EVALUATE OR MEASURE RESULTS.  
THE UNIVERSITY CASE

dott.sa Eliana Minelli  
Organization at University Carlo Cattaneo – LIUC,  
C.so Matteotti, 22 – 21053 Castellanza (VA) ITALY  
Tel. 0039.0331.572.269 – Fax 0039.0331.572.260  
eminelli@liuc.it

Prof. Gianfranco Rebora  
Rector University Carlo Cattaneo - LIUC  
C.so Matteotti, 22 – 21053 Castellanza (VA) ITALY  
Tel. 0039.0331.572.214 – Fax 0039.0331.572.260  
grebora@liuc.it

Dott. Matteo Turri  
Public Management at University Carlo Cattaneo – LIUC,  
C.so Matteotti, 22 – 21053 Castellanza (VA) ITALY  
Tel. 0039.0331.572.448 – Fax 0039.0331.572.260  
mturri@liuc.it

---

4-7 May 2005, Technische Universität
Abstract:
If it is true that in the last decades the whole issue of responsibility and accountability in public management has been rising, it is also true that there is a great confusion on how to appreciate the value of public action results. In this regard, the proposed paper illustrates the differences between the concepts of evaluation and measurement, verifying how the lack of clarity may imply relevant consequences in managerial field, in particular the outcomes observed in university field. After disclosing the consequences of a lack of awareness of this issue, the authors show the necessity to find rational solutions to meet the needs for responsibility and accountability, by means of differentiated approaches.

Keywords:
Evaluation, measure, management, university, responsibility

Track n.4: Responsibility in Public Management
“For with the judgment you pronounce you will be judged, and the measure you give will be the measure you get” (St. Matthew, VII, 2).

0. Introduction

The phenomena of the European integration and the consequences of the Bologna’s agreement retaken by the European Ministries in Berlin 2003, which will be the subject of discussion in Bergen, have encouraged the European states to adopt systems of evaluation in the area of higher education for themselves. Meanwhile for England, Holland and other Northern countries, they are already considering a second generation system of quality assurance (Jeliazkova and Westerheijden, 2002), other nations deal with these topics without experience but with the need to recover lost time.

The consequence of this is a certain bewilderment and the diffusion of practices and rituals. This paper refers particularly to the Italian experience and examines a distortion which may characterize the most recent evaluation systems: the difficulty in distinguishing between the evaluation and the measurement of results. It is a phenomenon rather common which has its origin in the rushed adhesion to methodologies of survey without a full comprehension of their exact potentiality.

1. Ambiguity of measurement
The etymology of the term “measure” appears to be full of ambiguity. The Italian ‘misura’ and the English ‘measure’ derive obviously from Latin ‘mensura’ and, therefore, the verb ‘metiri’ and from its participle ‘mensus’. This seems to mean judgment, measurement in a sense more psychological than physical, and also prudence. Therefore, it is not just a recall to an objective method of calculation but also a well-pondered and reflexive exercise of evaluation. The root of the verb ‘metiri’ and the Greek equivalent ‘metrein’ refer also to ‘mes-s’ (Italian ‘mese’ or month) which evokes a basic unit of time measuring, and which recalls also ‘moon’, a term indicative of the lunar cycle. (Mari, 2003).

However, the title of the Shakespearean play ‘Measure for measure’, gives entirely this ambiguity, reminding implicitly one of the verse of the Matthew Gospel reported in the epigraphy, where the terms “judge, you will be judged” (on a subjective level) and “measure, will be the measure you get” (on an objective level) seem at the same time to contrast, coincide and complement.

The awareness of this ambiguity is present in the field of the science of measurement itself, in that a recent article on Epistemology of measurement expressed the problem explicitly: “the measurement is a subjective evaluation or an objective description? (Mari, 2003)”.

The ambiguity, in any case, is a normal condition of the functioning of all organizations, because different meanings and interpretations can be given to their input and output, as well as to the currents of activities which they came across (Weick, 1979).

The ambiguity is more so for at least a part of the spheres of activities which pertains to the public management (Noordegraaf & Abma, 2003).
Our paper aims at approaching the theme of measurement of organization results, focusing on those contexts, like universities, characterized by a limited efficacy of the market and by conditions of strong complexity and relational valence of products or service.

2. Evaluation and measurement

To correctly enter into the subject, it is necessary to start from a science of measurement point of view, in whose ambit it is possible to reflect on some conceptual assumptions more deeply than it normally happens in the essays about managerial disciplines.

Taken for granted the ambiguity of the term measure and also the historical evolution of some basic conceptions, today it seems reasonable to talk about measurement in the ambit of an empiric science when some conditions are present and more precisely, in extreme synthesis, “that the measurements are evaluations provided with inter-subjective and objective character” (Mari, 2003). Thus, it is necessary that:

- the results of measurement give the same information to different observers;
- the results of measurements give information linked to the object to be measured only, and not to its environment, which can include the observer, and that measurements of the same thing in different conditions and in presence of different observers give the same results (Mari, 2003).

This means to consider the role of the systems of measurement not only as operative instruments but also as essential components which give epistemological foundation to the measurement.

This way the measurement would be differentiated from a generic evaluation, which is still basically a form of subjective judgment, and assume, in respect to it, a superior value, in that evaluation whose results can in a proper way be considered as objective. According to this approach, the science and the theory of measurement have their own future in facing the
problems of formalization of the structure, and of the operative function of the systems of measurement referred to a phenomena more and more complex.

3. Management, measurement and evaluation

However, management is not a science, it is a practical activity which also has aspects of art. It differs from other types of intellectual activities because it implies “a practical art”, more than an “applied science” (Carlsson, 1951; Mintzberg, 1973). It is always hard to give an adequate and objective description to artistic work. It is not just a matter of applying well-defined principles to determined cases, as would a designing engineer. Practical art cannot be based on well-defined principles and on the application of a systematic knowledge because it does not have a complete and exact foundation of this kind. The knowledge on which it is based is fragmentary, not completely defined, continuously struggling with the sphere of the unknown. The rules of reference ask for an active interpretation from the subject and therefore judgment, intuition and personal ability of creative type. Managers act in a disorderly way to absorb the ambiguity of the context and to permit the rest of the company to operate systematically; they act in a context of calculated chaos, where it is needed to give a convincing interpretation of a task which does not have defined borders, with open ends (Mintzberg, 1991). The superficiality is the central problem that management has to face, the lack of depth in thinking and acting which derives from multiplicity of pressures, from work overload, from the need to handle simultaneously different roles, projects, contacts, information. And in this matter, managers do not find help in the prevalent concept of the managerial professionalism as a whole of practices, concepts and instruments to be used independently from the diversity of the company contexts (Mintzberg, 1991; Mintzberg 2004). This refers more and more also to the analysis and interpretations of the company results, as demonstrated by the widely used instruments of analysis and evaluation showing
multi-dimensional characters, which reflect a versatility of references (Coda, 1989; Kaplan & Norton 1996). It is true in fact, in order to carry on their tasks, managers need to have additional analysis instruments, which reflect the criteria of objectivity and inter-subjectivity which are commonly practiced among the students of the science of measurement. They also need to take decisions in conditions of imperfect and incorrect information; and not always is it the best conditions to count exclusively on intuition requirements, or instinct. In this order of ideas, the instruments based on the logic of evaluation gain importance, which is not a measurement in a scientific sense but it comes from a rational insight and is expressed sometimes also in a formal way. From this point of view, in a managerial ambit, it seems to be unacceptable to consider the measurement as an instrument of superior quality comparing to evaluation, a kind of reference which the latter should tender. In reality, for many phenomena where the human components and, most of all, the interpersonal interactions have a central role, where there is a system of ‘calculated chaos’, the evaluation, with all its load of subjectivity and imperfect communicability, or of non-transferability in other contexts, is without alternative in measurement and even it is preferable. It is the same concept expressed in another ambit, the one of the determination of economic value of the company capital, by an acute scholar: “the more founded conceptual methods and, for this, more rational, outcomes mainly are influenced by subjective elements” (Pignatelli, 1990).

This is more so in the case of public management, and in particular in some areas of activity in the public sector, characterized by processes at the same time not routine, unclear, innovative and conflicting (Noordegraaf & Abma, 2003). It is especially in these contexts, where practices defined as ‘inappropriate’ are brought in, that the pretense of the movement of the new public management to solve the problems of uncertainty through managerial methods based on measure of performance, is faced with insuperable limits. In these cases, in fact, the definition of objectives loses its own meaning of reference and the relationships
between ends and means dissolve, also beyond the limitations clarified by the classic analysis of Simon on judgment of fact and judgment of value (Simon, 1947): in the sphere of action as the intervention to develop employment, or to defend the natural environment, or also to limit the traffic congestion, the interconnection between facts and values is too narrow to prevent a definition of objectives and a consequent measurement of the results which is not sensible to a divergence of interpretations and nor is it exposed to many visions. The relations between input and output cannot be represented in a stable way; and the results researched are often indirect comparing to the activated operative processes; furthermore, some of the fundamental components of outcome are ‘invisible’ because of long term or interpretable in conflicting manners and so the measurements of the performance become strongly arbitrary. The different measurements are not in anyway comparable in time and in space because the practices are unique and moving, they are unstable.

It is necessary to discuss the presence of these conditions in the activities object of our interest. However, in general, we can say that, based on the analysis carried out in the ambit of managerial science, management cannot be based only on measurement, and not only on this and the creative and spontaneous intuition; it is also needed the contribution of interpretation and evaluation, formally expressed and communicated to an inter-subjective level, even if without the requisite of the objectivity: different observers do not draw the same conclusions elaborating the same data through the same procedures.

4. The evaluation of universities and measurement of results

The university is interesting in order to reflect upon the relationship among management, evaluation and measurement. In the last years, a strong movement of reform is spreading all over Europe and has brought forth the institution of bodies, systems and procedures aimed towards evaluation of main results of the university activity, such as teaching, research and
different additional services offered to students and to other stakeholders. The concept of results evaluation and the quality assurance have become obliged references for the reform process of university systems, also as a necessary counter balance of the autonomy given to the athenaeums. Many countries have instituted national agencies with the aim to evaluate the output of the university activities, though in different forms. (Felt, 2004). The policies of regulation of universities promoted by various governs naturally are combined and integrated with systems of performance management and other practices connected to the movement of the new public management, starting from leading experiences as in Great Britain.

In the ambit of evaluation activities, even if in the multiplicity of methods proposed and experimented, a wide space has been taken in attempts to measure specific aspects of products and services given.

Many people sustain the importance of the indicators, variably composed and combined, as objective measures of efficiency and efficacy of diverse activities. The indicators are considered also as a possibility and natural reference for a ranking of athenaeums according to the quality level and therefore for the allocation of public resources destined to superior education.

In Italy, since 1989, different legislative interventions have made more valuable and also promoted the autonomy of universities, gradually modifying the pre-existing strongly centralized situation. These norms have progressively built a national system of evaluation of university. The National Committee for the Evaluation of the University System (CNSVU) is the central organ of this system, with the task to address the evaluation activities. Autonomy and evaluation have been conceived as complementary aspects, destined to fulfil the role of spearheading ideas which are pivoting points of a complex process of transformation.

According to this reasoning, the evaluation is conceived as strictly connected to a process of programming and controlling of the entire system; and the national body has tried., in the first
place, to activate a wide informative patrimony, systematically conceived and comprehensive of quantitative data banks and qualitative survey. The effort has been to give to governing bodies of the system, but also for to universities, students and their families, and third parties involved, an instrument of transparent and reliable information and, at the same time, timely and continuously updated.

The same information has also been utilized for the aim to re-orientate the athenaeum system of financing, within a logic of re-balancing the financing in respect to the parameters considered more indicative of size and quality of activity carried out. This purpose, imperfectly and only partially achieved in the first phase, has become one of the fundamental objectives of the development of evaluation activity. The CNSVU has elaborated documents which propose a radical revision of the system of financing for athenaeum, and it has more and more orientated its efforts towards the example offered by systems of financing and government of Anglo Saxon universities which, quoting the words of one of the CNSVU members, “They are meticulously governed through a prudent distribution of funds based on parameters previously known” (Figà Talamanca, 2001). The indications of the Minister of University and of the National Government are directed in this sense, too.

In a special issue of the MIUR magazine dedicated to the evaluation and directly edited by the CNSVU, it is written: “….. gathering data is not just for the intransgressible need to have a global knowledge of the system which we did not have before (…) the main objective is to understand how data can be useful for an effective management of the autonomy and for central government of public resources”. (CNSVU, 2002a).

The same CNSVU tends to present its role also in another way, not as “controllers, or external evaluators, but as a partner and this should eliminate the situation from suspicion, from the accusation to be the one who gives evaluations, who sets terms for others to work” (CNSVU, 2002b).
On the other hand, the experience of countries which have faced for a long time and with more determination this way enlightens the difficulties and the obvious contradiction to render compatibility the two diverse logics of the thrust towards efficiency, through forms of regulations and interventions from above, and of the scheme of professional and organizational development at university level, gearing to the improvement of quality and efficacy (Pollitt, 1987; Barzelay, 1996).

Comparing the different international experiences, we can acknowledge that the process of quality assurance does not function in the same way within universities as inside enterprises (Felt, 2004). In general, the results of universities are not solely internal but they are also considered at a political level and have a wider repercussion (Felt, 2004). Even more so, this happens when these same results are being linked to formal processes which influence the distribution of resources, or in forms of ranking which position in different ways the different universities and which are reported to the public. It is more than natural, in this context, that universities develop strategies to hide their own weaknesses and to emphasize their points of strength: these are behaviors which obstruct and frustrate the conditions of an open process of quality improvement (Felt, 2004).

In Great Britain, the system which seems to have become the main reference for the most recent Italian experiences has raised criticism and is being considered the manifestation of a regulatory state and of an audit society which, for efficiency reasons, are striving to erode the traditional autonomy of athenaeums (Brown, 2004). In any case, the Anglo Saxon system represents an attempt to regulate higher education by means of inducing a kind of simulated market: in this, the extensive use of parameters and measurements carries out its own role, which is also the one to maximize the widespread information about the results of different universities with transparency and objectivity.
Anyhow, this process seems to be incapable of overcoming, at the end, a fundamental contradiction: a careful exam of methodologies and instruments of evaluation used to demonstrate that none of these in practice corresponds to the double criteria to have a value of true measurement, in the sense indicated by Mari, and to have a meaning fully sustainable in a finalist point of view.

It is clear in fact that a good part of techniques utilized in the various European experiences for the evaluation of didactic and research can be converted in a formal expression of quantitative parameters that, however, do not constitute measures equipped with inter-subjective and objective character. In conclusion, it does not occur that the formal expression of quantitative parameters gives the same information to different observers; and that the measurements of the same thing in different conditions and in presence of different observers produce the same results (Mari, 2003).

This is true for example for all the group of methodologies based on peer review, applied both to didactic and research, which have very wide use all over Europe. The external experts, who carry out visits and inspections, examine data and publications, participate in representative moments of evaluation activities, can follow detailed instructions, use homogeneous guidelines and reference manual, participate in common learning processes, deal with bodies and roles of coordination. All this places a tension on expressing objective evaluations but the center of this method consists in relying, at final instance, on the experience and capability of judgment of selected people as peer; and therefore there is a limit to the degree of coherence of diverse judgments (Clark, 1996). The analysis of the English empiric experience has shown that the countermeasures adopted are not enough to guarantee the uniformity of judgment, more so when diverse disciplinary spheres are evaluated, in research or didactic. There is no guarantee that these references can build systems of measurements in a proper sense, at the
most, they can achieve a satisfactory coherence of the evaluations expressed by the same group of peers, or referred to the same institution (Underwood, 2000; Brow, 2004).

All this is even more obvious for the other relevant group of methodologies, the one based on evaluations, in this case declaratively subjective, from students, or other addressees of different activities. The next paragraph deals with issues related to the evaluation of didactic by students. The risks in assigning, when forcing, an objective value to evaluations which have different nature are well represented from the controversy developed in England on the choice to prefer a logic of grading to a threshold. If the judgment, regarding an academic institution or its program or project, is limited to recognize the achievement or not of a quality threshold-level considered as necessary requirement, the problems of measurement objectivity are limited if not minimized. By insisting on a grading of judgments, which has the consequence of institutionalizing forms of ranking, the problems of reliability and defensibility of the opinions expressed are being emphasized. When summative judgments are reached, univocally expressed on a quantitative scale, we inevitably end up attracting the attention on that part of evaluation which has less value and validity, for simple technical reasons, with effects on the credibility of the system adopted well foreseeable by the same evaluators, by the evaluated and by the other possible addressees of the evaluation reports (Sharp, 1995).

Regarding, instead, that part of methodologies adopted which suits the requirement of a real and proper measurement, the limits of an excessive simplification emerge. The parameters and the quantitative indicators, whatever chosen, do not reveal sufficiently as meaningful as general expressions of those characteristics which the regulator bodies would like to bring the value up.

Therefore the critic of Trow is valid (1994): “may well be that vigorous effort by agencies of central government to assist the quality of university work lead to his decline, as more and
more energy is spent on bureaucratic reports, and as university activities themselves begin to adapt to the simplifying tendencies of the quantification of outputs. Our research suggests that departments and individuals shape their activities to what ‘counts’ in the assessments, to the impoverishment of the life of university, which is always more complex and varied than assessment of ‘outputs’ can capture”.

In the end, the conclusion of Power (1997) results to be convincing, which sees it being put into place the formal practices of assessment from a point of view of the legitimization of the decisional bodies. The extension of these practices in different contexts and, most of all, in more complex professional activities, such as education and health, contributes to rationalize and strengthen the public image of control, more than representing a real technical answer to problems of responsibility and management. The elements present of the of these practices in objective measurements fits completely into this necessity.

If these critics are valid for experiences consolidated and pursued with determination, like in England, they fit even more for initial and ingenuously conducted approaches like the Italians’ in the recent years. The compromise and the attempt to conceal the contradiction through behaviors of central bodies of evaluation, which swing between the authoritative style and the paternalistic one, are obvious here.

5. An example: the evaluations of students on didactic

What is stated in the previous paragraph is true for the evaluations expressed by students on quality of didactics (SRT, Student’s rating of Teaching). It is one of the main instrument established all over the world for the evaluation of didactics and the most emphasized in the communication towards the students and towards the public opinion.

On the other hand, the use of these methods, at least in their hard version as a fundament of a reward system, and of a conduct of teachers with managerial criteria, has been the object of
criticisms in international literature. The practices of the SRT have, as a result, the ‘commodification’ of the students, the reification of the relationship teacher-student and a very strict concept of the professionalism of teachers and of academic responsibility (Lawrence & Sharma, 2002; Singh, 2002). At the end, the will to assimilate students into ‘consumers’ with a very narrow view of the university experience is apparent.

A part of these reactions derives from, probably, ideological reasons, or from forms of academic conservatism but it is in any case difficult to accept the projection of the SRT outside the ambit of single university towards the entire system. What makes sense as a managerial practice, with its reasonably acceptable load of subjectivity, loses surely meaning in a wider ambit.

The attempts of the Italian University Minister, declared by the press, to assume the results of SRT as an important reference for the allocation of financing to the different universities and for the other ministerial interventions, arouse astonishment. The fortnightly gazette of the MIUR on the 30th of June 2004 stated: ‘the complex of evaluations, which will be available regarding the didactic within September, will allow the MIUR to acquire important and significant indications about the correspondence of the services delivered by universities to the students’ needs in a new frame of students satisfaction surveys. (…) The periodical results of these surveys – has been explained by the Minister– will give an objective picture of the quality of university services in a way to allow, also through measures of support, incentive and discouragement, the adoption of concrete initiatives for the government of the university system in an optic to pursue the quality and the efficacy of the services’ (MIUR, 2004).

The collection of cards through which the students express their evaluations on didactic, provided by law and oriented by an official model elaborated by the CNSVU, misses in fact any kind of premise which justifies this choice.
On one hand, all the process does not offer any serious guarantee of reliability since it is left to the autonomy of athenaeums and their board of evaluation.

On the other hand, and this is the fundamental objection, which is not recoverable with more accurate organizational modalities, it cannot be seen how the formal expression of opinions from students can satisfy the requirements for a reliable measurement, given the diversity of contexts which are confronted on a score scale.

6. The consequences of a managerial orientation for evaluation of universities

The base research for objective measurements for the evaluation of universities has to reckon not only with the weak and scarce persuasiveness, just from the point of view of measurement, of the methodologies more frequently adopted, but also, and most of all, with the different requirements which put pressure today on the university management.

The same autonomy given to universities implies that each of these should commit itself to relationships with the external reality, establishing links and obtaining a position in the environment, until a definition of a proper specific strategic profile, and to orientate, as a consequence, their own actions in diverse sectors of interest. Other trends, as the competition, the internationalization, the acquisition of financed researches on order, the inclusion in network of collaboration both at national and at local level, push towards the construction of a profile and a specific identity of each university.

In answering to these pressures and opportunities, the athenaeums cannot avoid handling the fine line in relationships and facing negotiation processes with other players. The autonomy reveals more and more as the result of a network of negotiations which involves partners of different levels, bodies and differentiated subjects. The university management has, this way, to define strategies which orientate the selection of the relationships considered worthy of investment, the links which are useful and worthwhile to undertake (Felt, 2004).
It is not a surprise that everything results in a differentiation of the university profiles, in a strong heterogeneity of the respective characteristics and lines of behavior.

It seems also more difficult, in this context, to direct the attention to the scale of measurement of the university activities and of its results which should be valid independently from the contexts.

This is the case, frequent in Italy, of recent universities, born in suburban local contexts and finalized, above all, to a function of cultural incentive, of offering innovative formative services, of animating the social economical reality and rather isolated. A comparative evaluation of the research results, in respect to, for example, metropolitan universities with strong traditions, seems to have insignificant meaning, due to the obvious difference of objectives. The problem is not the probable penalization of these peripheral universities, but the foreseeable distorted effects, for the orientation of research. In many thematic sectors, in order to pursue their objectives, universities of this kind should orientate over thematic of local interests, privilege the inclusion in networks of relations proper to the adjacent social – economic area. The orientation of teachers, encouraged also by the priorities of the national evaluation, towards the reference of different circuits, also of national and international importance, will produce an eradication from the territory, substituted by a probable vague and inconclusive research of achievement in wider ambit, or by a ‘colonization’ from the part of research groups expressed from stronger athenaeums, which these teachers will consider as main reference at the end, in all these cases, with scarce benefit for the strategy of inclusion in territorial context.

In reality, the same concept of quality tends to change the meaning depending on the specific order and on the positioning strategies chosen by each university (Felt, 2004). This, in any case, perfectly corresponds to the most evolved interpretations of the evaluation results of company management, according to a strategic and multidimensional perspective.
It is acknowledged that the identical result expressed by one indicator for two different companies, or also for different markets and competitive positions experienced by the same company, does not imply the same evaluation.

For example, indicators referred to the costs of contractual agreements and to the costs of the company conflicts, with reference to union relations, do not seem to constitute representative elements of a generated value without considering the context. In different environments, the correct reading can give different interpretations.

To correctly read the results of a process and to understand the value generated, it is always needed to focalize the attention on more critical and essential aspects.

However, these are firm specific, which are connected to the typicalness of the operative processes of a company and also to the list of clients, competitors, intermediaries really present and involved in a certain time and place.

Instead of talking about measurement, it seems to be correct to refer to an evaluation of results obtained in relation to a process; evaluation implies judgment, assignment of value; it is a result of analysis and reasoning and it does not want to be objective.

In this there is, most of all, the reference to a perspective of management, which is different from the one of ‘measurement’: in a managerial perspective, the level (meaning the value expressed quantitatively) of a variable, is just relatively involved. Let’s think about a survey on personnel satisfaction, for example on a scale 1-5: to know that an average level has been reached, let’s say 3.6, compared to an average level for other companies of 3.8, it says little in a managerial perspective. Disregarding the uncertain reliability of surveys of this kind carried out in different contexts, not even the links are clear of this variable with the levels of productivity of work and the company competitiveness - it is difficult to give a positive or negative meaning to the gap measured in this way. This use of measurement, in a managerial key, requires attention not only for the level or the value observed but, most of all, for the
cause and effect relationship between the analyzed result and the objectives that the company pursues (Becker, Huselind & Ulrich, 2001).

We give meaning to numbers and to surveys from a point of view of the company finalism. And to give meaning, it is needed to place the result which we are trying to define in a context, the one important for the company, underlying the relation with a perspective of improvement which a positive value is assigned to.

A consideration in a strategic perspective of the critical processes for a company, or for an institution like a university, makes it really difficult to single out general references: the predominance of qualitative and immaterial elements tends to produce situations firm specific, up to the point that the same benchmarking experiences have scarce value, if orientated towards the simple comparison of parameters (Becker, Huselind & Ulrich, 2001).

7. How to combine evaluation and measurement with reference to universities

The methodological proposal which is intended to delineate tries to take into account the thrust towards the quality improvement, which cannot be contested beyond the limits and the defects of the current applications of evaluation and of practices and techniques inspired by the new public management. There is no doubt, in fact, that universities have to give credible answers to the request for responsibility and reliability which come from the society and students, leaving from a proud state of self-referenciability. (Hoecht, 2004).

This does not mean that we need to specifically follow the way of measurement and indicators, demonstrated to be sterile, at least as a main reference, in all the cases in which we had to deal with complex objects, influenced by multiple inter-subjective relationships, uncertainty and variability of the context. Neither do we need to insist in the search of false measurements only to respond to a generic need of security and objectivity at a superficial level.
Instead, we think it is useful to propose a trace of reference for the analysis and evaluation of the university quality which takes into account and leaves space for the ambiguity and the demand for interpretation, which are inherent to the core of academic activities, research and teaching.

It is really the research of conceptual strictness which pushes us to point out this aspect. We need to make an effort to comprehend what really qualifies the university core business, the didactic and research, considered also in their connection, in order to understand which space can be taken by measurements, or statistic surveys, even sophisticated, and which role can have other ways to express the attention towards the results of these activities.

A good package of structures, services, professional resources, the analysis of volume of developed activities and other aspects which are used for a real measurement constitute, in this logic, a necessary foundation, a base, a kind of pedestal or platform, which can be referred to as “static quality”. It is a task for university management to guarantee these conditions, also through proper systems of management control and quality assurance.

It is above this ‘threshold of acceptability of performance’ which we meet the critical problems for the university management and for the evaluation itself. It is also necessary to have clearly in mind that once these conditions are guaranteed, through proper quantitative measurements taken cared by bodies technically reliable, the problem to promote the real quality of the university study meets with other aspects, which involve more directly people and active and creative interaction among different subjects, such as teachers, students, technicians and professionals of the various services.

The real quality is dynamic, interactive, evolving; it begins with a combination of relationships; it requires to be continuously renovated (Pirsig, 1981 and 1992). It is influenced by environmental factors, often incomprehensible and without meaning if taken out of context.
As it was documented in another ambit, didactic and research can be measured, as far as the elements which bring the static quality are concerned/which is very useful for elements which constitute their pedestal, their base platform (Rebora, 2003; Minelli, Rebora & Turri, 2002). Therefore, the static quality can also be measured. The dynamic quality belongs to another conceptual order, and it can be evaluated, according to methods and logics which are inevitably filled with subjectivity.

It is a sphere of subjectivity which can be tempered, as it happens by adopting the group of methodologies of the peer review (Minelli, Rebora & Turri, 2004). But it is necessary to be aware that at this point we are operating in the horizon of the evaluation, expression of human and subjective faculty of the judgment, of the profound intelligence of things and situations. It’s a faculty which is convenient to support and link with all possible measurements but which does not have, in respect to these, lesser relevance and value, as autonomous modality of development of the knowledge and of the capability of governing and management of complex systems.

The application of these concepts is useful at both levels significant for the intervention on quality, which are needed to distinguish with clarity: the level of the national and international system of universities, in which the central bodies of promotion and quality assurance have sense/a main role; and the level of the single university and of the actions of the correspondent governing and management bodies.

The whole system of results evaluation and quality assurance should foresee a limited use of rules, standard and requirements, based on objective and incontrovertible measurements, imposing a general respect. Also the recent proposals of Brown (2004) go towards this sense: “the way to raise quality is to make accountability – in the sense of compliance with standards and rules – only a small part of quality assurance, and to appeal instead to the
professional motivation of staff to do better by engaging in a constructive, professional dialogue...”

For all the rest, it is needed to rely on the capability of management of different universities, which the central bodies can only promote and support with soft interventions, But also at this athenaeum level, the need to combine in a proper way the interventions which guarantee the pedestal or the platform of the static quality emerges, through suitable measurements and attention for standards and rules also severely imposed. At this level, it will be important, for example, to guarantee the respect of the hours for the different activities, to monitor the real attendance of students, to guarantee the correct procedure of exams.

For the rest, the reliance on the professionalism of teachers can find stimulus by adopting the instruments of self-evaluation and also a collection of opinions of various stakeholders, and therefore also instruments as the SRT. The peer review can have an important role. But these are instruments which are important within the university and in the ambit of its relationships with reference to the environment. If they become steps and instruments of an external regulation, these methods can be distorted and lose relevance for the processes of improvement and organizational learning.

The European University Association seems to have adapted this position, in the official documents: “Evaluation approaches that are based on standards, quantitative methods, sets of criteria, or checklists will not improve quality meaningfully and may not even control it significantly because they will not capture the complexity of the educational enterprise” (Eua’s quality assurance policy position in the context of the Berlin communique: code of principles for external qa in Europe - 12 April 2004)

The proposal seems to be for internal quality system fitted for the objectives autonomously adopted and geared towards gathering the quality elements of academics process and leading towards improvement “Universities are committed to developing, embedding and
"mainstreaming an internal quality culture that fits their institutional mission and objectives” (EUA Conference, Glasgow 2005 "Strong universities for a strong europe").

It is advisable to gear towards introducing an internal quality system, associated with external evaluation, at the level of system, through institutional audit, which are to stimulate and guarantee the quality level of single athenaeum, also focusing on checking that internal monitoring has been effectively done.

In conclusion, it is correct to state that confronting with numbers and statistics is important for universities, like any other public institutions, because they are forced to deal with facts and real data. Also the effort to achieve measurements endowed with a good degree of reliability and objectivity in the sense above explained is of great importance for institutions which are asked to demonstrate their own accountability and responsibility. We do not want to debate this. But it is needed that measurements are placed and comprehended in their right area without being confused with other instruments, which are not measurements, and which represent diverse modalities to meet the reality. From the management point of view, this second kind of instruments is destined to assume relevance in contexts as the university, which carries out a function unlikely referable to univocal standard.
References


