Measuring the Quality of Life and the Construction of Social Indicators

10

Filomena Maggino and Bruno D. Zumbo

Introduction

Complexity and the Process of Measurement

As is evident from even a cursory review of the research literature and current practices, the well-being of societies represents a multidimensional concept that is difficult and complex to define. Its quantitative measurement requires a multifaceted approach and a multipurpose methodology that is a mix of many approaches and techniques founded upon statistical indicators. The main notion that should be kept in mind in order to measure societal well-being from a quantitative perspective, using statistical indicators, is *complexity*. The complexity stems from the reality to be observed, and affects the measuring process and the construction of the indicators. Therefore, complexity should be preserved in analyzing indicators and should be correctly represented in telling stories from indicators.

In considering the topics we wished to include in this chapter, we chose to be inclusive with an eye toward integrating a vast body of methodological literature. Our aim in this chapter is to disentangle some important methodological approaches and issues that should be considered in measuring and analyzing quality of life from a quantitative perspective. Due to space limitations, relative to the breadth and scope of the task at hand, for some issues and techniques, we will provide details, whereas for others, more general integrative remarks. The chapter is organized as follows. The first section (comprised of three sub-sections) deals with the conceptual definitions and issues in developing indicators. The aim of this first section, like the chapter as a whole, is to provide a framework and structure. The second section (comprised of three sub-sections) is an overview of the analytic tools and strategies. The third, and final, section (comprised of two sub-sections) focuses on methodological and institutional challenges.

Given that our primary purpose is to catalog and organize the complex array of foundational methodological issues, analytic tools, and strategies, we will make extensive use of figures and tables whose primary purpose is to list and contrast concepts, issues, tools, and strategies. Table 10.1 provides an overview of the questions and issues one faces one when one is dealing with the first stage in developing indicators: the conceptual definitions, framework and structure. Table 10.2 provides an overview of the questions and issues surrounding the analytic tools and strategies. Tables 10.1 and 10.2 also provide a type of "advanced or graphic organizer" for the first and second sections of the chapter and as such are meant to help the reader catalog and retain some order in the complex array of ideas, tools, and strategies found when one aims to measure quality of life and one considers the construction of social indicators.

F. Maggino (✉)
Università degli Studi di Firenze, Florence, Italy
e-mail: filomena.maggino@unifi.it

B.D. Zumbo
University of British Columbia, Vancouver, Canada
e-mail: bruno.zumbo@ubc.ca

K.C. Land et al. (eds.), *Handbook of Social Indicators and Quality of Life Research*,
DOI 10.1007/978-94-007-2421-1_10, © Springer Science+Business Media B.V. 2012

Table 10.1 An overview of the questions and issues when dealing with conceptual definitions

Conceptual definition (framework and structure)	
How can the complexity be conceptually designed?	
1. Hierarchical design	Indicators should be developed through a *logical modeling 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 basic indicators
	The hierarchical design is completed by defining the relationships between:
	• *Each variable and the corresponding indicators*. These relations define the *model of measurement*
	• *Basic indicators*. 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*. These relations are defined in the ambit of the conceptual model and identify the structural pattern. The analysis of this kind of relationships is accomplished by *modeling the indicators*
How can the indicators be conceptually defined?	
2. Model of measurement	The model of measurement can be conceived through two different conceptual approaches:
	• *Reflective approach*. the basic 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 • *Formative approach.* a latent variable construct can be defined as being determined by (or *formed* from) a number of basic indicators
How can the indicators be consistently organized?	
3. system of indicators	A *system of indicators* represents the fulfillment of the conceptual framework and allows an organizational context to be defined in order to allow methodological supports and structured and systematic data management in a long-term longitudinal perspective
	This is particularly demanding with reference to subjective data, which require a great use of resources (beyond a solid survey research methodology)

Developing Indicators, Conceptual Definition, Framework and Structure

An Introduction to This Section: Developing and Managing Indicators

The *process of measurement* in the social sciences requires a robust conceptual definition, a consistent collection of observations, and a consequent analysis of the relationship between observations and defined concepts. The measurement objective that relates concepts to reality is represented by *indicators*. From this perspective, an indicator is not a simple crude bit of 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 a sphere of operation or influence (i.e., the ambit) of a conceptual model and are connected to a defined aim. As such, indicators can be considered *purposeful statistics* (Horn 1993). As Land (1971, 1975) reminds us, a statistical index can be considered an “indicator” when: (1) it represents a component in a model concerning a social system, (2) it can be measured and analyzed 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), and (3) it can be aggregated with other indicators or disaggregated in order to specify the model.

Far too often, however, indicators are developed and used without consideration of the conceptual

Table 10.2 An overview of the questions and issues surrounding the analytic tools and strategies

II. Analytic tools and strategies		
The consistent application of the hierarchical design actually leads to a parceled picture, with reference to the conceptual model, and consequently produces a compound data structure. In order to reconstruct a meaningful and interpretable picture, data needs to be managed pursuing different technical goals:		
– Reducing data structure – Combining indicators – Modeling indicators		
The different analytic and technical strategies to be adopted in these respects constitute a "composite" *process*, carried out through subsequent/consecutive steps (MULTISTAGE) and different/alternative analytic approaches (MULTITECHNIQUE).		
How can the observed picture be reconstructed?		
1. Reducing data structure	Since data structure shows:	
	– Basic indicators, stratified with reference to the identified variables – Cases, stratified with reference to the standard unit (e.g., Individuals vs. Cities)	
	Data reduction has the following goals:	
	(1) *Reconstructing the conceptual variables by aggregating basic indicators* through different logics:	
	a. Aggregating basic indicators referring to the same variable (*reflective logic*);. b. Aggregating indicators creating a new conceptual variable (*formative logic)*,	
	Since both kinds of aggregation process are carried out at micro level (e.g., for each individual), the following reduction step needs to be accomplished	
	(2) *Defining macro-units by aggregating single 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 analytic approaches	
	Traditional approach:	The two goals are usually carried out through traditional and consolidated analytic approaches and based upon linear statistics
	Alternative approach: This stage allows the system of indicators to be operational.	New methodologies have been proposed allowing discrete ordinal data to be dealt with, especially when evaluation, comparisons and rankings are of concern. Such methodologies are based on Partially Ordered SEt Theory (POSET theory), part of Discrete Mathematics that offers many tools and results to explore and analyze the structure of discrete datasets, like that of interest in the present study. Posets of finite cardinality can be conveniently depicted by means of certain directed acyclic graphs, called Hasse diagrams

(continued)

Table 10.2 (continued)

II. Analytic tools and strategies	
How can the whole picture be simplified and shown?	
2. Combining indicators	Sometimes, also after the data reduction process has been accomplished, the complexity of the system of indicators may require particular combinations 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 multidimensional phenomena to be transformed into uni-dimensional – Allow situations to be more easily compared across time – Compare cases (e.g., Nations) in a transitive way (ranking). – Allow 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?
	Depending on the particular need, different approaches can be adopted:
	– *Dashboards* allow indicators to be represented in a single graphical solution and the complex relationships among indicators to be communicated
	– *Benchmarking* through
	◦ *Composite indicators* can represent useful approaches aimed at summarizing indicators ◦ *Partial order sets*: new approaches based upon the *POSET* theory can be fruitfully applied through getting over the methodological critical aspects shown by composite indicators
How can the whole picture be explained?	
3. Modeling indicators	This stage is aimed at analyzing different aspects of the defined model (e.g., objective and subjective indicators) in order to find explanations by identifying the proper analytic approaches

Table 10.3 A structured plan to aid in developing and managing indicators

CONCEPTUAL DEFINITION (FRAMEWORK AND STRUCTURE)	Developing indicators	1. through a **hierarchical design**
		2. by defining a **model of measurement**
		3. by developing a **system of indicators**
ANALYTIC TOOLS AND STRATEGIES	Managing indicators in terms of	4. **data structure reduction**
		5. **indicators combination**
		6. **indicators modelling**

definition of the phenomenon and a logical cohesion of the conceptual definition and the analytic tools and strategies. In our experiences, the lack of any logical cohesion is often masked by the use and application of sophisticated procedures and methods that can deform reality producing distorted results.

Table 10.3 is an organization tool and structured plan to aid in developing and managing indicators that are able to (1) represent different aspects of the reality, (2) picture the reality in an interpretable way, and (3) allow meaningful stories to be told. We can see in Table 10.3 that the conceptual definition (framework and structure) shapes both how one develops indicators and the analytic tools and strategies. In terms of developing indicators, one does so through a hierarchical design, which leads to defining a measurement model and eventually to developing a system of indicators. Likewise, one manages indicators in terms of reducing the data structure, combining indicators, and modeling the indicators.

Table 10.4 is the advanced organizer for the developing indicators, conceptual definition (framework and structure) section. We can see that there are three sections: (1) hierarchical design which leads to (2) the choice of a measurement model, and eventually to (3) the system of indicators.

Defining the Hierarchical Design

Indicators should be developed, following Lazarsfeld's model (1958), through a *hierarchical design* requiring the definition of the following components: (a) conceptual model, (b) areas, (c) latent variables, (d) basic indicators, and (e) observed variables. We will describe each of these in turn below.

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 wherein they can be applied. In the social sciences, the description of concepts varies according to (1) the researcher's point of view, (2) the objectives of the study, (3) the applicability of the concepts, and (4) the sociocultural, geographical, and historical context. Examples include concepts such as health, education, well-being, income, production, and trade.

The process of conceptualization allows us to identify and define the:

(a) Model aimed at data construction
(b) Spatial and temporal ambit of observation
(c) Aggregation levels (among indicators and/or among observation units)
(d) Approach aimed at aggregating the basic indicators and the techniques to be applied in this perspective (weighting criteria, aggregation techniques, etc.)
(e) Interpretative and evaluative 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 time-consuming and exacting, especially with complex constructs, and requires a systematic review and analysis of the relevant research literature.

Latent Variables

Each variable represents one of the aspects to be observed and confers an explanatory relevance onto the corresponding defined area (see Zumbo 2007, 2009). The identification of the latent variable is founded on theoretical and statistical assumptions (e.g., homogeneity, dimensionality) as empirical commitments so that the defined variable can reflect the nature of the considered phenomenon consistently

Table 10.4 On overview of developing indicators, conceptual definition (Framework and structure)

	Components	Questions	Components´ definition
1. Hierarchical design	Conceptual model	What is the phenomenon to be studied?	It defines the **phenomenon**, its domains and its general aspects.
	Investigating areas/pillars	What aspects define the phenomenon?	Each area represents each **aspect** allowing the phenomenon to be specified consistently with the conceptual model
	Variables	What elements have to be observed?	Each variable represents each **element** that has to be observed in order to define the corresponding area. The variable can be observable - directly → *observable variable* - indirectly (through indicators) → *latent variable*
2. Model of measurement			Relationship between latent variables and indicators
	Basic indicators (b.i.)	In which way has each element to be measured?	Each indicator 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.
3. System of indicators			

with the conceptual model. However, even if we are able to identify a variety of diverse variables, we have to accept the possibility that maybe no set of variables can perfectly capture the concept to be measured (e.g., social or economic well-being; Sharpe and Salzman 2004).

Basic Indicators

Each basic indicator (e.g., an 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 (De Vellis 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 relatively 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 rarely possible to define the direct correspondence between one latent variable and one indicator. In other words, the variable is not always directly observable through

[1] In specific cases, some variables can be directly measured (e.g., some objective information). In this case, variable and indicator coincide.

[2] In data analysis, indicators/items are technically defined "variables"; consequently, these are conceptually different from "latent variables."

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 error[3]
(b) *Accuracy* (*validity*), since the chance that one single indicator can describe one latent complex variable is highly dubious and questionable
(c) *Relationships* with the other variables
(d) *Discriminating* and *differentiating* among observed cases, for example, individuals

This is precisely why, in many cases, the presence of complex latent variables requires the definition of several basic indicators. This can be done by adopting the *multiple indicators approach,* which considers the multiple indicators as *multiple measures* (Sullivan and Feldman 1981). Multiple indicators contribute to the measurement of the major aspects of the variable because each basic indicator may correspond to one particular aspect of the latent variable. This approach allows for the inherent variability in the defined latent variable to be covered. In addition, 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 basic indicators referring to one variable represents a *set of indicators*, while the complete group of indicators defining an area is called a set of *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 the micro and macro level.

Observed Variables

Some variables can be observed and directly measured. Consequently, they do not need any indicator (e.g., age, level of education).

Defining the Model of Measurement

The model of measurement can be conceived through two different conceptual approaches (Blalock 1964; Diamantopoulos and Siguaw 2006): models with reflective or formative indicators. Figure 10.1 is a statistical description of the two models.

Model of reflective indicators. This model is also sometimes referred to as the *top-down* explanatory approach. In this case, latent variables 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.[4] Structural relationships are identified among latent constructs by statistically relating covariation between the latent variables and the observed variables or indicators, measuring these latent, unobserved variables. 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 latent variable 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 variable to the indicators.

Models with formative indicators. This model is sometimes referred to as 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 latent variable can be defined as being determined by (or *formed* from) a number of indicators. In this case, causality flows from the indicator to the latent variable. A classic example of formative indicators is socioeconomic 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

[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.

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

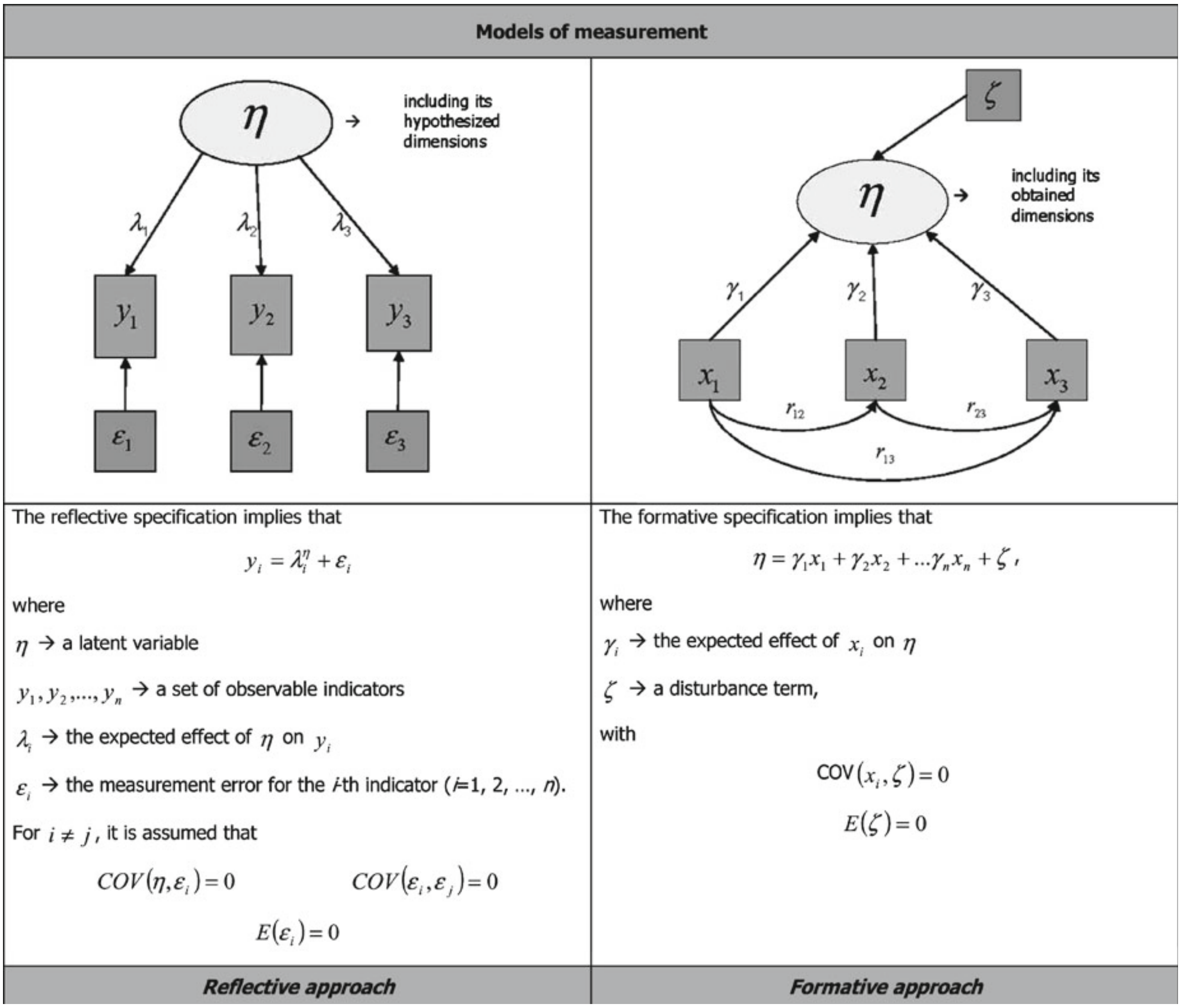

Fig. 10.1 Description of formative and reflective measurement models

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. As Zumbo (2007) notes, the reflective model is most often cast as factor analysis whereas the formative models as principal components analysis.

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. As Zumbo (2007) states, there are no empirical tests of whether a latent variable is reflective or formative; the exception is the vanishing tetrads test of Bollen and Teng (2000). It should be noted that, although it is often presented as evidence, computing a principal components analysis (PCA) is not sufficient evidence that one has formative indicators, nor does fitting a factor analysis model provide sufficient evidence to claim one has reflective indicators—that is, as is often evidenced in practice, both PCA and factor analysis may fit the same data equally well. Bollen and Lennox (1991) suggest that a good place to start, and often the only thing available, is a literal thought experiment.

Table 10.5 Possible outcomes in deciding between reflective and formative indicators

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

Zumbo (2007) added that one can also supplement this thought experiment with a content validation study wherein one asks subject matter experts to consider and rate whether the items (or indicators) are effects or causes; that is, whether the variable is a measure or index, respectively. One can build on the methodologies described for content validity by incorporating questions about whether an item should be considered a cause or effect indicator using methodology in content validity including the coefficients, designs, etc. Also, one could investigate the source of the decision of effects vs. causes by talk-aloud protocols and/or by conducting multidimensional scaling of the subject matter experts' judgments. These approaches aid one in determining whether one has reflective or formative indicators. What Zumbo was suggesting is an extension of Bollen and Lennox's thought experiment to include data from subject matter experts.

In deciding between formative and reflective indicators, four different situations can be theoretically identified (Diamantopoulos and Siguaw 2006), as represented in Table 10.5.

Two outcomes are desirable and correspond to the correct adoption of the measurement perspective (operationalization) following the correct conceptualization of the construct of interest. The other two outcomes correspond to wrong choices. In particular, two types of error may 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 operationalization (a synthetic indicator construction procedure is adopted in place of a scaling model). This error can lead to identification problems.

Developing a System of Indicators

The application of the hierarchical design, strictly connected to the definition of a proper conceptual framework, leads to the consistent definition of a set of indicators (single and synthetic 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. If the structure is systematized also in time perspective, the set of indicators can be characterized as a *system of indicators*.

The basic requirements defining a system of indicators are synthesized by Noll (2004) and depicted in Table 10.6.

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. There are several risks one may face in developing a system of indicators. That is, the set of identified indicators may be poor (i.e., limited) or poorly defined and unable to fit the conceptual framework, goals, and objectives; also, the data are not reliable; the indicators may not allow local realities to be compared (e.g., explanatory variables are not measured); and the system's results are not able to produce effects on the strategic, decision, and planning processes.

Systems of indicators can be utilized for both scientific and operational (e.g., public policy) goals. In particular, systems of indicators turn out to be useful whenever a process involves a composite evaluation (policy and technique). In this sense, a system of indicators can represent an important and valid support to individuals 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.

Main Functions

Systems of indicators can be distinguished according to the functions for which they have been created (Berger-Schmitt and Noll 2000; Land 2000; Noll 1996). The different functions, illustrated in Table 10.7, can be thought of in cumulative terms since each of them requires the previous one/s.

Crucial Elements

The main elements that make a system of indicators work are (1) aims, (2) structure, (3) analytic approaches, and (4) the interpretative and evaluative models (Noll 1996; Berger-Schmitt and Noll 2000).

Table 10.6 Noll's requirements defining a system of indicators

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 analyzed 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 and exportability*. The system has to allow its generalization to other similar context – *Joint development*. The system has to be developed in a shared way by all the "actors"	
Formal criteria to respect:	– Comprehensiveness – Consistency	– Nonredundancy – Parsimoniousness
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*	

Aims

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

1. *Conceptual aims* (*goals*) that represent broad statements concerning what has to be achieved or what is the problem to be faced. Usually goals are placed at a macro level (national, international, etc.).
2. *Operational 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.).
3. *Planning aims* (*actions*) that represent the specific activities identified to accomplish objectives. They can include developments and infrastructural changes in policies, in institutions, in management instruments, etc.

Each goal, objective and action has:

1. Corresponding *targets*, representing those elements allowing each goal, objective and action to find measurable criteria and to define a *timetable*.
2. Corresponding *indicators* defined in order to assess progress towards the target with goals and objectives and the accomplishment of actions[5]; these indicators can be distinguished in Table 10.8.

These indicators can be combined in order to define composite measures (efficacy/efficiency indicators).

Structure

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

1. *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 be implemented, according to local information.
2. *Horizontal*. Data are collected only at one level (e.g., regional) and allow particular observational

[5] Another nonalternative classification distinguishes them 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).

Table 10.7 The various functions of systems of indicators

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 plays an important role of explanation by meeting the need to
	– *Describe* the 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?" – *Analyze* 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 2009; Berger-Schmitt and 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 consequence of the reporting function, increases the probability of reaching some results by allocating resources and planning efficient procedures *ex-ante*. (Cannavò 2009)
	Accounting. A system can represent a useful means of *accounting*, by which it is possible to measure and make systematically available data in order to support decisions concerning the allocation and destination of resources (financial and others)
	In particular, this function allows the development of a system allowing decision makers to (Cannavò 2009):
	– Control ex post the suitability of the defined standards and of the planned resource flows – Evaluate efficiency and correctness of the defined procedures – Test adequacy and actual attainment of results
	Program management and performance evaluation. Systems of indicators represent valid supports to *project management since* they allow specific strategic programs 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 programs, 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 programs, they cannot be 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)
	Assessment. A system can represent valid support to assessment procedures (certification and accountability). In this case, the goal may be to certify or judge subjects (individuals or institutions) by discriminating their performances or to infer functioning of institutions, enterprises, or systems

ambits (environment, education) to be monitored and compared.

3. *Local*. This structure is typically designed in order to support local decision processes. This kind of system is characterized by two levels:
   (a) Internal, when the indicators are aimed at monitoring the internal organization of the level
   (b) External, when the indicators refer to parameters existing at higher levels (e.g., transportation)

Analytic Approaches

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

Interpretative and Evaluative Models

The observed results can be interpreted only according to a specific frame of reference. This can also include particular *standard-values*, which can be defined a

Table 10.8 Indicators and corresponding function

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 levels
– Output/outcome	→	Monitoring direct results of actions
– Impact	→	Monitoring progress and improvement towards goals and objectives achievement

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.

The Indicators in a System

Selection

Different issues need to be addressed when selecting and managing indicators, especially when this is carried out within a complex system allowing for functions such as 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 states, the issues collectively yield over 200,000 possible combinations representing at least that many different kinds of systems (Sirgy et al. 2006). The 15 different issues are:

1. Settlement/aggregation area sizes: e.g., the best size to understand air pollution may be different from the best size to understand crime.
2. 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.
3. Population composition: e.g., analyses by language, sex, age, education, ethnic background, income, etc. may reveal or conceal different things.
4. Domains of life composition: e.g., different domains like health, job, family life, housing, etc. give different views and suggest different agendas for action.
5. Objective vs. subjective indicators: e.g., relatively subjective appraisals of housing and neighborhoods by actual dwellers may be very different from relatively objective appraisals by "experts."
6. Positive vs. 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.
7. Input vs. 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.
8. 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.
9. Measurement scales: e.g., different measures of well-being provide different views of people's well-being and relate differently to other measures.
10. Report writers: e.g., different stakeholders often have very different views about what is important to monitor and how to evaluate whatever is monitored.
11. Report readers: e.g., different target audiences need different reporting media and/or formats.
12. Conceptual model: e.g., once indicators are selected, they must be combined or aggregated somehow in order to get a coherent story or view.
13. Distributions: e.g., because average figures can conceal extraordinary and perhaps unacceptable variation, choices must be made about appropriate representations of distributions.
14. Distance impacts: e.g., people living in one place may access facilities (hospitals, schools, theaters, museums, libraries) in many other places at varying distances from their place of residence.
15. 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 others can be identified. Dealing with these issues is merely a technical problem to be solved by statisticians or information scientists. However, the construction of

Table 10.9 Attributes of quality of an indicator

(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 (1) Professionalism in statistical policies and practices (2) Transparency (3) 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 (1) Clarity of presentations and documentations concerning data and metadata (with reference to the information environment: data accompanied with appropriate illustrations, graphs, maps, and so on, with information on their quality, availability and—eventual—usage limitations) (2) Impartiality of access (3) Pertinence of data (4) 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.

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.

Quality

Many international institutions, such as the World Bank and UNESCO (Patel et al. 2003) and Eurostat (2000) have 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. Tables 10.9 and 10.10, respectively, list the attributes of a good indicator and what a good indicator should be.

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 allowing for quality of statistics. In other words, indicator 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:

1. Legal and institutional environment, allowing
   (a) Conceptual framework to be defined
   (b) Coordination of power within and across different institutions to be framed
   (c) Data and resources to be available for statistical work
2. Quality awareness informing statistical work

Table 10.10 What a good indicator should be

An indicator should be	Clear	Meaningful	*With reference to its capacity and possibility to*	*Define and describe* (*concepts, definitions, and scopes*)	
	Appropriate	Accurate			
	Exhaustive	Well-designed			
	Measurable	Stable		*Observe* unequivocally and stably (in terms of space and time)	*(I) Methodological soundness*
	Reliable	Rigorous		Record by a *degree of distortion as low as possible* (explored through statistical and methodological approaches)	
	Valid	Precise			
	Repeatable	Exact			
	Robust	Faithful			
	Transparent	With ethical standards		*Adhere* to the principle of objectivity in the collection, compilation, and dissemination	*(II) Integrity*
	Consistent	Pertinent		*Reflect adequately the conceptual model* in terms of aims, objectives and requirements underlying its construction (knowing, monitoring, evaluating, accounting, …)	
		Coherent			
	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	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)	*(III) Serviceability*
	Revisionable				
	Well-timed	Timely		Reflect the length of time between its availability and the event or phenomenon it describes	
		Periodic			
	Regular	Punctual		Reflect the time lag between the release date of data and the target date when it should have been delivered	
	Comparable	Disagregable		*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	
	Discriminant	Thrifty			
	Believable	Comprehensible		*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*
	Accessible	Simple			
	Interpretable	Manageable			

Analytic Tools and Strategies

The consistent application of the hierarchical design actually leads to a parceled picture, with reference to the conceptual model, and consequently produces a compound data structure. In order to reconstruct a meaningful and interpretable picture, data needs to be managed pursuing different technical goals: reducing data structure, combining indicators, and modeling indicators

The different analytic and technical strategies to be adopted in these respects constitute a "composite" *process*, depicted in Table 10.11, carried out through subsequent/consecutive steps (multistage—MS) and different/alternative analytic approaches (multitechnique—MT). We discuss each of these strategies in turn below.

Reducing Data Structure

When indicators are developed according to a conceptual framework, dealing with a multidimensional construct and evaluating multiple aspects to be observed at different levels (individual, community, national, and global), the collected data produce a subsequent data structure which turns out to be very complex and needs to be reduced in some way. In particular, the information

Table 10.11 The compositive process of the different analytic and technical strategies

Goals	Stages	Aims	By	Level of analysis	Analytic issues: Traditional approach	Analytic issues: Alternative approach
4. Reducing data structure:	(a-i)	**reconstructing conceptual variables**	aggregating basic indicators	micro	From basic indicators to synthetic indicators through different logics (reflective & formative)	New methodolgies allowing discrete ordinal data to be dealt with, based on Partially Ordered SEt Theory (POSET theory).
	(a-ii)	**defining macro-units**	aggregating single cases		From micro units to macro units, by following different criteria (homogeneity, functionality).	
5. Combining indicators:	(b-i)	**joint representation of indicators**	dashboards	macro	Comparing over time / across units	
	(b-ii)	**benchmarking**	merging indicators		Composite indicators: useful approaches aimed at summarising indicators	POSET theory can be fruitfully applied through getting over the methodological critical aspects shown by composite indicators
6. Modelling indicators:	(c)	**analysis of indicators**	exploring explanations	macro	Different solutions (consistently with conceptual framework)	

Table 10.12 Traditional approach to data reduction

aggregation goes through	the columns	→	indicators	with reference to	each case	in order to obtain	a synthetic value
	the rows	→	units		each indicator		a macro-unit

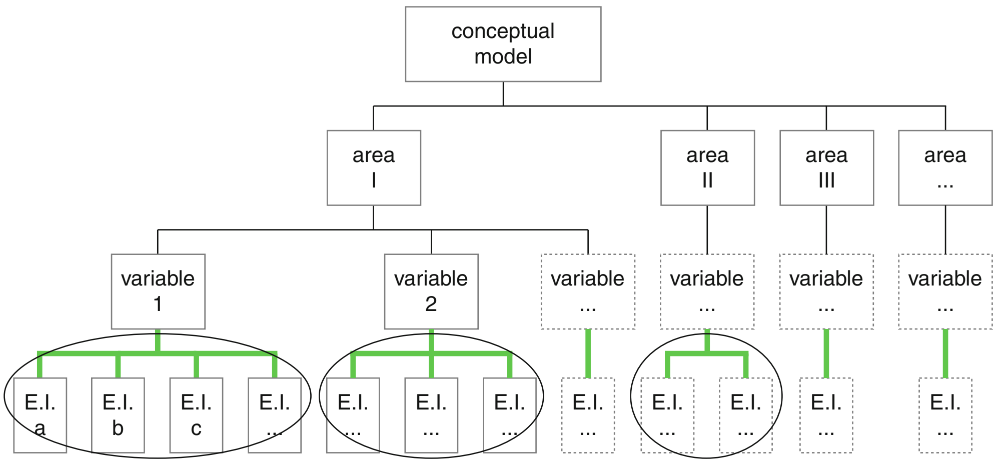

Fig. 10.2 An example in which the indicators that will make up three different synthetic indicators

collected at the micro-level needs to be aggregated at a proper scale (spatial or temporal), in order to accomplish a correct analysis and obtain a composite picture (e.g., national).

With reference to this goal, two different approaches can be identified. While the first one turns out to be very traditional (and known), the second one applies a different analytic approach, quite new with reference to data reduction perspective.

Traditional Approach

In reducing the data structure, the traditional approach proceeds through the following logic (Table 10.12).

Aggregating Indicators and Creating Synthetic Indicators

In order to better manage the complexity of the measured data, analytic 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 particularly important since aggregation of indicators has to be consistently accomplished. In other words, indicators can be aggregated into complex structures through a consistent methodology according to two different criteria: (1) *reflective criterion* and (2) *formative criterion*. In both cases, the condensation of basic 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 by the corresponding latent variable. In Fig. 10.2, one finds the indicators that will make up three different synthetic indicators.

In this context, the traditional distinction between formative and reflective is particularly important because aggregation of indicators has to be consistently accomplished. In other words, indicators can be aggregated in order to define a synthesis through a consistent methodology according to two different criteria: reflective and formative criteria.

1. The reflective criterion

   Since the indicators are seen as functions of the latent variable, 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 and Winklhofer 2001):

   (a) Indicators are interchangeable (the removal of an indicator does not change the essential nature of the underlying construct).
   (b) Correlations between indicators are explained by the measurement model.
   (c) Internal consistency is of fundamental importance: two uncorrelated indicators cannot measure the same construct.
   (d) Each indicator has error term (ε).
   (e) 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, the reflective criterion 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 latent factors. The fundamental equation of the factor model (for *m* indicators) is the following:

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

where

$\sigma_{x_i}^2$ total variance of indicator x_i

$\lambda_{x_i\xi_j}$ factor loading of indicator x_i with reference to latent variable ξ_j

$\delta_{x_i}^2$ uniqueness (specific variance + error) of indicator x_i

2. The formative criterion

   Since the indicators are viewed as causing—rather than being caused by—the latent variable, 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 and Winklhofer 2001):

   (a) The indicators are not interchangeable (omitting an indicator is omitting a part of the construct).
   (b) The correlations between indicators are not explained by the measurement model.
   (c) 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.
   (d) Indicators do not have error terms; error variance is represented only in the disturbance terms (ζ).

As a result, the formative criterion can be accomplished through a statistical approach consistent with a principal components specification, where the latent variable is defined as a linear combination of basic (manifest) indicators:

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

where

η latent variable

γ_i the expected effect of x_i on η

$\eta\zeta$ the disturbance term

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

In both cases, the aggregation of basic indicators, considered multiple measures, produces new synthetic values. Each synthetic indicator tries to re-establish the unity of the defined concept described by the corresponding latent variable.

Aggregating Observed Units and Defining Macro Units

This aggregation perspective aims at condensing values observed at micro/lower levels (usually, individual) to higher levels in order to produce new meaningful units, identified according to different kinds of scales. Generally, the macro units refer to preexistent/predefined partitions, such as identified *groups* (social, generation, etc.), *areas* (geographical, administrative, etc.), and *time periods* (years, decades, etc.).[6]

The aggregation can be accomplished through either an additive or compositional approach. The *additive approach* is characterized by a single-value synthesizing the values observed at micro level; this is usually done by averaging individual values at the level of interest (country, region, social group, and so on). According to the number of involved indicators, the single synthetic value could be represented by a simple descriptive statistical index, univariate (mean, median) or multivariate (centroid). The *compositional approach* is characterized by obtaining macro-units' values by aggregating individual values in a certain number of homogeneous subgroups. This approach is based upon the *homogeneity* criterion: within each level of aggregation (area, group, and so on), individuals' values are aggregated (or averaged) only if cases are homogeneous according to the involved indicators. Each level is then represented by a profile of values, component values (generally proportions or incidences) describing the subgroups. Each subgroup represents a macro unit defined in terms of a *typology*.[7] The sum of component values is constant. Each typology will be considered in the context of the successive higher-level analysis through the component value.

As seen in Table 10.13, in both cases the solution has to be reached consistently with the nature of data (qualitative or quantitative) and by taking into account the number of indicators to be aggregated.

Simultaneous Aggregation of Indicators and Units

Through particular combined analytic processes, the simultaneous aggregation of indicators and cases can be accomplished. These approaches have great potentialities since they simultaneously allow data reduction and synthesis to be reached, simultaneously for both cases and indicators:

(A) *A 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, b).This approach could turn out to be difficult since the identification of homogeneous groups relies on the quality of the synthetic scores previously obtained.

(B) *A 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. The use of a fast alternating least-squares algorithm allows applications to large data sets (Nardo et al. 2005a, b).

Combining Indicators

Joint Representation of Indicators: 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, and by representing them on a color scale (e.g., a green-to-red color scale). Several software programs (free or not) can be used in

---

[6] 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 (e.g., 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).

Other attempts aimed at weighting average values by different criteria can be identified (Kalmijn and Veenhoven 2005; Veenhoven 2005).

[7] Identification of typologies requires particular analytic approaches, allowing homogeneous groups among individual cases to be identified (Aldenderfer and Blashfield 1984; Bailey 1994; Corter 1996; Hair et al. 1998; Lis and 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 MethodsEach analytic approach produces results that vary according to the decisions made in terms of (1) selected indicators, (2) measures used in order to evaluate proximities between individual-points, (3) method used in order to assign individual-points to a group, (4) criterion used in order to determine the number of groups, and (5) criterion used in order to check the interpretability of the groups.

Table 10.13 An overview of aggregation approaches based on the nature of the data

				aggregation approach			
				additive		compositional	
				involved indicators			
				single	multiple	single	multiple
nature of data	qualitative	disjointed	labels	mode		incidences	typologies
	qualitative	ordinal	natural / conventional order	median	L1 – median	incidences	typologies
	quantitative	discrete	natural numbers	median	L1 – median	incidences	typologies
	quantitative	continuous	real numbers	mean	centroide	incidences	Typologies

order to carry out the graphical representation through different images:

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

1. Separated values (values are not aggregated), allowing weak and strong points to be identified.
2. Colors, allowing the analysis of relative performance (value to be displayed relatively to an expected value or a given level/targets).
3. Distributions, allowing assessment indicators' meaningfulness, outliers identification, etc.
4. 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 programs, performances or policies, since:

1. Highly complex systems of indicators can be represented by taking into account the hierarchical design.
2. Easy communications are possible through a catchy and simple graphical representation.
3. Indicators can be related to weights interpreted in terms of:
   (a) *Importance* (reflected by the size of the segments)
   (b) *Performance result* (reflected by the color, interpretable in terms of "good vs. bad")
4. Performances of different cases can be compared.

Of course, a 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 creating composite indicators.

Benchmarking: Merging Indicators

Traditional Approach: Composite Indicators

The previous procedures allow one to reduce the complexity of data by aggregating basic indicators (*construction of synthetic indicators*), and aggregating units/cases (*definition of macro units*).

Although the reduction process has been accomplished, the indicators consistently obtained through the hierarchical design remain a complex system. Sometimes, the complexity of the system of indicators may require indicators allowing measures that are more comprehensive. This need can emerge in order to (Noll 2009):

(a) Answer the call by "policy makers" for condensed information
(b) Improve the chance of getting into the media (compared to complex indicator systems)
(c) Allow multidimensional phenomena to be converted to unidimensional
(d) Allow situations to be compared across time more easily
(e) Compare cases (e.g., nations) in a transitive way (ranking and benchmarking)
(f) Allow clear cut answers to defined questions related to change across time, difference between groups of population or comparison between cities, countries, and so on

Composite indicators can provide useful approaches. A composite indicator synthesizes a number of values expressed by the indicators that constitute it (Booysen

2002; Nardo et al. 2005a; Sharpe and Salzman 2004) and re-establish the unity of the concept described in the hierarchical design. The aggregating process allows a somewhat 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.

Functions of Composite Indicators

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 across time, differences between geographical areas, and connections between social processes to be pointed out.
- *Explicative*, when the indicators are aimed at interpreting reality.
- *Predictive*, when the indicators help to delineate plausible evolutionary trends that it is possible to describe in terms of development or decrement; these indicators require strong prediction models and continuous observations across 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 reference standards defined in terms of time, territory, etc.; the reference values allow the evaluation of 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, and severities of specific problems (for example the lack of quality of life conditions among immigrants).
- *Evaluating*, which can be distinguished as:
  - *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. From 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 sensitivity.
- *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.

Perspectives of Observation

The indicators can be distinguished according to 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 and Sen 1997).

Anand and Sen (1997) argued that the conglomerative and deprivational perspectives are not substitutes for each other, and 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 a 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 and 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 different realities cannot 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 across 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. From this perspective, indicators can be classified as:
– *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 difficulty 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.

Methodological Issues

The construction of composite indicators requires a particular methodology and specific techniques aimed at:
1. Verifying the dimensionality of selected indicators (*dimensional analysis*)
2. Defining the importance of each indicator to be aggregated (*weighting criteria*)
3. Identifying the technique for aggregating the indicator 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*)

Selecting Indicators Leading to Dimensional Analysis

This analysis aims at selecting the indicators to be included in the composite, showing the best statistical characteristics.

From this perspective, *dimensional analysis* mainly allows the *dimensionality* of the conceptual construct, which the composite is based on, to be identified. In other words, dimensional analysis allows the analysts to investigate the level of complexity by which the composite indicator has to be constructed.

Actually, the results lead to a further selection of indicators before going through the construction of the composite indicator. From the statistical point of view, the selection should avoid superimposition and redundancies among indicators. However, in selecting the indicators also other criteria should be taken into account. In short, the criteria are:
- *Redundancy*. In building a composite indicator, two indicators showing a very high correlation are considered redundant; it is recommended to select only one of them.
- *Comparability*. When two indicators are redundant, it is recommended to select the one allowing trend analysis and wide comparisons.
- *Political impact*. If two indicators 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.

Dimensional analysis can be performed through different approaches (Alt 1990; Anderson 1958; Bolasco 1999; Cooley and Lohnes 1971; Corbetta 1992, 2003; Cox and Cox 1994; Hair et al. 1998; Kruskal and Wish 1978; Maggino 2004a, b, 2005a; Sadocchi 1981). Among them, the following methods are the more commonly used:
– *Correlation analysis*. It is useful in order to select indicators that are not redundant and to avoid multicollinearity (*double counting*) in composite indicator construction (Nardo et al. 2005a).
– *Principal component analysis*. The main goal of principal component analysis is to describe the variation of a data set using a number of scores that is smaller than the number of the original indicators. This approach is very often applied to test dimensional structures, even though this practice is strongly criticisable. 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.

– *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 indicators (Cox and Cox 1994; Kruskal and Wish 1978; Torgerson 1958).
– *Cluster analysis*. In this context, it can be useful to identify meaningful groupings among indicators (Aldenderfer and Blashfield 1984; Bailey 1994; Corter 1996; Hair et al. 1998; Lis and Sambin 1977; Maggino 2005a).

In some cases, methods related to a reflective model of measurement can be carefully used, like:

– *Performance analysis (Item Response Theory)*. When the indicators refer to performance variables, a particular analysis, derived directly from the application of *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 and Aitken 1981; Hambleton et al. 1991; Lord 1974, 1984; Ludlow and Haley 1995; McDonald 1989; Rasch 1960; Rupp et al. 2004; Rupp and Zumbo 2003, 2006; Sijtsma and Molenaar 2002; Swaminathan and Gifford 1982, 1985, 1986).
– *Factor analysis*. It allows the hypothesized dimensional structure underlying the group of indicators (latent structure analysis) to be tested; it is based upon the assumption that the total variance of each indicator is produced by a linear combination of different variance components (additive assumption), *common variance* (due to the dimensional structure), *specific variance* (due to the specificity variance of each indicator), and *error*. Actually, factor analysis allows the common variance (*communality*) to be estimated (Kim and Mueller 1989a; b; Marradi 1981).

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

Weighting Criteria

Since not necessarily all the identified 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 indicator before proceeding to the indicators aggregation.

When an implicit weighting system cannot 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 indicator to the construction of the composite indicator. From 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 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 indicator.

In order to proceed with the definition of a differential weighting system, the analyst needs to take into account (Nardo et al. 2005a):

- The defined rationale and theoretical structure which the conceptual construct and, consequently, the composite indicator are based on
- The meaning and the contribution of each indicator to the aggregation
- The quality of data and the statistical adequacy of the 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*)[8] and *differential weighting* (*DW*). The composite indicator will be strongly influenced by

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

whichever choice is made concerning this. Cases' positions can sharply change by simply changing the weights assigned to each indicator.

The adoption of the *differential weighting* procedure does not necessarily correspond 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. *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 and Salzman 2004).
2. *Multiattribute models*:
   (a) *Multiattribute decision making* (in particular, *Analytic Hierarchy Processes*—AHP) (Yoon and Hwang 1995),
   (b) *MultiAttribute Compositional Model* (in particular, *Conjoint Analysis*, CA),[9]
3. *Subjective methods*. New perspectives have been introduced recently showing the possibility involves more individuals (experts or citizens) in the process of defining weighting systems for social indicators. These approaches are defined from 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.[10]

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

– Theoretical reflections that endow a meaning on each indicator or consider its impact on the synthesis
– Methodological concerns that help to identify the proper techniques, consistently with the theoretical structure

[9]Hair et al. (1998); Louviere (1988); Malhotra (1996). A particular example of Conjoint Analysis application to QOL measurement see Maggino (2005b).

[10]Hagerty and Land (2007); Maggino (2008a, b, 2009); Maggino and Ruviglioni (2008a, b, 2009).

Table 10.14 Aggregating table according to a typical compensatory approach (additive technique)

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

In any case, we have to consider that a whole set of weights able to express in a perfect way the actual and relative contribution of each 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 analyze the evolution of the examined QOL ambit. The latter can be adopted when the aim—for example—concerns the definition of particular priorities. Please see Russell et al. (2006) for a discussion of whether weighting captures what is important in the phenomenon.

Techniques for Aggregating Indicators

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 composite 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 account the specific characteristics of each technique; in particular, we have to consider if the technique:

(a) Admits compensability among the indicators to be aggregated
(b) Necessitates comparability among indicators
(c) Necessitates a homogeneity in the levels of measurement of the indicators

(a) An aggregating technique is *compensatory* when it allows low values in some indicators to be compensated by high values in other indicators. In the typical aggregating table (see Table 10.14), we can observe all the possible synthetic values, obtainable by aggregating two indicators (*A* and *B*) using simple addition (additive technique).

Table 10.15 Aggregating table according to a typical compensatory approach (multiplicative technique)

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

Some of the obtained synthetic values, even if completely identical, are obtained through different original indicators. This means that obtained aggregated values do not allow us to return to the original unit's profile since the same synthetic values are obtained through different combinations of scores. In other words, two units with different realities turn out to be identical and indistinct.

By using the same data reported in the 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)—see table 10.15.

Table 10.15 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. From this perspective, a compensatory technique can be useful in some contexts especially when the purpose of applying indicators is to stimulate behaviors aimed at improving the overall performance by investing in those areas showing lower values.

All this highlights how important the choice of the aggregating technique is in order to avoid inconsistencies 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 noncompensatory aggregating procedure has to be preferred, such as a noncompensatory multicriteria approach, like multi-criteria analysis (MCA) (Nardo et al. 2005a).

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

– *Directionality* concerns the direction by which each indicator measures the concept (i.e., positive or negative). In some cases, it could be necessary to make the directionality of the whole group of indicators uniform before starting the aggregation process. In order to make the directionalities uniform, the indicators to be transformed should be submitted to the reflection procedure:

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

– *Functional forms* represent the changes in a variable that are valued at different levels. If changes are valued in the same way, regardless of level, then the functional form is linear. If changes are valued differently, according to the level, the functional form is not linear. In other words, in some cases the same absolute differences between observed values are valued differently and consequently can have different meaning (e.g., 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 interpreting the level of a variable, two issues arise:

– Are absolute values of a variable proportional in importance with reference to the measured concept?
– Are changes in the values of a variable of equal importance at various levels of the variable?

According to the response to these questions, the functional forms will be linear or nonlinear (Sharpe and Salzman 2004). Consequently, the most convenient interpretation and analytic treatment can be identified. 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 *n*th 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).[11] Both the functional forms are nonlinear by definition.

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

– Applying the appropriate functional form helps to better interpret the changes in the indicator. Many indicators commonly taken into account in social and economic indices show nonlinear functional forms, 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 (Sharpe and Salzman 2004).[12]

(c) *Homogeneity* refers to the level of measurement adopted by the whole group of indicators. Almost all the aggregating techniques require homogeneous scales. Some techniques exist allowing the 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 technique) is the most widely used. By contrast, multiplicative techniques (following the geometrical approach) and the technique based upon multicriteria analysis (following the noncompensatory approach) allow the difficulties caused by compensation among the indicators to be overcome (Table 10.16):

Assessing Robustness

(A) Uncertainty and sensitivity analysis

As we have seen, in order to proceed in aggregating multiple measures many choices have to be taken; these decisions can influence the robustness, which is the capacity of the composite indicator to produce correct and stable measures (Edward and Newman 1982; Nardo et al. 2005a; Saisana et al. 2005; Saltelli et al. 2004; Sharpe and 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 with reference to: (a) the model to estimate the measurement error; (b) the procedure for selecting the indicators; (c) the procedure for data management (missing data imputation, data standardization and normalization, etc.); (d) the criterion for weight assignment; and (e) the aggregation technique used.

In order to evaluate the robustness of the composite indicator, a specific analytic 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 construction process of the composite indicator. 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, and 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 analytic methodology (Nardo et al. 2005a):

1. *Uncertainty analysis*. This method aims at analyzing 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*. This method aims at evaluating 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 and Sandbu 2001). The iterative and synergic application of both the procedures have been revealed to be

[12] 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 and Salzman 2004).

Table 10.16 An overview of aggregation approaches

			Aggregating approaches			
			1. Linear aggregation		2. Geometrical aggregation	3. Noncompensatory aggregation
			Simple additive	Cumulative		
Assumptions	Dimensionality	Relationships between indicators	Uni	Uni	Uni	Multi
	Model of measurement	Relationship between indicators and latent variable	Monotonic	Differential relationship	Monotonic	
	Compensation	Among indicators	Admitted	Not admitted (scalability of indicators)	Admitted	Not admitted
	Homogeneity	Of the level of measurement	Requested	Requested	Requested	Not requested

useful and powerful (Saisana et al. 2005; Saltelli et al. 2004; Tarantola et al. 2000) in developing aggregated measures.[13]

(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 a 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-points* or *cut-offs*, referring respectively to continuous and discrete data. The selection of these reference scores is particularly useful when the composite indicator is applied for diagnostic and screening purposes.[14]

Criticisms of Composite Indicators

Despite its spreading, the composite indicator approach is currently being deeply criticized as inappropriate and often inconsistent (Freudenberg 2003). Critics point out conceptual, methodological, and technical issues especially concerning the difficulty in conveying into unidimensional measures all the relevant information pertaining to phenomena which are complex, dynamic, multidimensional, full of ambiguities, and nuances, and which are represented by data being sensitive and qualitative (even when quantitatively measured) and containing errors and approximations.

In other words, a composite indicator is hardly able to reflect the complexity of a socioeconomic phenomenon and to capture the complexity of the variables' relationships. This incapacity is related to the

[13] 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.

[14] *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 area 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**).

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.

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., & Fox, W. C. (1954). *The theory of signal detectability*. Institute of Radio Engineers Transactions, PGIT-4, 171–212.).

comprehensiveness and complexity of the phenomenon that should be covered by the composite indicators.

Those who maintain composite indicators stress they are simple to build and to communicate and based on "objective" computation tools. Although objectivity is always invoked as an essential requirement, in practice the procedures for computing composite indicators are far from being "aseptic." Generally, they comprise different stages (Nardo et al. 2005a, b; Sharpe and Salzman 2004), each introducing some degree of arbitrariness to make decisions concerning:

– The *analytic approach* to determine the underlying dimensionality of the available elementary indicators and the selection of those to be used in the evaluation process (*dimensional analysis*)
– The choice of the *weights* used to define the importance of each elementary indicator to be aggregated (*weighting criteria*);
– The aggregation technique adopted to synthesize the elementary indicators into composite indicators. (*aggregating-over-indicators techniques*).
– *The choice of the models and conceptual approaches* in order to assess:
  (a) The robustness of the synthetic indicator in terms of capacity to produce correct and stable measures (*uncertainty analysis, sensitivity analysis*)
  (b) The discriminant capacity of the synthetic indicator (*ascertainment of selectivity* and *identification of cut-point or cut-off values*)

Even though some decisions are strictly technical, it is quite difficult to make these decisions objective since they may involve different kinds of concerns. Generally, they are taken through a process accepted and shared by the scientific community.

Indicators selection. Selecting the indicators to be included in the composite represents a fundamental stage of the construction process since it operationally defines the latent concept that the composite is supposed to measure (formative logic). From the statistical point of view, this stage aims at:

– Exploring the level of complexity of the concept (dimensionality) as it is measured by the identified indicators
– Selecting the indicators showing the best statistical characteristics

The two goals are pursued contextually through traditional analytic approaches. Beyond the criticisms previously expressed concerning the metrics of data, the application of the traditional dimensional procedures puts other doubts, especially from the statistical logic point of view.

– *Factor Analysis*. It can be applied only to test the hypothesized dimensionality and to select the indicators that best fit the dimensional structure. In particular, it allows the hypothesized dimensional structure underlying the group of indicators (latent structure analysis) to be tested; it is based upon the assumption that the total variance of each indicator is produced by a linear combination of different variance components (additive assumption), *common variance* (due to the dimensional structure), *specific variance* (due to the specificity variance of each indicator), and *error*. Actually, factor analysis allows the common variance (*communality*) to be estimated (Kim and Mueller 1989a, b).
– *Principal Component Analysis*. The main goal of principal component analysis is to describe the variation of a data set using a number of scores that is smaller than the number of the original indicators. This approach is very often applied to test dimensional structures by assimilating it to factor analysis, even though this practice is strongly criticizable. In fact, 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.

Irrespective of the statistical tool adopted, dimensionality reduction raises some relevant questions, concerning its consequences on the composite indicator construction. If the concept to be measured turns out to be actually unidimensional, computing a single composite indicator could be justifiable. But when concepts are truly multidimensional, then singling out just one, albeit composite, indicator is very questionable. The nuances and ambiguities of the data would in fact be forced into a conceptual model where all the features affecting the multidimensionality are considered as noise to be removed. Moreover, synthetic scores could be biased towards a small subset of elementary indicators, failing to give a faithful representation of the data. Please see Zumbo and Rupp (2004) and Zumbo (2007) for a synthesis of the field and recommendations from a psychometric point of view.

Weighting indicators. When constructing indicators, particular attention is paid to the weighting process, which aims at assigning different *importance* to the elementary indicators to be aggregated. The necessity of choosing weights based on objective principles is frequently asserted (Nardo et al. 2005a, b; Ray 2008; Sharpe and Salzman 2004), leading to a preference for statistical tools like correlation analysis, principal component analysis, or data envelopment analysis, to mention a few. However, adopting purely statistical methods in weighting components of social indices must be carefully considered. Removing any control over the weighting procedure from the analyst, it gives a possibly false appearance of objectivity that is actually difficult to achieve in social measurement (Sharpe and Salzman 2004). Moreover, since defining weights is often interpreted from the perspective of identifying personal and social *values*, the procedure should necessarily involve individuals' contributions in attributing importance to different domains. Sometimes, the choice and decision could be shared by a larger community (involving individuals in the process of social indicators construction). If indicators concern societal well-being, their construction turns out to be not just a technical problem, being part of a larger debate aimed at obtaining a larger *legitimacy*. From this perspective, the weighting issue can be even considered as a leverage of democratic participation in decisions ("res publica"). Hagerty and Land (2007) stressed how building composite indicators should take into account and maximize the agreement among citizens concerning the importance of each elementary indicator. Choosing consistent weighting criteria is thus a subtitle issue, largely subjective and possibly data independent.

Aggregating indicators. Further criticisms concern the aggregating process, which raises methodological difficulties (Munda and Nardo 2008) encountered to get unidimensional scores out of multidimensional data, and which raises methodological difficulties when dealing with ordinal data. The process is in fact quite controversial since:

– The indicators to be aggregated are rarely homogeneous in many respects (metrics, directionality, functional form, …) and need not share common antecedents (Howell et al. 2007).
– The aggregating technique might introduce implicitly meaningless compensations and trade-offs among indicators.
– It is not clear how to combine ordinal variables, using numerical weights.

Methodological difficulties rise particularly when ordinal indicators are to be aggregated into a composite indicator, to get unidimensional scores for comparing and ranking statistical units. Unidimensional scores are usually computed through weighted averages of the ordinal evaluation variables, as in the quantitative case. As a matter of fact, this leads to highly controversial results, since weighted averages cannot be consistently computed over ordinal variables and different choice of the scaling tools would imply very different final conclusions (moreover, scaling tools tend to impose a quantitative latent model to data, which is often forced, arguable, and not fully justifiable on any epistemological basis).

Composite indicators represent the mainstream approach to socioeconomic evaluation (Maggino and Fattore 2011; Fattore et al. 2011), yet the discussion above shows how many critical issues affect their computation. The difficulties are even greater when ordinal variables are dealt with, since statistical tools based on linear metric structures cannot, strictly speaking, be applied to nonnumeric data. In a sense, socioeconomic analysis faces an impasse: (1) implicitly or not, it is generally taken for granted that ''evaluation implies aggregation''; thus (2) ordinal data must be scaled to numerical values, to be aggregated and processed in an (formally) effective way; Unfortunately, (3) this often proves inconsistent with the nature of the phenomena and produces results that may be largely arbitrary, barely meaningful, and interpretable. Realizing the weakness of the outcomes based on composite indicator computations, statistical research has focused on developing alternative and more sophisticated analytic procedures, but almost always assuming the existence of a cardinal latent structure behind ordinal data. The resulting models are often very complicated and still affected by the epistemological and technical issues discussed above. The way out of this impasse can instead be found realizing that evaluation need not imply aggregation and that it can be performed in purely ordinal terms. This is exactly what poset theory allows.

Assessing robustness. This stage aims at proving that the results obtained through the composite are not affected by the choices made along the process. The assessment is accomplished by applying uncertainty and sensitivity analysis. It could be interpreted as an attempt to objectify the choices, turned which were inevitably subjective. Actually, this stage aims at defending the choices through evidence. However, this approach does not require a methodological defense of the choices, in terms of scientific responsibility.

Traditional statistical data analysis procedures, based upon linear mathematical instruments, are hardly applicable for data discrete in their nature. New challenges and perspectives are emerging aimed at improving technical tools and strategies with reference to:

- Reducing data structure in order to aggregate units and indicators.
- Combining indicators.
- Communicating the "picture" obtained through the indicators (correctly and significantly representing and showing results).

These new challenges and perspectives should take into account:

- Nature of data (generally ordinal)
- Process and trends of phenomena (not always linear but more frequently monotonic)

By considering all this, new challenges and perspectives can be identified in order to improve the technical strategies allowing social indicators to be constructed and managed.

Modeling 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.

From this perspective, a proper analytic approach should be identified according to the defined conceptual framework. The feasibility of 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 modeling* (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 an exploratory approach. It usually starts with a hypothesis, represented as a model, operationalizes 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 with empirical data. 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 adjustment 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.

Given its specific assumptions, this approach can be adopted only in the presence of a strong 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 relationships among 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 a model allowing bidirectional effects to be estimated, has to be used on with extreme caution (Scherpenzeel and Saris 1996) and requires longitudinal data and analyses. The caution should increase especially in the presence of both objective and subjective indicators.

Multilevel Approach

Multilevel analysis refers to statistical methodologies, first developed in the social sciences, which analyze outcomes simultaneously in relation to determinants measured at different levels (for example, individual, workplace, neighborhood, nation, or geographical region existing within or across geopolitical boundaries) (Goldstein 1999; Hox 1995; Krieger 2002).

This approach can be applied from 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 contribute together with the micro-level living conditions (objective quality of life) to 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 analytic framework could be multiple regression; 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 a 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 in 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 modeled by a positive within-territory correlation among individual levels of subjective well-being in the same territory. This problem can be avoided by applying a variance components 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* (Raudenbush and Bryk 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 nonlinear models)[15] 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 analyzed at multiple hierarchical levels, whereas in simple linear and multiple linear regression all effects are modeled to occur at a single level.

For example, in educational research, where data are 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 analyze these kinds 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:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \tag{10.1}$$

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

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \tag{10.2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \tag{10.3}$$

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

Substitution of (10.2) and (10.3) in (10.1) gives:

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{01} Z_j + \gamma_{11} Z_j X_{ij} + u_{1j} X_{ij} + u_{0j} + e_{ij} \quad (10.4)$$

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 variables X at the lowest level, indicated by the subscript p ($p=1,\ldots, P$), and Q explanatory variables Z at the highest level, indicated by the subscript q ($q=1, \ldots, Q$). Then, Eq. 10.4 becomes the more general equation:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + u_{pj} X_{pij} + u_{0j} + e_{ij} \quad (10.5)$$

Multilevel analysis generally uses maximum likelihood (ML) estimators, with standard 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 and 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 the same living conditions at a 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 an individual's life, having a lifelong significance. The interest could be oriented to analyzing which of these processes are reversible and what the role of objective micro or macro level characteristics is.

This perspective deserves particular attention and consideration. Its limit is mainly represented by the difficulty of obtaining detailed and consistent individual longitudinal data and by the complexity of managing, analyzing, and modeling these kinds of data. According to its characteristics, this approach turns out to be useful for studying clinical data.

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.[16]

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.

[16]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 the 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 the Bayesian perspective, probabilities are described in terms of beliefs and degrees of uncertainty.

4. It offers an efficient and principled approach aimed at data overfitting.
5. Since a Bayesian 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*. Bayesian networks are easily extended to computing utility, given the degree of knowledge we have on a situation, and so 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, there is the selection of the statistical distribution induced in modeling the data. Selecting the proper distribution model to describe the data has a notable effect on the quality of the resulting network.

Traditional exploratory approaches, such as clustering and mapping approaches, multidimensional analysis, correspondences analysis (Aldenderfer and Blashfield 1984; Bailey 1994; Corter 1996; Hair et al. 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.

Closing Remarks

Methodological Challenges in Indicators Construction for the Measurement of Societal Well-Being

Actually, even a quick check of the academic literature allows us to see 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 in the hard economic perspective that considers 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)
   (b) *Subjective indicators* should not be seen as antithetical to objective indicators but as an important tool allowing information to be added, which cannot be provided by objective measures. In both perspectives, 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 a 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 levels 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 for exchanging information and dialog on these issues between different actors and within different research contexts.

Institutional Challenges: National Statistical Offices and the Measurement of Societal Well-Being

As we have seen, measuring and monitoring well-being of societies requires a complex and comprehensive framework and integrated approaches at conceptual and methodological levels. This perspective is urged not only by researchers belonging to academics but also by other organizations and institutions.

Also, the awareness aroused by many people directs us toward a more comprehensive approach in measuring societal well-being. The Report of Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al. 2009)—chaired by Joseph E. Stiglitz—represents further evidence of that and proposes the following 12 recommendations (Table 10.17):

As a consequence, measuring and monitoring societal well-being creates a great need for statistics but statistics with new and shared working models.

Moreover, this will require huge investments in order to carry out needed 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 organizations, and academic experts to work

Table 10.17 Twelve recommendations from the report of the commission on the measurement of economic performance and social progress

1	When evaluating material well-being, look at income and consumption rather than production
2	Emphasize 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 nonmarket 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)

alongside representatives of their communities to produce high-quality, facts-based information that can be used by all segments of society to form a shared view of societal well-being and its evolution over time.

A possible model could be that aimed at involving different public corporations operating in statistical areas and interacting in order to define an organic system, operating as a coordinated network organization (*statistical offices network*). Such a network's activities should be structured in nodes and needs to be:

– Aimed at defining clear statistical goals and programs
– Organized at different levels (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 statistical functions, by overcoming fragmentation, diversity and superimpositions at different network levels
– Adjusting forms of communication and involvement for different actors

These actions could be conceived at a:

– *General level,* since they should define norms concerning the statistical functions to be considered as a public service providing common and multifunctional wealth. Statistics should be considered in terms of knowledge and assessment
– *Specific level,* since they should promote (1) increasing the production of data and indicators at local levels; (2) interacting and integrating different data bases and data sources; and (3) developing appropriate analytic methods.

Some risks arise, 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."

References

Abbey, A., & Andrews, F. M. (1985). Modeling the psychological determinants of life quality. *Social Indicators Research, 16*, 1–34.

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: Sage.

Alt, M. (1990). *Exploring hyperspace. A non-mathematical explanation of multivariate analysis*. New York/London: McGraw-Hill.

Anand, S., & Sen, A. (1997). *Concepts of human development and poverty: A multidimensional perspective* (Human development papers 1997). New York: UNDP.

Andersen, E. B. (1972). The numerical solution of a set of conditional estimation equations. *Journal of the Royal Statistical Society, Series B, 34*, 42–54.

Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika, 38*, 123–140.

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York/London/Sidney: Wiley.

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: Sage.

Asher, H. B. (1983). *Causal modelling* (Sage university paper series on quantitative applications in the social sciences, series no. 07–003). Newbury Park: Sage.

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.

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: Sage.

Bartholomew, D. J., & Knott, M. (1999). *Latent variable models and factor analysis, Kendall's library of statistics* (Vol. 7). London: Arnold Publishers.

Bottarelli, E., & Parodi, S. (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*, XXIII: 49–68 (www.unipr.it/arpa/facvet/annali/2003/49.pdf)

Berger-Schmitt R., & Noll, H. -H. (2000). *Conceptual framework and structure of a European system of social indicators* (EuReporting working paper no. 9). Mannheim: Centre for Survey Research and Methodology (ZUMA) – Social Indicators Department

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

Blalock, H. M. (1964). *Causal inferences in nonexperimental research*. Chapel Hill: University of North Carolina Press.

Blalock, H. M. (Ed.). (1974). *Measurement in the social sciences. Theories and strategies*. Chicago: Aldine Publishing Company.

Bock, R. D., & Aitken, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrika, 46*, 443–459.

Bohrnstedt, G. W., & Knoke, D. (1994). *Statistics: A tool for social data analysis*. Itasca: Peacock Publishers.

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

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 10*, 305–314.

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

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: SAGE.

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

Cannavò, L. (2009). Dell'incertezza e della complessità: gli indicatori tra ricerca valutazione. In C. Bezzi, L. Cannavò, & M. Palumbo (Eds.), *Gli indicatori sociali e il loro uso valutativo*. Milano: Franco Angeli Editore.

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

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

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

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

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: 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* (pp. 65–102). New York: Academic.

Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., et al. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics, 61*, 267–276.

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

Cummins, R. A. (1993). *Comprehensive quality of life scale for adults (ComQol-A4)*. Melbourne: Deakin University, School of Psychology.

Cummins, R. A. (1995) On the trail of the gold standard for subjective well-being. *Social Indicators Research*, *35*: 170–200.

De Vellis, R. (1991). *Scale devolopment. Theory and applications* (Applied social research methods series, Vol. 26). London: SAGE.

Diamantopoulos, A., & Siguaw, J. A. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management, 17*, 263–282.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, 269–277.

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

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

Drenowski, J., & Wolf, S. (1966). *The level of living index*. Geneva: United Nations Research Institute for Social Development.

Drewnowski, J. (1970). *A planning model for social development*, Unrisd, (Report n.70.5), Geneva.

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* (pp. 89–125). New York: Academic.

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: Sage.

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

Engel, U., & Reinecke, J. (1996). *Analysis of change*. Berlin/New York: Walter de Gruyter.

Epidemiological Bullettin. (2002). Introduction to social epidemiology. *Epidemiological Bullettin, 23*(1). http://www.paho.org/English/sha/be_v23n1-socialepi.htm

Erikson, R. (1993). Descriptions of inequality: The Swedish approach to welfare research. In M. Nussbaum & A. Sen (Eds.), *The quality of life* (pp. 67–83). Oxford: Oxford University Press.

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

Fattore, M., & Maggino, F. (2011). Socio-economic evaluation with ordinal variables: Integrating counting and poset approaches in *Statistica & Applicazioni* (forthcoming).

Fattore, M., Maggino, F., & Colombo, E. (2011). From composite indicators to partial orders: Evaluating socio-economic phenomena through ordinal data. In F. Maggino & G. Nuvolati (Eds.), Quality of Life: reflections, studies and researches in Italy. *Social Indicators Research Series* (forthcoming).

Firebaugh, G. (1997). *Analyzing repeated surveys* (Sage university paper series on quantitative applications in the social sciences, series no. 07–115). Newbury Park: Sage.

Freudenberg, M. (2003). *Composite indicators of country performance: a critical assessment* (STI working paper, 2003/16, Industry Issues). Paris: OECD. url: http://www.olis.oecd.org/olis/2003doc.nsf/43bb6130e5e86e5fc12569fa005d004c/8bb0f462911c2cc6c1256ddc00436279/$FILE/JT00153477.DOC

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

Ghiselli, E. E. (1964). *Theory of psychological measurement*. New York/London: McGraw-Hill.

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: 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, H. (1999). *Multilevel statistical models*. London: Arnold Publisher. http://www.arnoldpublishers.com/support/goldstein.htm.

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

Greenley, J. R., Greenberg, J. S., & Brown, R. (1997). Measuring quality of life: A new and practical survey instrument. *Social Work, 42*, 244–254.

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

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, 35*, 455–496.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River: Prentice-Hall.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Measurement methods for the social sciences series, Vol. 2). London: SAGE.

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

Headey, B., Veenhoven, R., & Wearing, A. (1991). Top-down versus bottom-up theories of subjective well-being. *Social Indicators Research, 24*, 81–100.

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

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods, 12*, 205–218.

Hox, J. J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties.

Jamison, D., & Sandbu, M. (2001). WHO ranking of health system performance. *Science, 293*, 1595–1596.

Kalmijn, W. M., & Veenhoven, R. (2005). Measuring inequality of happiness in nations: In search for proper statistics. *Journal of Happiness Studies, 6*, 357–396.

Kim, J.-O., & 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: Sage.

Kim, J.-O., & 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: Sage.

Kozma, A., Stone, S., Stones, M. J., Hannah, T. E., & McNeil, K. (1990). Long- and short-term affective states in happiness: Model, paradigm and experimental evidence. *Social Indicators Research, 22*, 119–138.

Krieger, N. (2002). A glossary for social epidemiology. *Epidemiological Bullettin, 23*(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: Sage.

Lance, C. E., Mallard, A. G., & Michalos, A. C. (1995). Tests of the causal directions of globallife facet satisfaction relationships. *Social Indicators Research, 34*, 69–92.

Land, K. C. (1971). On the definition of social indicators. *The American Sociologist, 6*, 322–325.

Land, K. C. (1975). Social indicator models: An overview. In K. C. Land & S. Spilerman (Eds.), *Social indicator models* (pp. 5–36). New York: Russell Sage.

Land, K. C. (2000). Social indicators. In Borgatta, E. F., & Montgomery, R. V (Eds.), *Encyclopedia of sociology*, rev edn. New York: Macmillan

Lazarsfeld, P. F. (1958). Evidence and inference in social research. *Daedalus, 87*, 120–121.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton M Company.

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

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data* (Wiley Series in Probability and Mathematical Statistics). New York: Wiley.

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: 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: Sage.

Lord, F. M. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika, 39*, 29–51.

Lord, F. M. (1984). Standard errors of measurement at different ability levels. *Journal of Educational Measurement, 21*, 239–243.

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

Ludlow, L. H., & Haley, S. M. (1995). Rasch model logits: Interpretation, Use and trasformation. *Educational and Psychological Measurement, 55*, 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: Firenze University Press, Archivio E-Prints.

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

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

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

Maggino, F. (2005b). *The importance of quality-of-life dimensions in Citizens' preferences: An experimental application of conjoint analysis*. Firenze: Firenze University Press, Archivio E-Prints.

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

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. & Ruviglioni, E. (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) at the VII International Conference on Social Science Methodology (1–5 September 2008, Campus di Monte Sant'Angelo, Naples).

Maggino, F. & Ruviglioni, E. (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).

Maggino, F. & Ruviglioni, E. (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).

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

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

Marradi, A. (1981). Factor analysis as an aid in the formation and refinement of empirically useful concepts. In E. F. Borgatta & D. J. Jackson (Eds.), *Factor analysis and measurement in sociological research*. London: SAGE.

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

Max-Neef, M. (1992). Development and human needs. In P. Ekins & M. Max-Neef (Eds.), *Real life economics: Understanding wealth creation* (pp. 97–213). London: Routledge.

McDonald, R. P. (1989). Future directions for item response theory. *International Journal of Educational Research, 13*, 205–220.

Menard, S. (1991). *Longitudinal research* (Sage university paper series on quantitative applications in the social sciences, series no. 07–076). Newbury Park: Sage.

Michalos, A. C. (1985). Multiple discrepancies theory (MDT). *Social Indicators Research, 16*, 347–413.

Michalos, A. C. (1992). Use and abuses of social indicators. *Sinet* 32.

Michalos, A. C. (2008). Education, happiness and wellbeing. *Social Indicators Research, 87*, 347–366.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005a). *Handbook on constructing composite indicators: Methodology and userguide*. OECD, Statistics Working Paper

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005b). *Tools for composite indicators building*. European Commission, EUR 21682 EN, Institute for the Protection and Security of the Citizen, JRC, Ispra.

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

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., & Glover, J. (1995). *Women, culture, and development: A study of human capabilities*. Oxford: Oxford University Press.

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

Patel, S., Hiraga, M, Wang, L., Drew, D., & Lynd, D. (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, 44*, 5–39.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park: Sage.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in survey* (Wiley Series in Probability and Mathematical Statistics). New York: Wiley.

Rupp, A. A., & Zumbo, B. D. (2003). Which model is best? robustness properties to justify model choice among unidimensional IRT models under item parameter drift. *Alberta Journal of Educational Research, 49*, 264–276.

Rupp, A. A., & Zumbo, B. D. (2006). Understanding parameter invariance in unidimensional IRT models. *Educational and Psychological Measurement, 66*, 63–84.

Rupp, A. A., Dey, D. K., & Zumbo, B. D. (2004). To bayes or not to bayes, from whether to when: Applications of Bayesian

methodology to modeling. *Structural Equation Modeling, 11*, 424–451.
Russell, L. B., Hubley, A. M., Palepu, A., & Zumbo, B. D. (2006). Does weighting capture what's important? Revisiting subjective importance weighting with a quality of life measure. *Social Indicators Research, 75*, 141–167.
Sadocchi, S. (1981). *Manuale di analisi statistica multivariata per le scienze sociali*. Milano: Franco Angeli Editore.
Saisana, M., Saltelli, A., & Tarantola, A. (2005). Uncertainty and sensitivity techniques as tools for the analysis and validation of composite indicators. *Journal of the Royal Statistical Society A, 168*, 1–17.
Sakitt, B. (1973). Indices of discriminability, *Nature, 241*: 133–134.
Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice. A guide to assessing scientific models*. Chichester: Wiley, 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*. Amsterdam: Sociometric Research Foundation.
Scherpenzeel, A., & Saris, W. (1996). Causal direction in a model of life satisfaction: The top-down/bottomup controversy. *Social Indicators Research, 38*, 161–180.
Sen, A. (1993). Capability and well-being. In M. Nussbaum & A. Sen (Eds.), *The quality of life* (pp. 30–53). Oxford: Clarendon.
Sharpe, A., & Salzman, J. (2004). *Methodological choices encountered in the construction of composite indices of economic and social well-being*. Ottawa: Center for the Study of Living Standards.
Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory* (Measurement methods for the social sciences series, Vol. 5). London: SAG.
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*, 401–420.
Simpson, A. J., & Fitter, M. J. (1973) What is the best index of detectability? *Psychological Bulletin, 80*: 481–488.
Sirgy, M. J., Michalos, A. C., Ferriss, A. L., Easterlin, R. A., Patrick, D., & Pavot, W. (2006). The quality-of-life (QOL) research movement: Past, present, and future. *Social Indicators Research, 76*, 343–466.
Stiglitz J. E., Sen, A., & Fitoussi, J.-P. (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
Stoker, T. M. (1993). Empirical approaches to the problem of aggregation over individuals. *Journal of Economic Literature, 31*, 1827–1874.
Stones, M. J., Hadjistavropoulos, T., Tuuko, H., & Kozma, A. (1995). Happiness has trait-like and state-like properties: A reply to Veenhoven. *Social Indicators Research, 36*, 129–144.
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: Sage.
Swaminathan, H., & Gifford, J. A. (1982). Bayesian estimation in the Rasch model. *Journal of Educational Statistics, 7*, 175–192.
Swaminathan, H., & Gifford, J. A. (1985). Bayesian estimation in the two-parameter logistic model. *Psychometrika, 50*, 349–364.
Swaminathan, H., & Gifford, J. A. (1986). Bayesian estimation in the three-parameter logistic model. *Psychometrika, 51*, 589–601.
Tarantola, S., Jesinghaus, J. M., & Puolamaa, M. (2000). Global sensitivity analysis: A quality assurance tool in environmental policy modelling. In A. Saltelli, K. Chan, & M. Scott (Eds.), *Sensitivity analysis* (pp. 385–397). New York: Wiley.
Torgerson, W. S. (1958). *Theory and methods of scaling*. New York/London/Sydney: Wiley.
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*. New York: Oxford University Press.
Veenhoven, R. (2002). Why social policy needs subjective indicators. *Social Indicators Research, 58*, 33–45.
Veenhoven, R. (2005). Inequality of happiness in nations. *Journal of Happiness Studies, 6*, 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, 91*, 5–21.
Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1974). Quantifying unmeasured variables. In H. M. Blalock (Ed.), *Measurement in the social sciences: Theories and strategies*. Chicago: Aldine Publishing Company.
Yoon, K. P., & Hwang, C.-L. (1995). *Multiple attribute decision making* (Sage university paper series on quantitative applications in the social sciences, series no. 07–104). Thousand Oaks: Sage.
Zapf, W. (1975). Systems of social indicators: Current approaches and problems. *International Social Science Journal, 27*, 479–498.
Zapf, W. (1984). Individuelle Wohlfahrt: Lebensbedingungen und Wahrgenommene Lebensqualität. In W. Glatzer e & W. Zapf (Eds.), *Lebensqualität in der Bundesrepublik* (pp. 13–26). New York/Campus: Frankfurt a. M.
Zumbo, B. D. (2007). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Psychometrics, Vol. 26, pp. 45–79). Boston: Elsevier.
Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions and applications* (pp. 65–82). Charlotte: Information Age Publishing.
Zumbo, B. D., & Rupp, A. A. (2004). Responsible modeling of measurement data for appropriate inferences: Important advances in reliability and validity theory. In D. Kaplan (Ed.), *The SAGE handbook of quantitative methodology for the social sciences* (pp. 73–92). Thousand Oaks: Sage.