Measurement awareness: the use of indicators between expectations and opportunities[1]

Con senso di misura:l'uso di indicatori tra aspettative ed opportunità

Lorenzo Bernardi
Dipartimento di Scienze Statistiche, Università di Padova
lorenzo.bernardi@unipd.it

Vincenza Capursi
Dipartimento di Metodi Quantitativi per le Scienze Umane, Università di Palermo
capursi@unipa.it

Leonardo Librizzi
Facoltà di Economia, Università di Palermo
leonardolibrizzi@inwind.it

"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science."
Lord Kelvin

In questo lavoro sono discussi problemi di merito e di metodo che si presentano allorquando ci si imbatte nell'articolato processo di costruzione di un indicatore composto. Alcuni temi discussi con riguardo agli indicatori sociali sono, quindi, richiamati, pur senza analizzare la varietà e la specificità degli apparati definitori e dei sistemi classificatori reperibili nella vastissima letteratura. Una particolare attenzione è rivolta all'analisi del rapporto fra indicatori e costruzione delle scale, che si conclude con l'illustrazione di un esempio, ancora in progress, che mette in risalto la distanza che tuttora separa volontà conoscitiva e fattibilità operativa.

Keywords: indicators, measurement, scaling

1. Introduction

The debate on the cultural-scientific approaches for the study of social phenomena has been for a long time rooted on the harsh confrontation between prophets on the discovery at all costs and hardworking followers of the praxis of explanation.
These are two strategic visions of research that seem to imply that an ideological separation as regards to epistemology, methods and techniques of work, modes of

[1] The present paper is financially supported by MURST funded Research (ex 60%, 1998) "Tecniche di costruzione di indicatori semplici e composti", directed by V. Capursi.

production of the results, the meanings themselves of the knowledge value achieved; this separation seems to have found a justification in a set of dichotomies apparently radical, which in concrete action tend still to run alongside, finding contextual, even though diversely dosed, the adoption[2] of quality versus quantity, the subjective versus the objective, uniqueness of the events (and of the conditions of their happening) versus repetition, deduction versus induction, the search for meaning versus empirical research.

The incipit acknowledged as the principal inspiration of this work explicitly assesses the choice of branch of learning even if it is assumed as a warning and not as the exclusive path. To give value to this declaration of belonging, as social scholars, there is also another simple but acute admonition[3]: the attention for the processes of measurement begin in the social sciences late, in the 17th century, while it has always been necessary in the government of civil society[4] and in the world of commerce and trade. In other words, the importance of measurement in every occasion of social exchange is highlighted no matter what the terrain of communication/confrontation between men is: the rules and the forms need to assure conditions of transparency, reproductions, modality of assessment.

From this point of reflection the theme of indicators is collocated with particular emphasis, as the functions for which they have in the course of time been evoked: informative, predictive and programme evaluative; also for the reason of the delicate terrains on which they should be carried out: health, education, social and economic policies , law, the environmental quality, the existence of man and his organisations in all their complexity. Above all Land's strong indication (1971) is still valid for which an indicator assumes meaning only when it possesses an informative value within a theoretical model however defined – mathematically, operationally, logically, in words – by analysis and interpretation 'of the functioning of the social sectors and segments'. Linking Land's suggestion to one of the possible functional classifications it seems evident that the construction, the role played, the meaning itself of the indicators, these are solved with different modalities and with different procedures according to the ratio intrinsic to the nature of the model[5].

To Land's precious statement, however, the history of the indicators, which has developed in these last decades, seems to have added a supplementary principle: the same single indicator is usually the result of a process of disarticulation of a complex event / social phenomenon that, seen in its elementary components, has to be successively reassembled through justified procedures of aggregation, weigthing, transformation. Clearly utilizing modalities of observation / description / analysis / interpretation that refer to the qualitative and the quantitative, to the subjective and objective, to external sources and to original surveys, to the research of connections (casual or not) and to reasonable inclusive visions.

---

[2] De Lillo A. (1980) analysing Durkheim and Weber's knowledge courses, that move from assumptions epistemologically opposite, he tries to demonstrate that, in the end, they utilise similar methodologies in different phases of the research development.

[3] Leonardi F. (1991)

[4] The affirmation by the Chinese Emperor of the 5th century BC is well known; he recommended the men of government to keep trace of at least 10 figures: the number of the employed, the old in need of assistance, the fat-stock, etc.

[5] Bernardi and Torelli (1992) identify, for example, 5 kinds of models, within which the indicators assume diverse functions; Buratto (1986) discusses the role of the indicators in a process/model of empirical assessment of a theory.

In conclusion, what we want to assess is that the indicator exists in a model and that the indicator itself is the product of a model, and following this line (or a certain process of its own construction) it sometimes furnishes content and meaning to the concept[6]. The aim in this contribution is that of discussing the problems regarding merits and methods – seen not as opposites but as factors that converge on the choice and definition of a useful interpretative and knowledge tool, both for the researcher and the decision-maker – which present themselves in the difficult process of defining a complex indicator. To this aim the work initially develops some arguments on social indicators, without going into details on the variety and specificity of the defining structures and the classifying systems available in the abundant literature, analysing in depth the methodological relationship between indicators and scaling, illustrating with an example, still in progress, the distance that still divides the want of knowledge from the operational feasibility.

2. Indicators: some references

The choice of using in the title of this work the word indicator without the attributive 'social' comes from the belief that most indicators can be labelled 'social'. The diffusion of the tool 'indicator' in every disciplinary field indicates the need to have 'statistics' that are able to describe the state of a system or to highlight changes, or simply to furnish signs of alarm to the public decision-maker. This could represent the standpoint of the user. From a scientific point of view if we follow the path offered by Horn (1993, p.15) the indicators are essential elements of scientific research, tools in a kit of mathematical techniques which can go beyond the inflexibility of closed models and provide interesting axioms with informative and illuminating contents in a simple and evident manner. Poincaré himself (1913) maintained that science conforms to a reasonable choice from a hierarchy of facts, if, in our opinion, description and theory move alongside. They may also not be so strict and exact as some methods of physical sciences, but they help to establish conjectural relationships. They can thus represent an analytic and interpretative tool of great importance if the techniques used to construct indicators, simple or composite, do not disobey the principles and axioms of the numerical proprieties they exhibit.
Curatolo (1972), in a broad report traced a complex overview of what could (and probably still can) be intended for Social Indicator passing through the analysis of the 'social dimension', of the responses of statisticians in various countries, of the meaning of Indicator and Social, of the many definitions found in the literature and, lastly, of a recognition of the voluntary omissions as well as the procedures to pursue for the growth of the indicators. It is our belief that not much has changed from the distant 1972 as regards to the 'formal acknowledgement of the discipline' of indicators. On the one hand we witness brilliant and philosophically acute disquisitions, and on the other an enormous dissemination of indicators in many disciplinary fields (social sciences, behavioural sciences, educational settings, environmental settings, scientometrics, health boards, town planning, etc.) and in many organisations, local authorities public administration etc. In the statistics area, however, something has changed in the attitude

[6] For example, it is sufficient to think about the I.Q. measurement tests, which are adopted as the measurement of intelligence itself.

towards the quantification on the part of statisticians as regards to what was reported by Curatolo (p.109): *'la posizione assunta a questo riguardo dagli statistici …….è decisamente limitatrice e critica di certe pseudo-quantificazioni operate dai cultori delle scienze dell'uomo'*. In fact the national and international literature sees many articles on technical issues and, in general, on the use of standard tools in the context of scales when the attention is focused on 'measuring' complex concepts as the 'level of living' (Panek, 1987), the 'quality of life' (Cox, Fitzpatrick e Fletcher, 1992; Fayers and Hand 2002) or the quality of teaching (Gori,1992, Petrucci and Rampichini, 2000) and other concepts also as complex. This only to quote published work in publications characteristically statistical, the overview, in fact, is much broader, but a review on the issue goes beyond the aims of this work. One can almost assert that there is a certain excitement in the construction and production of indicators that induce the reintroduction of many questions on the legitimacy of a set of choices taken both on the merit and the method. And some of the questions that seem to us compulsory are those that in Curatolo's work are referred to as not treated because they were not shared by the author. The strongest point seems to be that reported as a lack of 'convenience' in the construction of complex or 'aggregated' indicators. The author has widely delineated the reasons of the opportunity of constructing only simple or 'disaggregated' indicators. This is the reason why in his work the crucial problem of the passage from quality to quantity, apart from some very brief notes, is absent. However, even if we want to share, as best choice, the arranging of a series of simple indicators, we think that the problem quality-quantity cannot be ignored. The objectives that permit the identification of an indicator condition the choice in merit and in method, and one can construct an indicator for the aim of making simple ranking, evaluating, monitoring, comparing. But of what use is the comparison of single dimensions if the objective is to compare systems and not single elements? And, in any case, even the measurement of a single element can present the crucial problem of measurement: how to measure, for example, the element 'teacher's clarity during lessons' on the basis of a judgement expressed by a student?

Which indicators and to what purpose, then? An answer, linked to the difficulty of the concept which is to be measured through the indicator could be: let's try to find a way of 'measuring' that, taking into consideration the aims, turns out to be the most objective and fitting to describe in a quantitatively and qualitatively way the process we are interested in.

Which definition of indicator in this context? None, in the hypothesis that the objectives determine the indicators. We only want to affirm that indicators are the bond between statistical observations and the phenomena that one wants to measure; they are 'metadata', i.e. data that describe data, statistics with a purpose. Moreover the indicators help to bear light onto the concept that one wants to measure and finish up with assuming the dual role of the concept specification and sign (concept measurement) (Horn, 1993), as summarisable in the following table in which, even if with a certain forcing, we have tried to imagine a possible definition of a concept (a clever student in a subjective and objective version) moving from a plurality, likewise concurrent, of observable dimensions and from a discrete (but pertinent) list of predictable elementary aspects (simple indicators):

Table 1: *Possible connection between concept and measure*

Measure / Concept interpretation	Qualitative		Quantitative	
	Subjective	Objective	Subjective	Objective
Subjective	Self-assertion of cleverness; measurement modality (m. m.): Much, sufficient, poor, none.	Comparative evaluation of peers; m.m.: more, same, less.	Number of hours dedicated to study.	Rate of class attendance.
Objective	Teaching support to peers (recognized as such by these); m.m.: yes, no.	Knowledge of informative / teaching material; m.m.: all, partial, none.	Number of exams, evaluated complex, passed.	Number of credits gained.

It needs to be specified that the objectivity in the measure of a complex construct is more formal than substantial, thus in the table above it should be intended in the sense that the measurement is based on factual evidence opposed to a measurement based on an opinion – subjectivity.

3. Indicators and scaling

'Measurement is the hallmark of science. It not only enables us to classify and compare but also places the whole of mathematics at our disposal' (Bartholomew,1996). *'Statistics depends crucially on measurement yet statisticians have not been prominent among those actually constructing measures. This has largely been left to practitioners in, for example, psychology, economics and medicine.'* (Bartholomew, 2002) Bartholomew's suggestions bring us to the problem of scaling constructing for the measurement of a one-dimensional or multidimensional concept that juxtaposes itself to that of the construction of a simple or complex indicator. The debate on the meaning of the operation that is carried out when numbers are assigned to some elements, usually non-numerical, that define a propriety, in the attempt to transfer some characteristics of the numbers to the proprieties themselves (measurement theory) (Stevens, 1946), is long-running and can be in part found in Stevens (1974). Stevens's approach (1974, p.25) is *'that of invariance and to classify scales of measurement in terms of the group of transformations that leaves the scale form invariant'*. Measurement may be regarded as the constructing of scales (in the behavioural and social sciences, homomorphism in measurement theory) from empirical relational structures of interest into useful numerical relational structures on the basis on assumptions taken from foundational analysis (Krantz et al., 1971). *A relation structure is a set together with one or more relations on that set* (Krantz et al. 1971, p. 8). Thus, in general if we need to assign numbers to elements, often non-numerical, we need to find a real valued function, *T,* that is order preserving and additive[7]. If such a function exists we can affirm that *T* is a homomorphism of an empirical relation structure in a numerical relation structure. So, in general, if *T* is a homomorphism between two sets,

[7] See Krantz et al. (1971).

A and B, the identification of a group of transformations that leaves the scale form invariant is expressed in the identification of the class of admissible transformations of *T*, i.e. of functions of the type *f: T(A)→B* for every type of scale so that the set of the transformed measures preserve the proprieties of the original set. In this way, for a nominal scale, the admissible transformations *f* are of the one-to-one type; for an ordinal scale, *f* are strictly increasing on *T(A),* in the sense that only the order is preserved under these transformations; for an interval scale, *f* are $f(x)= ax + b$, for all *x* in *T(A),* with $a>0$ and $b\neq0$, which means that the ratios of interval between observations with a different origin *a* and a different scale parameter *b* are the same (i.e. ratios of intervals are invariant)[8]; for a log-interval scale, *f* are $f(x)= a\ x^b$ for all *x* in *T(A),* with $a,b >0$, a logarithmic transformation of such a scale results in an interval scale[9]; for a ratio scale, *f* are $f(x) =ax$, for all x in *T(A)* and with a>0, in the sense of equality of ratios, whichever the units of measure are of *A* and *B*[10].

If the scales must reflect the relation between concept (construct) and its measurement, the indicator too must reflect this relation.

In general, the construct can be represented by measurements on related variables which give rise to single indicators. When the construct is one-dimensional and directly measurable (height, weight, age, ...), we have the usual problems of respecting the mathematical-statistical proprieties for comparison (i.e. standardisation, or methods to adjust to the effect of some factors), or if the construct is not directly measurable and we use an indicator we must be careful to identify the specific relationship between construct and indicator and assess whether the transformation carried out on the original data is scale invariant (Leibowitz and Hyman, 1999). When the construct, ***X***, is multidimensional and both the general form of the relation between each variable component *(y,t,z…)* and the related indicator *(y',t',z'..)* and the function between ***X*** and *(y,t,z…)* are known, we have to focus attention on the transformations that are applied to single variables in order to obtain simple indicators and to what permits the aggregation process of these. In this case the problem is complex, but it can be resolved, provided that the mathematical proprieties are respected in the aggregation operation (Leibowitz and Hyman, 1999). When the construct is complex and of difficult definition there is a wide margin of arbitrariness in the identification both of the dimensions (or latent variables of these) and related indicators, and of a complex indicator that often measures a latent variable of the construct ***X***. Nevertheless, the effort to refine the method of measurement to avoid violating the statistical-mathematical proprieties cannot be ignored. This is because the complex indicator, that is already the result of large arbitrary choices, could result without meaning and, thus, useless. An example we can refer to is the composite indicator utilised by the Italian financial newspaper *Il Sole 24ore* in the annual report *Qualità della vita* (QV) (quality of life) on the 103 Italian Provinces. In this report they sum over 36 variables x_i, using two transformations $T(x_i)$:

$$T_1(x_i) = \frac{x_i}{Max\{x_i\}}1000; \qquad T_2(x_j) = \frac{Min\{x_j\}}{x_j}1000$$

[8] In interval scales two arbitrary choices are made: the zero point (usually conventional) and the unit of measure.

[9] This scale concerns the power transformations.

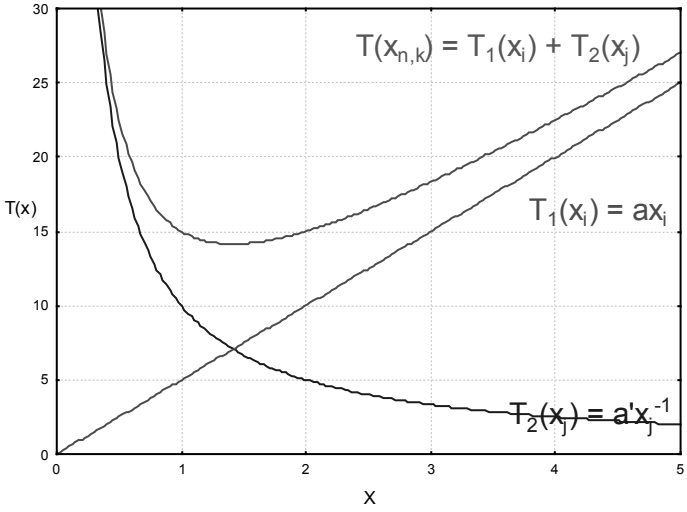
[10] In ratio scale an arbitrary choice is made: the unit of measure. The set $B\subseteq \{R\}$, in which $\{R\}$ is the set of real numbers

The first is of the type $T(x)= ax$, (a ratio scale), when x_i , which represents the reported value of each province for the single variable, identifies a dimension considered positive for the construct subject to study. The second is of the type $T(x)= a'x^{-1}$ (a weak log-interval scale), when x_i identifies a dimension considered negative for the construct. The function sum used for the composite indicator:

$$T(x_{nk}) = T_1(x_i) + T_2(x_j)$$

is the result of a linear combination of two functions, one linear and the other non linear. The question is: is it correct from a mathematical point of view? No. The obtained result can be synthesised graphically in the following way:

Figure 1: *Simple and composite indicators of QV*



One can clearly see that the resulting function is not monotone, that is it does not hold one of the most important assumptions. What is the meaning of $T(x_{nk})$? Which scale of measurement? Which proprieties of the original data have been preserved? (Attanasio, Capursi, 1997; Aiello, Attanasio,2004).

Even when there are no errors, as in the one mentioned, it is nonetheless difficult to identify the best approach for measuring a multidimensional and abstract construct, both from the point of view of the sense of measurement and from the point view of the field of application. The title of the article by Cox, D.R., Fitzpatrick, R., Flatcher, A.E. (1992) in this sense is instructive: *'Quality of life assessment: can we keep it simple?'* The paper is enlightening and extremely educational for statisticians and practitioners for the clarity with which many important issues are outlined and resolved (choice of instruments of measure, the desiderata of their proprieties, etc.). But in this case, modesty is necessary. Nanny Wermuth's comment – quoted in the paper – on the standard scales seems to be the most appropriate synthesis: '*Though desiderable properties of scales are readily listed, I think that we have to admit that statisticians have failed to provide widely accepted methods for obtaining scales with such properties ... I am convinced that no simple methods will be appropriate for constructing good scales and establishing their high quality, especially if it is likely that responses depend on cultural context, on knowledge, or on the education of a respondent.'*

4. Indicators between expectations and opportunities: an example

In the previous section some aspects of measurement have been highlighted, while other important aspects for the construction of complex indicators have been neglected. Below are the fundamental steps to follow in the construction of a composite indicator:

1. The definition of a construct we want to measure.
2. The specification of construct, which consists in the choice of the dimensions, components, items, and variables which compose the construct itself.
3. The identification and the evaluation of the instruments of measurement for the elements defined in point 2 above, especially when these are not directly measurable.
4. The choice of suitable transformations to apply to the elements in point 2 after 'measurement' in order to construct indicators of all dimensions.
5. Weighting: are all the dimensions equally important? If so, the problem does not exist. If they are not, weighting is an extremely important issue for the severity of the effects on the final result.
6. Lastly, the choice of the function in order to recompose the simple indicators in a single measure: the composite indicator.

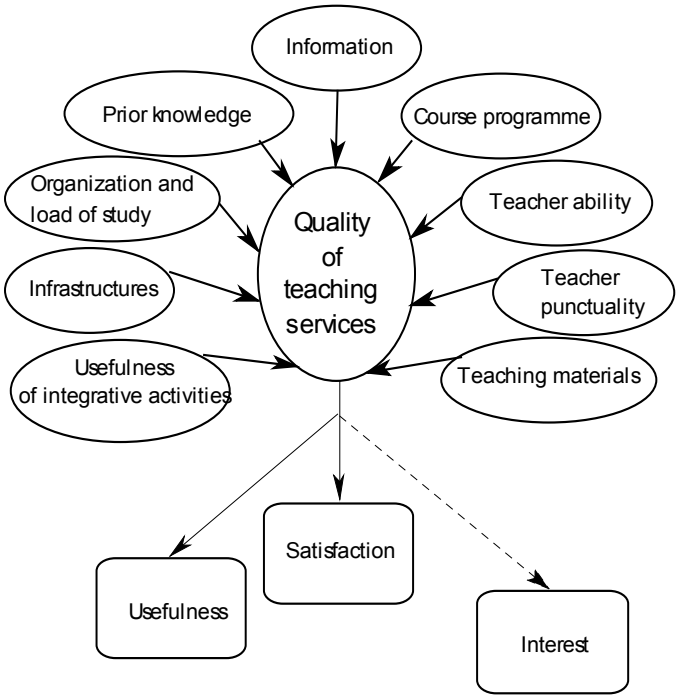
The six points above are difficult to resolve and, moreover, can be object of great arbitrariness when the construct we want to measure is complex and holds many elements of subjectivity. Moreover all the choices are heavily influenced by the proposed aim.

Let us analyse some of the six points with a case study (Capursi, Librizzi, 2004). The study aims to measure the construct 'quality of teaching at the University of Palermo'[11], comprising all the organisational aspects from facilities to curriculum programming of the whole course, based on student assessments of the teaching activities for each course. The construct is complex and based solely on elements of subjectivity. In fact what is measured is the quality of the teaching services (QT) oriented to students' needs and expectations, both conditioned by the heterogeneity of the 'evaluators' for different cultural backgrounds, for the different attitude in the choice of university and, in general, for the different social-economic conditions (point 1). The specification of the concept (point 2), through the identification of its components, was carried out with an analysis of the conceptual type, sustained *a posteriori*, even if not in a rigorous fashion, through the non metric multidimensional scaling method, which gave a configuration of the points, each of which represents an item, compatible with the components previously identified. The representation below (fig.2) has the construct to be measured at the centre, around it the components that specify it, and below the effects of the quality in the opinion of the students: satisfaction in the foreground, the usefulness of teaching in the background and marginally the interest in the topics treated.

The measure of proximity chosen to elucidate the relationship between the items of the questionnaire is Kendall's cograduation coefficient, τ_b, the goodness of fit of the adaptation was assessed graphically with Shepard's diagram on a two dimensional space.

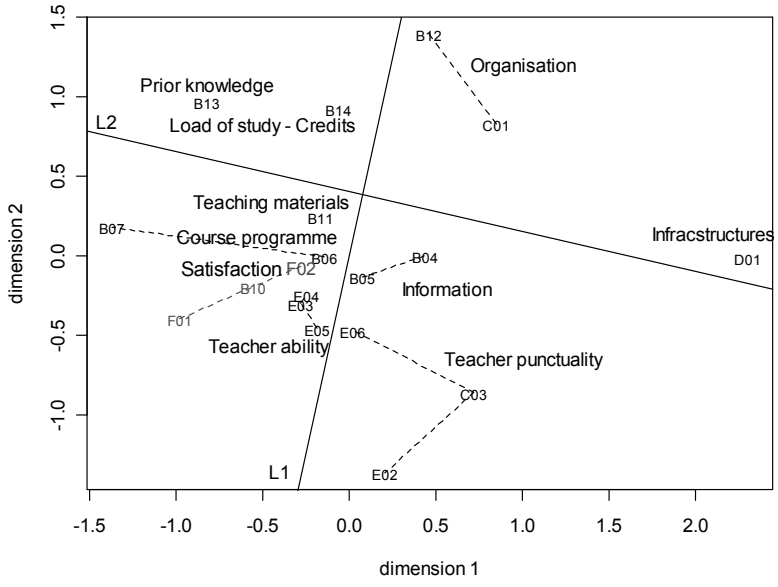
[11] For details on the evaluation system of University in Italy see Biggeri and Bini, 2001.

Figure 2: *Quality of teaching services components*



From the representation of the items of the questionnaire with a two dimensional non-metric scaling (fig.3) we can easily see the components in the picture.

Figure 3: *Nonmetric scaling of QT items*



The items in the fig. 3 are separated by two lines to help an interpretation of the dimensions: the first sets up the teaching activities items against, i.e. those on the left of L_1 are the proper ones, while those on the right of L_1 are the aside ones; the latter sets up the teaching activities items against, i.e. under L_2 there are those referring to the single discipline, while over L_2 there those referring to organization of the whole course (schedule, load of study – credits, etc.). What is important is the central position of the

satisfaction item (F02) in respect to all the items that identify the aspects linked closely to the dimension 'teacher ability' (items E03, E04, E05), with the dimension Information (on teaching, items B04 and B05), to the dimension Programme (of teaching, item B06 and B07), that is, all which is closely related to teaching. In other words, what emerges is that the quality of teaching as such is not influenced by the organisational aspects that remain marginal. The instrument of measurement used (point 3) is a self-filled questionnaire which, therefore, measures opinions that are characterised by uncertainty and subjectivity. [12] The choice of 'simple' indicators (point 4) aims to synthesise the distributions of the judgements expressed by the student for every item. Each item of the questionnaire is measured on the ordinal scale (four categories of answers). The starting point is represented by the following simple indicator (Capursi, Porcu, 2001):

$$1-\frac{1}{m-1}\sum_{i=1}^{m-1}F_{Ai}\,,$$

where A is a variable, m is the number of the categories of answers, F_{Ai}, the value of the cumulative function of the distribution of A, in correspondence of the i-th categorie. This indicator extracts the ordinal information and is invariant for all the symmetric distributions in respect to the median. This is a limit; it, in fact, does not take into consideration the variability of the distribution. The aim is then to construct an indicator that takes variability into consideration, and that is invariant for all the admissible transformations for ordinal scales. The suggestion (Capursi and Librizzi, 2004) is the following:

$$I=1-\left[\frac{1}{m-1}\sum_{i=1}^{m-1}\left(F_{Ai}\right)^{1/2}\right]^{2}.$$

This indicator has the following proprieties: it increases the more the quality of the 'item' increases; it is included between 0 and 1, in fact, $0\leq F_{Ai}\leq 1$, for the propriety of Cauchy also

$$0\leq\mu_{1/2}=\left[\frac{1}{m-1}\sum_{i=1}^{m-1}\left(F_{Ai}\right)^{1/2}\right]^{2}\leq 1\Rightarrow\ 0\leq I\ \leq 1\,.$$

Indeed $I=0$ in the case of the maximum concordance on the negative judgement is equal to 1 in the case of maximum concordance on the positive judgement. The weighting (point 5) is still subject to assessment as, plainly, all the dimensions do not have the same importance. A solution could be that of using Kendall's τ_b coefficient between each dimension and the item of satisfaction (F02); however, refinements still need to be found for the problems linked to the existence of possible associations between the items of a superior order in respect to the first.
The function which synthesises the indicators of dimensions in a composite indicator (point 6) should be sensitive to very low values even in a single dimension because the

[12] The questionnaire, prepared by the University, is available on the website www.unipa.it

bad functioning of this could soften the good functioning of the others. It is thought that such an effect could be had with a geometric mean or, if this should result too sensitive, with a power mean of order *r*, with $0 < r < 1$. The question, however, remains central: which measures for which objectives? If our expectation is the measurement of teaching quality and the objective is to receive the alert, attention should be drawn upon the measurement of the single components closely related to teaching, and, thus, leave information disaggregated. If the expectation is to measure the quality of the teaching services, with the aim of comparing each discipline or courses, the synthesis, even if with due caution, becomes necessary.

5. Conclusions

The course followed in the work, recalling the old debate on the principles and contents of the ways of knowing in social sciences, accepts in the indicators the instrument to bring closer the measurement of complex phenomena, in a systematic grouping of interpretative conjectures, of identification of relationships, of collocation in a functional model of explanation. Even in the awareness of the contingent value, of the uncertainty and fallibility of every restrictive approach of reality, conditioned by sensitiveness historically present and by the availability of instruments of observation, always perfectible, every construction of complex indicators must respect the accurate rules of measurement and aggregation on which it is, in particular, focused. The demonstrated example, in progress, points out the partial separation still existing between the convinced, needful, legislatively codified, intention to arrive at a judgement on a complex dimension (the quality of teaching on which many organisational and personal choices, often delicate, depend) and its realistic measurement, amply defined by the system of operational conditions effectively adopted, even if under the umbrella of a singular consideration on methodological rigorousness.

'Il pensiero semplificatore privilegia la disgiunzione o la riduzione. Il pensiero complesso deve praticare nello stesso tempo distinzioni e congiunzioni e deve tentare di concepire il molteplice' (Vianelli, 1989).

References

Aiello, F., Attanasio, M. (2004). How to transform a batch of simple indicators to make up a unique one? (to be published).

Attanasio M., Capursi V (1997). Graduatorie sulla qualità della vita: prime analisi di sensibilità delle tecniche adottate in *Atti della XXXV Riunione Scientifica della SIEDS*.

Bartholomew D.J. (1996) *The statistical approach to social measurement*, Academic Press, San Diego.

Bartholomew D.J. (2002) Discussion on the paper by Fayers, P.M., Hand, D.J.(2002) p.253.

Bernardi L., N. Torelli (1992), Le potenzialità offerte dalle statistiche sociali ufficiali dalla definizione di modelli e indicatori, ISTAT, *Avanzamenti metodologici e statistiche ufficiali,* Atti delle prime giornate di studio, Roma, 13-14 dicembre.

Biggeri, L., Bini, M. (2001) Evaluation at University and state level in Italy: need for a system of evaluation and indicators. *Tertiary Education and Management,* 7: 149-162.

Buratto, F. (1986) Indicatori e ricerca empirica in *Problemi di Statistica Sociale* edited by B.G. Resi . La Goliardica Casa Editrice, Roma, 17-60.

Capursi, V., Librizzi, L. (2004) Valutazione della 'qualità della didattica': quali misure per quali obiettivi? (work in progress).

Capursi, V., Porcu, M. (2001) La didattica universitaria valutata dagli studenti: un indicatore basato su misure di distanza fra distribuzioni di giudizi. *Atti Convegno Intermedio della SIS "Processi e Metodi Statistici di Valutazione"* Roma 4-6 giugno 2001

Cox, D.R., Fitzpatrick, R., Flatcher, A.E. (1992) Quality of life: can we keep it simple? *J.R. Statist. Soc. A*, 155, 353-393.

Curatolo, R (1972) Indicatori sociali. Atti della XXVII Riunione Scientifica della SIS. Vol. 1, 25-154.

De Lillo A. (1980), Cenni metodologici e schemi interpretativi della sociologia, in Livolsi M., *Sociologia, problemi e metodi,* Teti, Milano.

Fayers, P.M., Hand, D.J.(2002). Causal variables, indicator variables and measurement scales: an example form of quality of life. *J.R. Statist. Soc. A*, 165, 233-26.

Gori E., (1992)La valutazione dell'efficienza ed efficacia dell'istruzione, in *Atti della XXXVI Riunione Scientifica della Società Italiana di Statistica*, CEDAM, Padova.

Horn R.V. (1993), *Statistical Indicators for the Economic and Social Sciences*, Cambridge, Cambridge University Press.

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

Leibowitz, S.G., Hyman, J. B. (1999) Use of scale invariance in evaluating judgement indicators. *Environmental Monitoring and Assesment*. 58, 283-303.

Leonardi F. (1991), Contro l'analisi qualitativa, *Sociologia e ricerca sociale,* 35.

Panek, T. (1989). On the measurement of the level of living. *Metron* . **47** 361- 371.

Petrucci A., Rampichini C., (2000). Indicatori statistici per la valutazione della didattica universitaria, in *Valutazione della didattica con sistemi computer-assisted*, Cleup, Padova.

Poincaré, H. (1913) *The Foundations of Science.* Science Press, New York

Stevens Smith, S. (1974) Measurement, in *Scaling: a sourcebook for behavioural scientist*, edited by Maranell M. Aldine Publishing Company. Chicago.

Vianelli S. (1989) "Statistica e Scienze della complessità" in atti della XXXI Riunione scientifica della Società Italiana di Economia Demografia e Statistica, Palermo, Vol. XLIII, nn. 1-2, pp.15-67.

Wermuth, N. (1992) Discussion on the paper by Cox, D.R., Fitzpatrick, R., Flatcher, A.E.