snapp, (2007) “Quality of life: An approach

integrating opportunities, human needs, and subjective well-being", *Ecological Economics*, Vol. 61, Issues 2-3, Pages 267-276.
(http://www.sciencedirect.com/science/article/B6VDY-4JDN6F9-6/2/79e011309e81356e5c38c5f2863df485)

Cox T.F., Cox M.A.A. (1994) *Multidimensional scaling*, Chapman & Hall, London.

Cummins, R.A. (1993) *Comprehensive Quality of Life Scale for Adults (ComQol-A4)*, Deakin University, School of Psychology, Melbourne.

De Vellis R. (1991) *Scale devolopment. Theory and Applications*, Applied Social Research Methods Series, vol. 26, SAGE Publications, London.

Diamantopoulos, A. and J.A.A. Siguaw (2006) "Formative versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration" in *British Journal of Management*, Vol. 17, No. 4, pp. 263-282.

Diamantopoulos, A. and H.M. Winklhofer (2001) "Index Construction with Formative Indicators: An Alternative to Scale Development", *Journal of Marketing Research*, Vol. 38, No. 2, pp. 269-277.

Diener E. (1984) "Subjective well-being", *Psychological Bulletin*, n. 95, pp. 542-575.

Diener E., E. Suh (1997) "Measuring quality of life: economic, social, and subjective indicators", *Social Indicators Research*, vol. 40, 189–216.

Drewnowski J. (1970), *A Planning Model for Social Development*, Unrisd, (Report n.70.5), Geneva 1970.

Drenowski J., Wolf, S. (1966) *The Level of Living Index*. Geneva: United Nations Research Institute for Social Development.

Easterlin R.A. (1974) "Does economics growth improve the human lot? Some empirical evidence", in P.A.David, M.W. Reder (eds), *Nations and households in economic growth; essays in honor of Moses Abramowitz*, Academic Press, New York, pp.89-125

Edward W., Newman J.R. (1982) *Multiattribute Evalutation*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-026, Newbury Park, CA:Sage.

Engel, U. and J. Reinecke: 1996, Analysis of Change (Walter de Gruyter, Berlin – New York).

Epidemiological Bullettin (2002), "Introduction to Social Epidemiology", *Epidemiological Bullettin*, Vol. 23, No. 1. http://www.paho.org/English/sha/be_v23n1-socialepi.htm

Erikson (1993) "Description of inequality: the Swedish approach to welfare research" in M.C: Nussbaum & A.K. Sen (Eds.) *The quality of life*, Oxford University Press, Oxford.

Eurostat (2000) *Definition of quality in statistics* and *Standard quality report*, Eurostat.

Firebaugh, G. (1997) *Analyzing Repeated Surveys*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-115, Newbury Park, CA:Sage.

Frisch, M.B. (1998) "Quality of life therapy and assessment in health care", in *Clinical Psychology: Science and Practice*, n. 5, 19–40.

Ghiselli E.E. (1964) *Theory of Psychological Measurement*, McGraw-Hill, New York-London.

Giovannini E. (2007) *The role of statistics in a globalised world: risks and challenges, paper* presented at the DGINS (Directors-General of the National Statistical Institutes) Conference, 20–21 September 2007, Budapest, Hungary.

Glenn, N.D. (1977) *Cohort analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-005, Newbury Park, CA:Sage.

Global Water Partnership - Technical Committee (2004) *Monitoring and evaluation indicators for Integrated Water Resources Management strategies and plans*, http://www.gwpforum.org/gwp/library/Tec_brief_3_Monitoring.pdf

Goldstein G. (1999) *Multilevel Statistical Models*, Arnold Publisher. http://www.arnoldpublishers.com/support/goldstein.htm

Greenley, J.R., Greenberg, J.S., Brown, R. (1997) "Measuring Quality of life: a new and practical survey instrument", in *Social Work*, n. 42, 244–254.

Guilford J.P. (1954) *Psychometric methods*, McGraw-Hill, New York-London.

Hagerty, M.R., Land, K.C. (2007) "Constructing Summary Indices of Quality of Life: A Model for the Effect of Heterogeneous Importance Weights", *Sociological Methods and Research*, Vol. 35, n. 4, 455-496.

Hair J.F. jr., Anderson R.E., Tatham R.L., Black W.C. (1998, 5th ed.) *Multivariate Data Analysis*, Prentice-Hall Inc., Upper Saddle River, New Jersey.

Hambleton R.K., Swaminathan H., Rogers H.J. (1991) *Fundamentals of item response theory*, Measurement Methods for the Social Sciences series, vol. 2, SAGE Publications, London.

Headey B., R. Veenhoven, A. Wearing (1991) "Top-down versus Bottom-up Theories of Subjective Well-being", *Social Indicators Research*, vol. 24.

Horn R.V., (1993) *Statistical Indicators*, Cambridge University Press, Cambridge.

Hox J.J. (1995) *Applied Multilevel Analysis*, TT-Publikaties, Amsterdam.

Jamison D., M. Sandbu (2001) "WHO ranking of health system performance" in *Science*, 293, 1595-1596. http://lgst.wharton.upenn.edu/sandbu/Documents/Science2001.pdf#search=%22WHO%20ranking%20of%20health%20system%20performance%20filetype%3Apdf%22

Kalmijn W.M., Veenhoven R. (2005) "'Measuring inequality of happiness in nations'. In search for proper statistics", *Journal of Happiness Studies*, vol. 6, n. 4, pp. 357-396.

Kim Jae-On, Mueller C.W. (1989a) *Introduction to Factor Analysis: What it is and how to do it*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-013, Newbury Park, CA:Sage.

Kim Jae-On, Mueller C.W. (1989b) *Factor Analysis: Statistical Methods and Practical Issues*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-014, Newbury Park, CA:Sage.

Kozma A., S. Stones, M.J. Stones, T.E. Hannah, K. McNeil (1990) "Long and short term affective states in happiness", *Social Indicators Research*, vol. 22.

Krieger N. (2002) "A Glossary for Social Epidemiology", *Epidemiological Bullettin*, Vol. 23, No. 1. http://www.paho.org/English/sha/be_v23n1-glossary.htm

Kruskal J.B., Wish M. (1978) *Multidimensional Scaling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-011, Newbury Park, CA:Sage.

Lance C.E., A.G. Mallard, A.C. Michalos (1995) "Tests of the Causal Directions of Global-Life Facet Satisfaction Relationships", *Social Indicators Research*, vol. 34.

Land K. C. (1971) "On the definition of social indicators" in *American Sociologist*, n. 6, 322-325.

Land K. C. (1975) "Social indicator models: an overview", in Land K. C., Spilerman S., *Social indicator models*, New York, Russel Sage Foundation.

Land K. C. (2000) "Social indicators", in Edgar F. Borgatta and Rhonda V. Montgomery (eds.), *Encyclopedia of Sociology*, Revised Edition. New York: Macmillan.

Lazarsfeld P.F.(1958) "Evidence and Inference in Social Research", in *Daedalus*, Vol. 87, No. 4, On Evidence and Inference (Fall, 1958), pp. 99-130.

Lazarsfeld, P.J. and N.W. Henry (1968) *Latent Structure Analysis* , Houghton M. Company, Boston.

Lis A., Sambin M. (1977) *Analisi dei cluster*, CLEUP, Padova.

Little R.J.A., D.B. Rubin (1987) *Statistical Analysis with Missing Data*, John Wiley & Sons, Wiley Series in Probability and Mathematical Statistics, New York.

Long J.S. (1993a) *Confirmatory Factor Analysis. A Preface to LISREL*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-033, Newbury Park, CA:Sage.

Long J.S. (1993b) *Covariance Structure Models. An Introduction to LISREL*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-034, Newbury Park, CA:Sage.

Lord F.M. (1974) "Estimation of Latent Ability and Item Parameters when there are Omitted Responses" in *Psychometrika*, 39.

Lord F.M. (1984) "Standard Errors of Measurement at Different Ability Levels" in *Journal of Educational Measurement*, 21.

Louviere J.J. (1988) *Analyzing decision making: metric conjoint analysis*, Sage, Newbury Park

Ludlow L.H., Haley S.M. (1995) "Rasch Model Logits: Interpretation, Use and Trasformation" in *Educational and Psychological Measurement*, Sage Periodicals Press, London, Vol. 55, N. 6, pp. 967-975.

Lyubomirsky, S., King, L. A., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131,* 803-855.

Maggino F. (2003) *Method effect in the measurement of subjective dimensions*, Firenze University Press, Archivio E-Prints, Firenze.

Maggino F. (2004a) *La misurazione nella ricerca sociale. Teorie, strategie, modelli*, Firenze University Press, Archivio E-Prints, Firenze.

Maggino F. (2004b) *I modelli di scaling. Confronto tra ipotesi complesse per la misurazione del soggettivo*, Firenze University Press, Archivio E-Prints, Firenze.

Maggino F. (2005a) *L'analisi dei dati nell'indagine statistica*, Firenze University Press, Firenze.

Maggino F. (2005b) *The Importance of Quality-Of-Life Dimensions in Citizens' Preferences: An Experimental Application of Conjoint Analysis*, Firenze University Press, Archivio E-Prints, Firenze.

Maggino F. (2007) *La rilevazione e l'analisi statistica del dato soggettivo*, Firenze University Press, Firenze.

Maggino F. (2008a) *Choice of weights for subjective variables*, invited paper at the "Conference on Composite Score" (ESADA, Universitat Ramon Llull, Barcelona, 14-15 February 2008) – chair: prof. W.Saris.

Maggino F. (2008b) *Towards more participative methods in the construction of composite indicators. The case of obtaining weights: from "objective" to "subjective" approaches*, invited paper presented at the International Seminar on "Involving citizens/communities in measuring and fostering well-being and progress: towards new concepts and tools", Council of Europe, (Strasbourg, 27-28 November, 2008)

Maggino F. (2009) *Towards more participative methods in the construction of social indicators: survey techniques aimed at determining importance weights*, paper to be presented at the 62nd conference of the World Association for Public Opinion Research "Public Opinion and Survey Research in a Changing World" (Swiss Foundation for Research in Social Sciences – University of Lausanne – 11-13 September 2009 – Lausanne – Switzerland).

Maggino F. (forthcoming) *Obtaining weights: from objective to subjective approaches in view of more participative methods in the construction of composite indicators*.

Maggino F. & E. Ruviglioni (2008a) *Choice of subjective weights for subjective variables. Identifying subjective/individualized weights for comparing well-being among group and individuals*, paper presented at the session on "Social Indicators" organized by Heinz-Herbert Noll (GESIS-ZUMA Mannheim, Germany) at the VII International Conference on Social Science Methodology (1-5 September 2008, Campus di Monte Sant'Angelo – Naples – Italy).

Maggino F. & E. Ruviglioni (2008b) *Choice of subjective weights for subjective variables. Identifying subjective/individualized weights for comparing well-being among group and individuals*, paper presented at the IV International Conference on Quality of Life Research – QOL 2008 – "Quality of Life Improvement through Social Cohesion", Conference organized by the Department of Statistics of Wroclaw University of Economics (15-18 September 2008 – Wroclaw – Poland).

Maggino F. & E. Ruviglioni (2009) *Obtaining weights: from objective to subjective approaches in view of more participative methods in the construction of composite indicators*, paper presented at the Seminar on "*New Techniques and Technologies for Statistics (NTTS)"* – EUROSTAT (18-20 February 2009 – Brussels – Belgium).

Mallard A.G.C., C.E. Lance, A. Michalos (1997) "Culture as a moderator of overall life satisfaction relationships", *Social Indicators Research*, vol. 40, 259-284.

Malhotra N.K. (1996) *Marketing research: an applied orientation*, Prentice-Hall International, Inc., Englewood Cliffs, New Jersey

Marradi A. (1981) "Factor analysis as an aid in the formation and refinement of empirically useful concepts", in E.F. Borgatta - D.J.Jackson (ed.), *Factor Analysis and Measurement in Sociological Research*, SAGE Publications, London.

Maslow A. H. (1954) *Motivation and Personality*, Harper and Brothers, New York.

Max-Neef, M. (1992) "Development and human needs" in Ekins, P., Max-Neef, M. (Eds.) *Real life Economics: Understanding Wealth Creation*, Routledge, London, pp. 97–213.

McDonald R.P. (1989) "Future Directions for Item Response Theory", *International Journal of Educational Research*, 13(2).

Menard, S. (1991) *Longitudinal Research*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-076, Newbury Park, CA:Sage.

Michalos A. (1985) "Multiple Discrepancies Theory", *Social Indicators Research*, vol. 16.

Michalos A. (1992) "Use and Abuses of Social Indicators" in *Sinet*, n. 32.

Michalos A. (2008) "Education, Happiness and Wellbeing", *Social Indicators Research*, vol. 87, n. 3.

Nardo M., M. Saisana, A. Saltelli and S. Tarantola (EC/JRC), A. Hoffman and E. Giovannini (OECD) (2005a) *Handbook on Constructing Composite Indicators: Methodology and Userguide*, OECD, Statistics Working Paper.Nardo M., Saisana M., Saltelli A. and Tarantola S. (2005b) *Tools for Composite Indicators Building*, European Commission, EUR 21682 EN, Institute for the Protection and Security of the Citizen, JRC Ispra, Italy.

Netemeyer R.G., Bearden W.O., Sharma S. (2003) *Scaling Procedures: Issues and applications*, Sage Publications, Thousand Oaks.

Noll H.-H. (1996) *Social Indicators and Social Reporting: The International Experience*, http://www.ccsd.ca/noll1.html

Noll H.-H. (2004) *Social indicators and indicators systems: tools for social monitoring and reporting*, paper presented at OECD, World Forum "Statistics, knowledge and policy", Palermo, 10-13 November 2004.

Noll, H.-H. (2009) *Measuring and Monitoring the Quality of Life*, Lecture at the Università degli Studi di Firenze, April 23-24, 2009. http://www.gesis.org/forschung-lehre/veranstaltungen/veranstaltungs-archiv/zentrum-fuer-sozialindikatorenforschung/quality-of-life/

Nussbaum, M., J. Glover (1995) *Women, Culture, and Development: A Study of Human Capabilities*, Oxford University Press, Oxford.

Organisation for Economic Cooperation and Development (2007) *Istanbul Declaration*, II OECD World Forum on "Statistics, Knowledge and Policy", June 27-30 June, Istanbul (Turkey). http://www.oecd.org/document/51/0,3343,en_21571361_31938349_37115187_1_1_1_1,00.html

Patel S., M. Hiraga, and L. Wang (World Bank) D. Drew and D. Lynd (Unesco) (2003) *A Framework for Assessing the Quality of Education Statistics*, World Bank – Development Data Group and UNESCO – Institute for Statistics.

Rampichini C., Schifini S. (1998) "A Hierarchical Ordinal Probit model for the analysis of life satisfaction in Italy", *Social Indicators Research*, vol. 44, n. 1, 5-39.

Rasch G. (1960) *Probabilistic Models for Some Intelligence and Attainment Tests*, Copenhagen: Danish Institute for Educational Research.

Rubin D.B. (1987) *Multiple Imputation for Nonresponse in Survey*, John Wiley & Sons, Wiley Series in Probability and Mathematical Statistics, New York.

Sadocchi S. (1981) *Manuale di analisi statistica multivariata per le scienze sociali*, Franco Angeli Editore,

Milano.
Saisana M., A. Saltelli, S. Tarantola (2005) “Uncertainty and sensitivity techniques as tools for the analysis and validation of composite indicators”, *Journal of the Royal Statistical Society* A 168(2): 1-17.
Saltelli A., S. Tarantola, F. Campolongo, M. Ratto (2004) *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*, John Wiley & Sons publishers. http://webfarm.jrc.cec.eu.int/uasa/doc/forum-tutorial/WU082-FM.pdf
Saris W.E., Stronkhorst L.H. (1990) *Causal Modelling in Nonexperimental Research. An introduction to the LISREL Approach*, Sociometric Research Foundation, Amsterdam.
Scherpenzeel A., W. Saris (1996) “Causal Direction in a Model of Life Satisfaction: the top-down/bottom-up controversy”, *Social Indicators Research*, vol. 38.
Sen A. (1993) “Capability and Well-Being”, in M. Nussbaum and A. Sen (eds.) *The Quality of Life*, Clarendon Press, Oxford, p. 30-53.
Sharpe A., J. Salzman (2004) *Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being*, Center for the Study of Living Standards, Ottawa, CAN.
Sijtsma K., Molenaar I.W. (2002) *Introduction to Nonparametric Item Response Theory*, Measurement Methods for the Social Sciences series, vol. 5, SAGE Publications, London.
Sinden, J.A. (1982) “Application of Quality of Life Indicators to Socioeconomic Problems: An Extension of Liu’s Method to Evaluate Policies for 26 Australian Towns”, *American Journal of Economics and Sociology*. 41 (4). October, 1982, 401-420.
Sirgy M.J., A.C. Michalos, A.L. Ferriss, R.A. Easterlin, D. Patrick and W. Pavot (2006) “The Quality-of-Life (QOL) Research Movement: Past, Present, and Future”, *Social Indicators Research*, vol. 76, n.3, 343-466.
Stiglitz J. E., A. Sen & J.-P. Fitoussi eds (2009) *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Paris. http://www.stiglitz-sen-fitoussi.fr/en/index.htm
Stones M.J., T., T. Hadjistavvropoulos, H. Tuuko, A. Kozma (1995) “Happiness has Tratlike and Statelike Properties: a reply to Veenhoven”, *Social Indicators Research*, vol. 36.
Stoker T M (1993) ‘Empirical approaches to the problem of aggregation over individuals’, *Journal of Economic Literature*, XXXI, 1827-1874.
Sullivan J.L., Feldman S. (1981) *Multiple Indicators. An Introduction*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-015, Newbury Park, CA:Sage.
Swaminathan H., J.A. Gifford (1982) "Bayesian Estimation in the Rasch Model" in *Journal of Educational Statistics*, 7.
Swaminathan H., J.A. Gifford (1985) "Bayesian Estimation in the Two-Parameter Logistic Model" in *Psychometrika*, 50.
Swaminathan H., J.A. Gifford (1986) "Bayesian Estimation in the Three-Parameter Logistic Model" in *Psychometrika*, 51.
Tarantola S., J. Jesinghaus J., M. Puolamaa M. (2000) “Global sensitivity analysis: a quality assurance tool in environmental policy modeling” in *Sensitivity Analysis* (eds A. Saltelli, K. Chan, M. Scott) pp. 385-397. New York: John Wiley & Sons.
Torgerson W.S. (1958) *Theory and Methods of Scaling*, John Wiley & Sons, Inc., New York, London, Sydney.
Tucker, L. & MacCallum R. (1993) Exploratory Factor Analysis, Book Manuscript, Retrieved in 2008, from: http://www.unc.edu/~rcm/book/factornew.htm
United Nations Development Programme (2007) *Human Development Report 2007*, Oxford University Press, New York.
Veenhoven R. (2002) “Why social policy needs subjective indicators”, *Social Indicators Research*, vol. 58, n. 4, pp. 33-45.
Veenhoven R. (2005) “Inequality of happiness in nations”, *Journal of Happiness Studies*, vol. 6, n. 4, pp. 351-355.
Veenhoven R. (2009) “Well-Being *in* Nations and Well-Being *of* Nations. Is There a Conflict Between Individual and Society”, *Social Indicators Research*, vol. 91, n. 1, pp. 5-21.
Werts C.E., Linn R.L., Jöreskog K.G. (1974) ‘Quantifying Unmeasured Variables’ in Blalock H.M. jr., ed., *Measurement in the Social Sciences. Theories and Strategies*, Aldine Publishing Company, Chicago.
Yoon K.P., Hwang Ching-Lai (1995) *Multiple Attribute Decision Making*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-104, Thousand Oaks, CA:Sage.
Zapf W. (1975) Le système d’indicateurs sociaux: approches et problèmes", *Revue Internationale des Sciences Sociales*, Vol. XXVII, n. 3.
Zapf W. (1984) “Individuelle Wohlfahrt: Lebensbedingungen und Wahrgenommene Lebensqualität“, in W. Glatzer e Zapf W. (eds.) *Lebensqualität in der Bundesrepublik*, Frankfurt a. M. – New York, Campus, pp. 13-26.