Methodological challenges in measuring the wellbeing of countries: the case of subjective wellbeing

Пловдивски университет „Паисий Хилендарски“

Факултет по икономически и социални науки

ИЗСЛЕДОВАТЕЛСКИ МЕТОДИ И ТЕХНОЛОГИИ В ИКОНОМИЧЕСКИТЕ И СОЦИАЛНИТЕ НАУКИ

RESEARCH METHODS AND TECHNIQUES IN ECONOMIC AND SOCIAL SCIENCES

Международна научна конференция
6 – 7 октомври 2012 г.

гр. Пловдив

УНИВЕРСИТЕТСКО ИЗДАТЕЛСТВО
„ПАИСИЙ ХИЛЕНДАРСКИ“

ISSN 978-954-423-837-7

СЪДЪРЖАНИЕ

СЕКЦИЯ 5
Интелигентни методи за обработка на бизнес информация

SESSION 5
Intelligent methods for processing of business information

Международната
ветена на важни преди
въпроси като:

1. Какво представ
2. Възможно ли е
изучена на фо
случайността?
3. Какви са възмо
предложат аде
изследователск
4. Докъде се прос
ние и контекст
5. Какви са възмо
подходи на поз

Всичко това отп
културата.

Светът е сложен
бъде изследван. Точно
ледване, защото ние в
който формира собст
модели на поведение.

Възможно ли е,
вателски методи, да п
мичният интерес е са
процесите на разработ
тяхната практическа е
широки слоеве от насе

Не е достатъчно
то се. Те трябва да ра
дели в икономическа
Всичко това налага во
и икономическите нау
ходи и методи, търсещ
поддържането на конк
преминава през етап
емпиричното ниво на

Dear Colleagues, Dear Guests,

The International scientific conference organized by the Faculty of Economic and Social sciences is dedicated to important challenges of our time that are expressed in questions such as:

1. What is the surrounding world?
2. Could the complexity of the social world be studied and understood when uncertainty, unpredictability and risk exist?
3. What are the possibilities of social and economic sciences to provide adequate answers by developing new research methods to conduct applied researches?
4. How far do the bounderies of their instrumental use and contexts of interpretation lie?
5. What are the possible ways of constructing new methodologies and approaches to knowledge?

All these create challenges to science, education and culture.

The world is complicated and unstable, but that does not mean that it could and should not be studied. That creates the need for new approaches of research to be developed, because we are largely dependent on our own way of thinking, which forms our opinions, attitudes, ideas, expectations and behaviors.

Is it possible, by applying the full range of traditional research methods to get an objective picture of the social world? The academic interest is only one aspect of the problem. Other aspects are the processes of development and adoption of management decisions and the evaluation of their practical effectiveness. These are decisions that affect the interests of a great number of the population.

It is not enough the applied researches just to fix what is happening. They must develop and provide mathematical and software models in the economic, social, cultural and educational fields. All these require discussions on current issues of social and economic sciences to be held. In particular, this refers to traditional approaches and methods for reliable results. The creation and the maintenance of a constructive dialogue between science and social activity go through a stage where long lasting relationships between the theoretical and the empirical level of knowledge are created. The role of methodology is crucial here.

Is it possible by studying the everyday life to get to ideas, concepts and models that allow us to leave the Present and look at the Future? And most importantly, this desire should be transformed by a series of researches into a

cultural phenomenon that will put its imprint on the hearts and minds of students and become part of their nature.

As organizers of this conference our goal is the conference to become an intellectual space for sharing and promoting experiences for effective application of research methods in education, science and practice. The format and the organization of this scientific event are oriented towards creating a scientific atmosphere that favours discussion, criticism and exchange of ideas between supporters of different theories and approaches.

It is clear by the presented abstracts of the papers that the conference has attracted the attention of representatives of different generations – from senior researchers to students.

I would like to thank all the authors and participants.

Thank you for your attention and good luck.

Prof. Ivan Ivanov

Dean of the Faculty of Economic and Social Sciences

Темата на конфе
икономическите и соц
участие с доклади на н
икономическите и соц
лия, Израел, Полша, С
ции, които обхванаха г
в социалните и иконом
ниците и гостите на к
ни въпроса.

Първият – предс
дени изследвания в раз
ледващите дискусии –
ни характеристики на
живот на хората и по
и свободното време;
човека; (в) формиране
ация на постоянно на
само да преценят как
да селектират поток
притежава някаква це
много измерения е ин
Главното изменение е
ношения. Промени се
ното, защото се появ
да станат „видими" н

Можем да обоб
социалната стратиф
изискват да бъдат от
те и икономическите
успеем да отговорим
провокативен въпрос
тази цел?

Процесите на у
нарастване на риска,
т.н., не са от днес ил
даваме, защото днеш

Often comparisons based on single questions or on composite scores of the latent variables are made. The problem of such comparisons is that one can compare the results across different countries only if in fact the data are comparable, that is, if the measures used in the different countries have the same meaning. This topic is studied under the heading of *functional equivalence* or *invariance of measures* in different countries. Invariance refers to (i) *configural invariance* (the same standard factor analysis model should hold for all different groups); (ii) *metric invariance* (equality of the loadings); and *scalar invariance* (the intercepts should also be equal in the different groups). These requirements are too strict. There are two reasons for this.

a) A response model can be specified that makes a distinction between (a) the *interpretation of the questions* (cognitive part of the model): which should be the same across groups; and (b) the *response process* (measurement part of the model), which produces less fundamental differences.

b) Significant differences across countries test are done for parameters of single indicators while these indicators are combined to an index. Therefore, a significant deviation of one indicator across countries in a set of other indicators may have only a very minimal effect on the total score for the index and the deviation may be rather irrelevant evaluating the index as a whole.

METHODOLOGICAL CHALLENGES IN MEASURING THE WELLBEING OF COUNTRIES: THE CASE OF SUBJECTIVE WELLBEING

Filomena Maggino

Professor of Statistics applied to Social Research
University of Florence, Italy

1. Let's make it clear

1.1. What is the meaning of "subjective" / "objective"?

Before showing the measuring process step-by-step, it is helpful trying to clarify here the meaning of "**subjective**" adjective, as well as its opposite "**objective**", consistently with different concepts.

1.1.1. Modelling reality: a subjective activity

In this case, we are referring to the conceptual framework defined in order to observe and to interpret the reality. The conceptual framework is always yielded by a *subjective* hypothesis and view of the world made by the researcher. Concerning this, as Michalos (1992) noticed, the models defined to observe a reality are only apparently neutral. Actually, the conceptual model represents only a "small window" through which only some facets of the reality can be seen (*reductionism*); in this sense, the view is politically and socially distorted and can condition knowledge, evaluations, choices, actions, and policies. In this sense, subjectivity expresses the unavoidable working hypothesis helping in understanding the reality. The researcher, through the dialogue with the working hypothesis, can change the perspective in a continuously evolving knowledge path.[1]

1.1.2. Measuring reality: an objective process

The methodologies adopted to study the characteristics defined in the ambit of the conceptual framework have to be **objective**. In other words, the methodological objectivity concerns the capacity of a procedure to measure without alteration due to external factors and to be free from effects due to the

[1] In any scientific field, a meaning does not exist without "subjectivity". This dimension is present also in the ambit of hard sciences, where it helps in structuring the observed reality. In this perspective, data represent the "text" to be read by the researcher in search of a sense. This process cannot be considered arbitrary, since the sense cannot be "invented" regardless of the relationship with the reality.

observer; this notion spreads from the procedure of measurement to the data analysis to the interpretation of the results.

The measuring process should meet the requirements of any scientific, strictly connected to the possibility to take into account the *error* which represents a hypothetical component of any procedure of measurement. The observational error (with its additive *random* and *systematic* components) is the amount by which an observation differs from its expected value (Carmines & Zeller, 1992):

Accuracy and precision

In any scientific application, accuracy and precision are closely related; in particular

– the *accuracy* represents the capacity of the procedure to measure what we intend to measure; assessing accuracy requires the observation of a known process or the availability of a *reference value* (calibration); the correspondent concept of "accuracy" for subjective measurement is validity.

– the precision is the degree to which further measurements will show the same or similar results; by determining the precision of a measurement it is possible to verify, consequently, the capacity to measure through a degree of distortion as low as possible; the precision is related to the concepts of robustness and stability and can be distinguished into (i) repeatability (variation arising when all efforts are made to keep conditions of measurement constant and by repeating during a short time period), and (ii) reproducibility, that is the variation arising by using the same measurement process among different instruments and operators, and over longer time periods.

Precision is assessed by controlling the coherence of the model of measurement; the correspondent concept of "precision" for subjective measurement is reliability.

A procedure of measurement that meets these requirements not only gains scientific relevance but can also be standardized. We can assume that the greater the number of repetitions, the more accurate and precise the estimation of the error. By assuming repeated measures, the error is said **uncertainty**.[2] Moreover, reliability does not imply validity, while validity implies reliability. That means a valid measure must be reliable, but a reliable measure need not to be valid. That is

[2] In case of repeated measures, the value corresponding to the highest frequency is assumed to have the highest probability to be close to the true value; besides, positive errors are assumed to compensate negative errors (even if this is not actually always true). Consequently, the distribution of the values of the repeated measures is assumed to be *normal*.

a reliable measure is mea
is supposed to be measur

1.1.3. Components

In this case the ad
defined in the ambit of
measured and analyzed.
subjective characteristics
to the source – called *unit*
units can be represente
administrative areas, geo
distinguish between (i) *o*
(ii) *subjective informatio*

As known, definin
the measuring process.
definitions (i) *objective i*
observers, and (ii) *subje*
implicit criteria which ca

1.1.4. Objective/su

Sometimes, the di
equivalent to the distinct
not correct. In our perspe

- "objective – su
- "quantitative –

applied in order to observ

2. Let's define

2.1. What are the

In synthetic terms,
each individual lives (he
on). They can find differe

- micro-level, re
- macro-level,

demographic, geographic

Table 1 summarizes

a reliable measure is measuring something consistently, but not necessarily what it is supposed to be measuring.

1.1.3. Components of the reality: both objective and subjective

In this case the adjectives refer to the kind of information which has been defined in the ambit of a conceptual framework and subsequently objectively measured and analyzed. In order to make the distinction between **objective and subjective** characteristics more clear from the operative point of view, we can refer to the source – called *unit* – on which the characteristic of interest is measured. The units can be represented by individuals, institutions, social groups, services, administrative areas, geographical areas, nations, and so on. Consequently, we can distinguish between (i) *objective information,* collected by observing reality, and (ii) *subjective information*, collected only from individuals and their assertions.

As known, defining what we are going to measure represents the first step of the measuring process. This distinction can be developed into other resulting definitions (i) *objective indicator*, based upon explicit criteria, shared by external observers, and (ii) *subjective indicator*, based upon subjective evaluations and implicit criteria which can vary from individual to individual.

1.1.4. Objective/subjective vs. quantitative/qualitative

Sometimes, the distinction between objective and subjective is considered equivalent to the distinction between quantitative and qualitative. Of course, this is not correct. In our perspective, we can summarize the two dyads as follows:

- "objective – subjective" refers to what we are going to observe
- "quantitative – qualitative" refers to the methodological approach applied in order to observe the previous dimensions

2. Let's define

2.1. What are the "objective" and "subjective" characteristics?

In synthetic terms, **objective components** refer to the conditions in which each individual lives (health, working conditions, environmental situations, and so on). They can find different definition according to two major perspectives:

- micro-level, referring and taking into account the individual level
- macro-level, concerning and taking into account economic, demographic, geographical, administrative or social level.

Table 1 summarizes the different components of the objective characteristics.

	level				
Objective	micro	Demographic and socio-economic characteristics		– sex – age – civil/marital status – household	– educational qualification – occupation – birthplace / residence / domicile
		Observable acquired knowledge		– skills – cognition	– know-how – competences
		Individual living conditions and resources		– standards of living – financial resources (income) – housing	– working and professional conditions and status – state of health
		Social capital		– social relationships – freedom to choose one's lifestyle – geographical mobility	
		Observable behaviours and life style		– activities (work, hobby, vacation, volunteering, sport, shopping, etc.)	
				– engagements (familiar, working, social, etc.)	
				– habits (schedule, using of public transport and of means of communication, diet, etc.)	
				– public life (participation, voting, etc)	
	macro	Structure of societies	Social conditions	Social exclusion	Disparities, equalities/inequalities, opportunities
				Social inclusion	Informal networks, associations and organisations and role of societal institutions
				Social mobility	
			Political setting		Human rights, democracy, freedom of information, etc.
			Institutional setting		Educational system
					Health system
					Energy system
			Economical setting		Income distribution, etc.
			Decisional and institutional processes		
		Environmental conditions			

Table 1. *Objective characteristics observed at micro and macro levels*

Traditionally, **subjective characteristics** can be distinguished in three content areas (Nunnally, 1978): abilities, personality traits, and sentiments (Tab. 2).

The scheme is not exhaus
overlap one another.

2.2. What are the su

The widely shared m

– *quality of life*, ref
– *equity*, referring
community (macro) level
– *sustainability*, re
levels, the environment and

subjective	abilities / capacities	intellectual	
		special	
	personality traits		
	sentiments	Interests an	
		Values	
		Attitudes	

Table 2. *Subjec*

Each concept shows
evaluations, feelings, percep
to each individual life as a
objective characteristics, nc
can be defined in observing

The notion of *qual*
nents/dimensions: *living co*
subsequent sub-dimensions
to questions about happines

The concept of *equity*
population and refers to d
inclusion/exclusion, social c

The scheme is not exhaustive and the identified components for each area can overlap one another.

2.2. What are the subjective components in measuring wellbeing?

The widely shared main concepts defining the progress of a community are

- *quality of life*, referring to wellbeing observed at individual (micro) level
- *equity*, referring to the fair distribution of wellbeing and observed at community (macro) level
- *sustainability*, referring to the relationship between the two previous levels, the environment and the future

	components			
subjective	abilities / capacities	intellectual	– verbal comprehension and fluency – numerical facility – reasoning (deductive and inductive) – ability to seeing relationships	– memory (rote, visual, meaningful, etc.) – special orientation – perceptual speed
		special	– mechanical skills – artistic pursuits	– physical adroitness
	personality traits		– social traits – motives – personal conceptions	– adjustment – personality dynamics
	sentiments	Interests and preference		
		Values		
		Attitudes	cognitive → evaluations (beliefs, evaluations opinions)	
			affective → perceptions (emotional states)	
			behavioural intentions	

Table 2. *Subjective characteristics observed at micro level*

Each concept shows subjective aspects, which refer to and concern opinions, evaluations, feelings, perceptions, attitudes, desires, values, and motivations related to each individual life as a whole or in different specific contexts. Contrarily to the objective characteristics, no explicit standard is defined and no external reference can be defined in observing the subjective component.

The notion of *quality of life* can be structured in two macro components/dimensions: *living conditions* and *subjective wellbeing*, which in turn find subsequent sub-dimensions. It can be assessed by individuals' or groups' responses to questions about happiness, life satisfaction, utility, or benefit.

The concept of *equity* can be expressed as the distribution of wellbeing in the population and refers to dimensions like economic and social cohesion, social inclusion/exclusion, social capital, connection and social ties.

The concept of *sustainability* refers to the possible erosion / durability of those conditions (interpretable in terms of capitals) with reference to present generations' future and future generations. According to the World Bank's four capital approach, for example, sustainable development should enhance and preserve social, human, produced and natural capital of present generations and provide future generations with them.

The comparison that each individual can make with other persons, groups, neighbours, colleagues, and so on allow the perceived equity to be observed while retrospective and prospective comparisons allow perceived sustainability to be described.

The different levels of importance that each individual assigns to living aspects in (i) any person's life and (ii) one's own life, reveal respectively a component of individual's values and the individual investment in the different life's areas. The interpretation of the latter case allows individual's *internal equity*[3] and the sustainability of individual investment to be checked. Equity and sustainability of individual living conditions can be evaluated at micro level by observing the "perceived adequacy", "perception of the future", and sense/scope of one's own life. Social capital can be explored also by describing the interpersonal and systemic trust expressed by each individual

This description allows us to realize how the subjective perspective constitutes an integral part of those concepts' definition. Consequently, the measurement of the subjective perspective represents an important component of the national progress and well-being's description and assessment.

2.3. How can subjective wellbeing be interpreted and explained?

The interpretative and explanatory models allow determinants of subjective wellbeing to be identified. This is important especially if the obtained information allow subsequent policy intervention to be discussed and planned. The different ways through which the relationship between living conditions and subjective well-being is defined relate to different conceptualization of quality of life. *Subjective wellbeing* is subjected to continuous reflections in different fields (psychology, sociology, and economics). The different approaches (Sirgy, 2011) aim at giving a definition and subsequent explanation to subjective well-being, by referring to (a) personality traits, (b) evaluations and values, or (c) feelings and emotions, or (d) perceptions and functionings. Generally, psychologists include in subjective well-being also abilities/capacities while sociologists and economists consider them just conditions (Fischer, 2009). Recent orientations prefer mixed approaches.

3. Let's start measuring

3.1. How subjective characteristics can be observed?

The previously defined conceptual framework allows characteristics to be defined "subjective" and consequently measured. In this perspective, different methodological approached can be identified, in order to observe phenomena

[3] *Internal equity* is distinguished from *external equity*, observed by comparing different social groups (identified through different perspectives, age, gender, and so on).

consistently. In order to
(Tab. 3) can be identifie
producing quantitative or

Measuring approaches	Dimensio that can be measur
Performance	abilities
Inventory / self inventory	personali values interests
Self-reported measures	attitudes opinions abilities
Observational methods	behavio
Projective techniques	social t motivat adjustm attitude
Physiological measures	

Table 3. A

consistently. In order to measure subjective characteristics, different approaches (Tab. 3) can be identified and combined in various practical and functional ways, producing quantitative or qualitative information.

Measuring approaches	Dimensions that can be measured	Measures	Notes
Performance	abilities	– execution of an assigned task; – outcome evaluated with reference to specified criteria of success	intra-individual and inter-individual comparisons to be evaluated
Inventory / self inventory	personality values interests	individual answer to a certain number of submitted "stimuli"	
Self-reported measures	attitudes opinions abilities	subject's answer – expressed in terms of agreement, preference, etc. – to a particular statement referring to the characteristic to be measured	Problems: – instability of opinion and sentiments, especially those reflecting emotional feelings – incomparability of expressions among individuals (subjects can use different criteria, e.g., in evaluating their own life)
Observational methods	behaviour	observation made on the subject by an external and neutral observer	
Projective techniques	social traits motivations adjustment attitudes	individual's reaction to one situation constructed but not completed	– requires strong interpretative approaches. – interpretation of answers strictly connected to researcher's experiences (subjective method) – method considered not scientifically applicable and relevant (not standardizable) – useful to integrate other approaches
Physiological measures		relationship between subjective traits and physiological processes	scientific evidence of this connection is not completely demonstrable

Table 3. *Approaches to measure subjective characteristics*

The measurement of subjective characteristics presents some difficulties, produced by specific factors; in particular, the measurement of a certain characteristic can turn out to be falsified because of *individual factors* as social desirability, response set, researcher's attitude, *semantic factors* that can provoke interpretative discordance between subject and researcher, and *situational factors* when the observation occurs in different situations (presence or not of other persons) o in different context (at home, at work, in the street, etc.).

3.1.1. Quantitative and qualitative approaches

The approaches can be broadly classified into "quantitative" and "qualitative". The two perspectives differ in many ways (Tab. 4) and often these differences have been the source of considerable debates and divisions between different researchers.

3.2 What is the nature of subjective data?

3.2.1. Data theory

In order to convert empirical observation into understandable, interpretable and analysable information (datum), a theory is needed allowing nature of information to be clarified. The reference theory for subjective data definition is that defined by Coombs (1950, 1953, 1964; Flament, 1976; Jacoby, 1991; McIver & Carmines, 1979), based upon geometrical interpretation.

	Qualitative	Quantitative
Aim	Complete and detailed description.	Constructing statistical models in order to explain what is observed. Classifying and counting features.
To be adopted when	Recommended during earlier phases of research projects	Recommended during latter phases of research projects.
Hypothesis	Exploratory perspective and purpose (i.e., hypothesis-generating).	Confirmatory perspective: testing hypotheses.
Prior knowledge	Researcher knows roughly in advance what he/she is looking for.	Researcher knows clearly in advance what he/she is looking for.
Study design	The design emerges as the study unfolds.	All aspects of the study arc carefully designed before data collection.

Researcher role The basic underlying assumptions guide and sequence the types of data collection methods employed.		Rese – is strur – s bein situa – t imm Rese prob Hen shou rese
Measures quality		Qual estab
Data collection		Diffi obse obse jourr grou inter
Data	**form**	Wor
	characteristics	Mor less
	interpretation	Subj inter impo obse etc.
	analysis	Data – c – a ideas An e techr is tex conte
Sampling		Samp the s typic Case to ce conte

Table 4. *Key differen*

Researcher role The basic underlying assumptions guide and sequence the types of data collection methods employed.		Researcher – is the data-gathering instrument – studies by participating and/or being immersed in a research situation – tends to become subjectively immersed in the subject matter Researcher's 'neutral' position is problematic. Hence qualitative researchers should reflect on their role in the research and analytical process.	Researcher – uses tools, such as questionnaires or equipment to collect numerical data – is ideally an objective observer who neither participates in nor influences what is being studied – tends to remain objectively separated from the subject matter.
Measures quality		Qualitative perspective aims at establishing content validity.	Quantitative perspective aims at establishing reliable and precise measures through focused hypotheses, measurement tools and applied mathematics.
Data collection		Different approaches (participant observation, non-participant observation, field notes, reflexive journals, structured interview, group discussion unstructured interview, text analysis).	Models and hypotheses lead to development of instruments and methods for measurement.
Data	**form**	Words, pictures or objects.	Numbers, codes, statistics.
	characteristics	More 'rich', time consuming, and less able to be generalized.	More efficient, able to test hypotheses, but may miss contextual detail.
	interpretation	Subjective - individuals' interpretation of events is important, e.g., uses participant observation, in-depth interviews etc.	Objective – seeks precise measurement & analysis of target concepts, e.g., uses surveys, questionnaires etc.
	analysis	Data are – categorized into patterns – analyzed through observer ideas and impressions An example of quantitative techniques using qualitative data is textual data analysis, or content analysis.	Data analysis represents an important moment of the research process and is based upon statistics.
Sampling		Sample size is usually small and the sampling procedure is typically not probabilistic. Cases can be selected according to certain characteristics or contextual information.	Sample size is usually large and the sampling procedure is typically probabilistic.

Table 4. *Key differences between qualitative and quantitative research*

All empirical observations can be represented like comparisons, more or less explicitly performed, between at least two entities, which can be defined as points into a particular space. The relative positions of the two points depend on the interpretation given to the comparison between the two entities. For each observation, the datum can be defined in terms of geometrical relationship between two entities. The geometrical representation of data composes a model. Coombs (1953, 1964; Flament, 1976; Jacoby, 1991; McIver & Carmines, 1979) developed a theory based completely on the geometrical interpretation of data.

According to the theory, two entities in a single datum can vary with reference to two criteria:

a. the set to which the entities belong to (i) two *different sets* (e.g. a student and a test) or (ii) the *same set* (e.g. student *A* and student *B*).

b. the relation in which the entities are involved that can be (i) *dominance relation*: an individual answers a question by reporting a level exceeding a defined measure or (ii) *proximity relation*: two individuals share an event or two objects match or coincide with each other at different levels.

Combination between the two criteria (set and relation) produces four different kinds of data (Tab. 5): all empirical observations, regardless of their substantive nature, can be classified into one of the four types (Flament, 1976; Jacoby, 1991):

		Pairs of points in observation	
		same set	different set
Relation between points in pair	dominance	Stimulus comparison	Single stimulus
	proximity	Similarities	Preferential choice

Table 5. *Classification of data according to Coombs' theory*

4. *Let's construct subjective data*

The construction of subjective data represents a delicate stage of the measurement procedure and needs special care, in order to avoid excessive arbitrary elements and inaccuracy.

Quantitative measurement of subjective components needs to be founded on strong theoretical and methodological principles. These principles, grounded in psychometrics standards, states that the measurement of subjective characteristics requires a model in order to obtain interpretable and analysable information. The model is aimed at ensuring comparability within-individual and between-individuals. It allows *observation* (the collected information) to be transformed into *datum* (analysable information)[4] and is composed by:

[4] Data consist of portions of information extracted according to a reference model; in this sense, data represent a researcher's construction and interpretation. As well-known, "data" represents the plural form of the Latin neuter term "datum" which represents a form

1. a procedure aime
with reference to the obser
the observed characteristic
1980; Stevens, 1951, 1957;
2. a model for ass
assignment); this model all
data interpretable and th
measurement).

4.1. How the respon

In constructing the s
and, to some extent, to c
captured. Identifying the
definition is, however, stric
This procedure, rec
scaling, which can be def
Actually in general terms,
broken up into discrete
continuum is defined *unde*
scaling). However, since,
measurement, continuous
Any *scaling* procedure req
– the **criterion**, wh
(ii) affective (feelings, int
– the **reference**, w
compare two or more stin
previously defined schem

(participle) of the Latin verb
is to assign, to fix, to establi
[5] The wider used compa
comparison, rank order, con
[6] Generally, non-comparati
applied and can be classifie
– *kind of representati*
Koch-Weser, 1977), with
continuum with reference to
o the polarity (for th
o the identification o
o the meaning of the
o the orientation (e.g
also because of different cu
– *anchoring*, which co
– *number of segmen*
definitions.

1. a procedure aimed at defining the response score for each individual case with reference to the observed characteristic (scaling techniques) with reference to the observed characteristic (Andrews & Withey, 1976; Lodge, 1981; Marradi, 1980; Stevens, 1951, 1957; Weller & Romney, 1990),

2. a model for assigning data values (defining the *rules for numeric assignment)*; this model allows a value to be assigned that makes the constructed data interpretable and that may be treated in operative terms (system of measurement).

4.1. How the response can be captured? (defining the response score)

In constructing the subjective datum a problem is placed in order to define and, to some extent, to create the score allowing the subjects' response to be captured. Identifying the score corresponds to creating a continuum whose definition is, however, strictly theoretical.

This procedure, recalling what was defined at data theory level, is named ***scaling***, which can be defined according to different approaches (Marradi, 1980). Actually in general terms, every approach assumes a continuum, which can be (i) broken up into discrete ordinal categories (*discrete scaling*), in this case the continuum is defined *underlying continuum*, or (ii) metrically defined (*continuous scaling*). However, since, this identification would require, an infinitely precise measurement, continuous scaling is an abstraction that is difficult to be observed. Any *scaling* procedure requires the following aspects to be defined:

– the **criterion**, which can be (i) cognitive (evaluations and judgments) or (ii) affective (feelings, interests, preferences)

– the **reference**, which can be (i) comparative (the subject is asked to compare two or more stimuli)[5] or (ii) absolute (each respondent reacts, by using a previously defined scheme, to each stimulus independently of the others.[6]

(participle) of the Latin verb "dare" (do, das, dedi, datum, dare). The meaning of this verb is to assign, to fix, to establish, to frame, to place.

[5] The wider used comparative scaling techniques are comparative choice, pairwise comparison, rank order, comparative rating, constant sum.

[6] Generally, non-comparative scaling techniques are very simply to be constructed and applied and can be classified with reference to

– *kind of representation* (verbal, numeric or graphical) of the continuum (Aureli & Koch-Weser, 1977), with particular care to be paid to the numerically represented continuum with reference to

- o the polarity (for the lowest value up or from the highest value down)
- o the identification of the polarity (i.e. anchoring)
- o the meaning of the used graphic symbols
- o the orientation (e.g. horizontal or vertical)

also because of different cultural effects.

– *anchoring*, which could be explicit or defined by each respondent (self-anchoring)

– *number of segments* and their correct balance between positive and negative definitions.

Subsequently, in order to make data useable for statistical analysis, a value should be assigned to each segment. The set of values defines the system of measurement. Both criteria and references present applications in the scaling techniques applied in order to collect subjective wellbeing data [7]

4.2. How to assign an analysable value to the captured response?

After defining the model aimed at capturing the response (construction of the datum), a different model is to be defined aimed at assigning analysable values to observed data, consistently to the identified continuum. This model allows a value to be assigned that makes the constructed data interpretable and treated in operative terms. In other words, the model (*system of measurement*) is defined by

1. **rules**, allowing numbers/symbols to be assigned to each identified level in a standard and uniform way (*kind and criteria of measurement*) (Bruschi, 1999),
2. a **system of classification**, allowing status – with reference to the measured characteristic – to be assigned to each case (*type and level of measurement*) (Caracciolo in Siegel & Castellan, 1992; Stevens, 1946, 1951; Velleman & Wilkinson, 1993).

The whole group of values identified according to this model defines what is usually called a *scale*.

The definition of the "system of measurement" represents one of the most debated points concerning the subjective measurement. In fact, except for those rare cases in which the system is self-evident and does not require a detailed formulation, the definition of the "system of measurement" is never simple, unambiguous, clear, or intuitive, and it usually raises problems of arbitrariness.

5. Let's transform data into measures

The subsequent step is to transform the obtained subjective data into subjective indicators. Different subjective indicators are defined in order to cover the different concept and their dimensions with reference to different life domains.[8] This process leads to a [conceptual] matrix (Tab. 6), in which not each combination of conceptual dimension and ambit (*cell*) will be covered by indicators.

5.1. How can multi-indicator measures be transformed into complex indicators?

Investigating different aspects of subjective wellbeing (as well as other subjective dimensions) requires the definition of items (basic indicators) representing what is actually measured with reference to the corresponding subjective dimension.

[7] For examples, the *Satisfaction with life scale* (Diener et al., 1985) adopts the cognitive scaling criterion while the *Affect Balance Scale* (Bradburn 1969) applies the affective scaling criterion. Some approaches include both approaches (*Index of well-being*, Campbell et al., 1976).

[8] Life domains represent segments of the reality in which fundamental concepts should be observed and monitored.

According to a simple
single indicator (*single-indic*
just one item.

This strategy, undoubt
robust assumptions concernir
and one indicator. Such ass
produce a wide and consider
(i) *precision* (*reliability*), sin
strongly affected by random
one single indicator can desc
questionable; (iii) *relationsh*
discriminating and *differentia*

Actually, the adoption
able in case of complex conc
single-indicator approach to
measures allow the charact
multiple measures compens
particular, the items defined
measures, since each item c
dimension, whose variability

CONCEPTS ↓ **SUBJECTIV DIMENSION** ↓

Quality of life	Subjective evaluatio living conditions	
	Subjective wellbeing	Cog
		Affi
Equity	Inequalities	Per of e livi con
		...
	Social cohesion and social capital	Tru
		Ide
		Sol
		...
Sustainability	Expectations	in i ual con
		in sitt
		in tiv bei
		...

According to a simple and weak strategy, each dimension is defined by a single indicator (*single-indicator approach*), e.g., measuring "happiness" through just one item.

This strategy, undoubtedly thrifty and functional, requires the adoption of robust assumptions concerning the direct correspondence between one dimension and one indicator. Such assumption presents a risk since single indicator can produce a wide and considerable amount of error, leading to problems concerning (i) *precision* (*reliability*), since the measurement through one single indicator is strongly affected by random error; (ii) *accuracy* (*validity*), since the chance that one single indicator can describe one conceptual dimension is highly dubious and questionable; (iii) *relationship* with the other dimensions; and (iv) *capacity of discriminating* and *differentiating* among observed cases.

Actually, the adoption of several items (*multi-indicator approach*) is desirable in case of complex concepts. This approach allows the problems produced by single-indicator approach to be avoided or, at least, reduced. In fact, multiple measures allow the characteristic to be measured with more precision (since multiple measures compensate random) accuracy and discriminant capacity. In particular, the items defined in multi-indicator approach are considered *multiple measures*, since each item corresponds to one particular aspect of the conceptual dimension, whose variability can be covered.

CONCEPTS ↓	SUBJECTIVE DIMENSIONS ↓		LIFE DOMAINS (AMBITS) ↓ 1	2	3	4	5	6	7	8	9	10	11	12	13
Quality of life	Subjective evaluation of living conditions														
	Subjective wellbeing	Cognitive													
		Affective													
Equity	Inequalities	Perception of equity in living conditions													
		...													
	Social cohesion and social capital	Trust													
		Identity													
		Solidarity													
		...													
Sustainability	Expectations	in individual living conditions													
		in general situation													
		in subjective well-being													
		...													

1	households and families	6	education	9	health
2	housing	7	labour market and working condition	10	environment
3	transport			11	social security
4	leisure and culture	8	Income and standard of living	12	crime and safety
5	participation			13	total life situation

Table 6. *Measuring subjective characteristics of wellbeing: measures' conceptual organization*

5.1.1. The scaling models

Since, as we have seen, in order to define an accurate and precise measurement of each subjective characteristic, several items are needed, each cell of the previous conceptual matrix could include several basic indicators which need be synthesised. Aggregating subjective indicators requires the application of particular models allowing a group of items to be assessed and developing the corresponding scale. These models are named *scaling models*.

Each scaling model represents a theoretical design aimed at consistently developing a new measure (Nunnally, 1978) by verifying the main scale properties, including dimensionality, internal consistency and convergent/discriminant validity. In particular, each scaling model aims at testing the underlying conceptual model, re-establishing the unity of the concept of interest, synthesizing the multiple items (defining the synthetic value to be assigned to each case), and defining the continuum on which each case can be placed in a meaningful, interpretable and manageable way.

5.1.2. Selecting the appropriate scaling model

In order to select the scaling model that better fit the measurement approach, the following criteria (Tab. 7) should be taken into account (McIver & Carmines, 1979; Maggino, 2007; Netemeyer et al., 2003):

1) *Adopted scaling technique*, comparative or non-comparative.

2) *Nature of data* (single stimulus, stimulus comparison, similarities or preferential choice).

3) *Dimensionality*, which could be (a) uni-dimensional (the definition of the characteristic assumes a unique, fundamental underlying dimension), or (b) multidimensional (the definition of the characteristic assumes several underlying aspects/dimensions).

4) *Standard of measurement*, concerning the treatment of the multiple items and the assignment of the synthetic value (the final score can be assigned to individuals or to "stimulus").

For each conceptual model, the identified items should be submitted to item analysis in order to support and justify the aggregate them in a single interpretable score, according to the proper scaling models.

6. Methodological c

By concluding this w
could be pointed out, 1) m
for comparative research.

A further challenge
longitudinal research.

Concerning the first
a multi-level framework
various levels of a comple
aggregates.

In particular, he foc
aggregated and inferences
these cases inferences at a
nation) may carry with se
the inference (form of "fall

6. Methodological challenges in measuring subjective well-being

By concluding this work, two challenges in measuring subjective well-being could be pointed out, 1) measures' assessment, and 2) measurement requirements for comparative research.

A further challenge is represented by the measurement requirements for longitudinal research.

Concerning the first issue, Zumbo (2009, Zumbo & Forer, in press) proposes a multi-level framework for validity, using individual measurement results at various levels of a complex ecologically rich system, moving from individuals to aggregates.

In particular, he focuses on what happens, when the individual measures are aggregated and inferences are made at a different (higher) level in the system. In these cases inferences at a higher aggregate level (neighborhood, province, region, nation) may carry with secondary dimensions that may contaminate or confound the inference (form of "fallacies"). Zumbo and Forer (2011) note that:

Scaling models → selection table						
Criteria				**The choice**		
(1) How the response score has been captured?	(2) How data are interpreted?	(3) What is the nature of the measured characteristic?	(4) What will be the final score assignment?	**Scaling models**		**Criterion applied for testing the model (Item Analysis approach)**
scaling technique	**nature of data**	**dimensionality**	**standard of measure-ment**			
Non-comparative	Single-stimulus	Uni	Cases	**Uni-dimensional**	**Additive**	Internal consistency
Non-comparative	Single-stimulus	Multi	Cases	**Multidimensional**		Dimensionality of the items
Comparative (pair comparison or rank-order)	Stimulus comparison	Uni	Items	**Thurstone model (differential scale)**	**Cumulative**	Metrics between items
Comparative (rank-order or comparative rating)	Stimulus comparison	Uni	Items	**Q methodology**		
Non-comparative	Single-stimulus	Uni	Cases and items	**Guttman**	**Deterministic**	Scalogram analysis: reproducibility, scalability and predictability
		Bi	Cases and items	**Multidimensional Scalogram Analysis (MSA)**		Regionality and contiguity
		Bi	Cases and items	**Partial Ordered Scalogram Analysis (POSA)**		Correct representation
						parameters estimation (maxi-

		Bi	Cases and items	**Partial Ordered Scalogram Analysis (POSA)**			Correct representation
Non-comparative	Single-stimulus	Uni	Cases and items(without condensation)	**Monotone (one or more parameters)**	**Probabilistic**		parameters estimation (maximum likelihood) goodness of fit (misfit and residuals analysis)
Comparative (pair comparison)	Similarities	Multi	Items	**Multidimensional scaling**		**Perceptual Mapping**	Goodness of fit of distances to proximities (stress, alienation)
Comparative	Preferential choice	Uni & Multi	Cases and items	**Unfolding**			Goodness of fit of distances to ordinal preferences
Comparative (rank-order)	Preferential choice	Multi	Items at individual level	**Conjoint model**			Goodness of fit of the model (part-worth) to the ranking

Table 7. *Characteristics of the scaling models*

– any inferences from the individual level may not hold in the same way at higher (or lower) levels of aggregation

– systematic and coherent evidence (validation evidence) needs to assembled to support the inferences at the various levels

– the level of validation evidence needs to be in line with the level of inferences

– individual level validity evidence (which is what is traditionally done in validation research, e.g. criterion validity) does not provide sufficient validity evidence for inferences at higher levels in the system; and may actually be misleading because it may miss invalidity at the aggregate level.

– By summarizing, applying traditional individual differences validation methods (e.g., correlation with another wellbeing measure, or even cognitive response models) are insufficient evidence for support multi-level validation inferences like those often used in wellbeing research.

– In fact, these individual differences validation methods are susceptible to the cross-level inferential fallacies such as the (reverse) ecological fallacy or atomistic fallacy.

– Multi-level validation arises when one has a multi-level construct; that is, an individual level measure (or assessment) and aggregating it to make inferences at a higher level.

– Historically, multi-level constructs have not been an issue in measurement and validation because measurement has been immersed in and emerged from an individual differences psychological or sociological school of thought.

– In Zumbo's proposal, multi-level validation research might include, for example, the following issues:

– Is the aggregate score reflecting differences between measurement units (at that level) such as national differences in wellbeing?

– To what extent might we be measuring, unintentionally, other important constructs at the aggregate level that are not meant to be included in our measures (at that level), such as, construct irrelevant variance like neighborhood effects, or regional effects, or is our measure of wellbeing mostly a restatement of gross domestic product (or other such economic indicators)?

– Also a matter of determining what is that is (and is not) be measured by that aggregate variable.

– Concerning the second issue, in a recent work (2007), Saris & Gallhofer discuss problems of cross-cultural comparative research. It is well known that measurement error has strong effects on results of research. Therefore, when the effects of measurement error differ in the individual countries, comparisons across countries become quite challenging. Two types of comparisons are most frequently made, (i) comparison of means and (ii) comparison of relationships of different variables across countries.

Often comparisons based on single questions or on composite scores of the latent variables are made. The problem of such comparisons is that one can compare the results across different countries only if in fact the data are comparable, that is, if the measures used in the different countries have the same meaning. This topic is studied under the heading of *functional equivalence* or *invariance of measures* in different countries. Invariance refers to (i) *configural invariance* (the same standard factor analysis model should hold for all different groups); (ii) *metric invariance* (equality of the loadings); and *scalar invariance* (the intercepts should also be equal in the different groups). These requirements are too strict. There are two reasons for this.

a) A response model can be specified that makes a distinction between (a) the *interpretation of the questions* (cognitive part of the model): which should be the same across groups; and (b) the *response process* (measurement part of the model), which produces less fundamental differences.

b) Significant differences across countries test are done for parameters of single indicators while these indicators are combined to an index. Therefore, a significant deviation of one indicator across countries in a set of other indicators may have only a very minimal effect on the total score for the index and the deviation may be rather irrelevant evaluating the index as a whole.