PERGAMON

Higher Education Policy 14 (2001) 75–90

—HIGHER— EDUCATION —POLICY—

www.elsevier.com/locate/highedpol

Forum

Evaluating research performance: the strategy of the University of Naples Federico II (Italy)

Gianfranco Carotenuto[a,*], Mariagloria Lapegna[b], Giuseppe Zollo[c], Alberto Di Donato[d], Luigi Nicolais[a]

[a] *Department of Materials and Production Engineering, University of Naples 'Federico II', Piazzale Tecchio, 80-80125, Napoli, Italy*

[b] *Office for Research and International Relationships, University of Naples 'Federico II', Via Mezzocannone, 16-80134, Napoli, Italy*

[c] *Department of Information and Computer Science, University of Naples 'Federico II', Piazzale Tecchio, 80-80125, Napoli, Italy*

[d] *Department of Organic and Biological Chemistry, University of Naples 'Federico II', Via Mezzocannone, 16-80134, Napoli, Italy*

Abstract

The evaluation of research activities is a complex task. This complexity derives from two main factors: 1.) the difficulty of defining objective and reliable measures of the "advancement of knowledge"; 2.) the presence of many potential users of the results of scientific activity, such as scientists, technicians, engineers, manufacturers, students, and consumers, all of whom impose many different demands. In the absence of objective criteria, the University of Naples Federico II (UFII) has chosen to implement a set of measurements over time, beginning with the most reliable parameters, such as the Impact Factors, and then adding new quantitative and qualitative considerations as the development of the evaluation culture within UFII. This article describes some of the issues related to the first step of the implementation of this strategy.

Keywords: Evaluation; Scientific activity; Journal Citation Report; University of Naples 'Federico II'

1. Introduction

The evaluation of research activities is a problematic issue, as relationships between research activities and research results are difficult to establish for two main reasons:

1. The real concept of "result" is ambiguous, as the research market does not function in the same manner as does a common goods market. In other words, while traditional

* Corresponding author.

0952-8733/01/$20.00
PII: S0952-8733(00)00022-2

markets establish the value of goods through an exchange mechanism, this is not the case with research outcomes;
2. In many cases, the social impact of research activities can be assessed only after months and years following the first results, especially in the case of results that represent a breakthrough with respect to common knowledge.

As the value of research activities cannot be assessed on the basis of objective, timely and reliable indicators, we have to assume that the value of research activities is a socially constructed convention, and that this value is justified through different methods, such as retrospective methods, qualitative peer reviews, and quantitative Methods. The most important condition is that each scientific community should choose its methodology for assessing the value of its activity, and should choose the path to follow in order to improve that methodology. The following part of this article describes the first step taken in establishing the research evaluation system at UFII.

There is a wide variety of approaches that may be taken in defining organizational effectiveness. Generally, these different approaches emerge from different concepts of the meaning of an organization: organizations have been viewed as rational entities in pursuit of goals; as coalitions reacting (or proacting) to strategic constituencies; as individual need-meeting cooperatives; as meaning-producing systems; as information processing systems, and so on.

As the concept of "organization" changes, so do the definitions and approaches to organizational effectiveness.

Four approaches towards defining organizational effectiveness have received particular attention. The most widely used is the *goal model* (including both operative and official goals), which defines effectiveness as the extent to which the organization accomplishes its goals. One problem with this approach is that an organization may be judged to be effective in areas outside its goal domain. For example, NASA was very effective in the 1960s in producing useful consumer products aside from its primary goal of reaching the moon, and the Liggett and Myers Tobacco Company became so successful at selling dog food that it was sued in the early 1970s for holding a monopoly. In addition, the organization may be ineffective even when it accomplishes its goals if its goals are too low, misplaced, or harmful. For example, Boise Cascade set a goal of increasing company earnings by 20% each year and met that goal for 12 years. However, in order to do so the firm developed a norm of taking on risky projects that led to its demise and forced reorganization in 1972.

The second approach to measuring effectiveness is the system *resource model* which focuses on the ability of the organization to obtain needed resources. Inputs replace outputs as the primary consideration. Organizations, however, may prove to be effective even when inputs are not optimal, and when a competitive advantage in the resource market place does not exist. For example, the "no name" Seattle Supersonics did not succeed in attracting superstars for their team but still reached the NBA championship finals in 1978, and won in 1979. Furthermore, it has been suggested that in non-profit organizations the acquisition of inputs is not tied to the production of outputs. Consequently, resource acquisition (inputs) cannot be used as a legitimate criterion of effectiveness.

A third approach is the *process model*, wherein effectiveness is equated with international organizational health, efficiency, or well-oiled internal processes and procedures.

It still remains true, however, that an organization may be effective even when organizational health is low and internal processes are questionable. In addition, in turbulent external environments, the presence of organizational slack (unused, convertible resources) may indicate an inefficiency in internal processes while being nonetheless essential for long term organization survival and adaptability.

The fourth apporoach is the *ecological model*, or the participant satisfaction model. Effectiveness is defined in terms of the degree to which the needs and expectations of strategic constituencies are met by the organization. This approach can be viewed as a summary measure for an organization (e.g., where score variance becomes critical). It mainly emphasizes constituencies outside the organization, and the most effective organization at least minimally satisfies, or reduces, the regret of these strategic constituencies.

Each of these approaches has certain advantages as a research and theoretical tool, but each also has its own weaknesses. Because each of these models is analytically independent, one approach may be appropriate in certain circumstances, or with certain types of organizations for which other models are not appropriate. One major consideration in determining which model is most appropriate in assessing effectiveness is the domain of activity in which the organization operates. For example, the strategic constituencies approach may be most applicable in an organization operating in multiple domains, where outcomes are obscure, or when the organization is required to respond to a diverse group of constituency demands. The goal model, on the other hand, is not appropriate in those types of organizational settings, but is most appropriate where organizational domains are narrowly defined, where goals are consensual, or where outcomes are easily identifiable (including both operative and official goals).

1.1. The effectiveness of research activity

In the light of the previously described issues, the University UFII of Naples started the evaluation process of research activities by defining a general model of the research process. Variables to be evaluated are defined with reference to this model. The most general representation of the model is described in Fig. 1, where five main components appear:

1. Research Activity, which is performed by an Organizational Actor (OA). In the first implementation of the model, OA is the University Department, but in a further refinement of the model it might correspond to people belonging to the same Scientific Area, or Research Group;
2. Input Resources, which consist of financial and organizational resources, people, skills and competencies, technical facilities, research traditions, past experiences, and relationships with other research groups or with external agents, such as private companies. In the first application of the model only financial resources are considered;

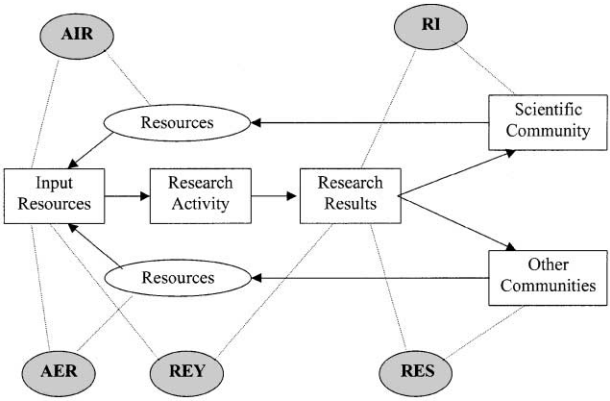

Fig. 1. The evaluation model.

3. Intangible and Tangible Results, which consist of scientific results, patents, prototypes, procedures, practical methods and techniques, software, and relationships with other bodies. In the first application of the model only scientific results are considered;
4. Scientific Community, which represent the first user of research results;
5. External Relevant Community, consisting of companies, institutional bodies, and consumers, which are the end users of research results.

These five elements define a closed system within which the research process takes place. According to this model, the activity of evaluation focuses on the following set of measures, which implement some of the issues raised by the above four models:

(i) *Research impact* (*RI*), which measures the acceptance of the research results by the Scientific Community, within the purview or perspective of the goal model;
(ii) *Research effectiveness* (*RES*), which measures the acceptance of the research results by the External Relevant Community, within the purview of the ecological model;
(iii) *Research efficiency* (*REY*), which measures the amount of tangible and intangible results per unit of input resource, within the purview of the process model;
(iv) *Access to internal Resources* (*AIR*), which measures the ability to attract resources from the Scientific Community, within the purview of the resource model;
(v) *Access to external Resources* (*AER*), which measures the ability to attract resources from External Relevant Community, within the purview of the resource model.

This approach implicitly assumes that the results of research activity should be considered from several points of view, as it is impossible to synthesize within a single measure of value the cultural, social and economic meaning of scientific results. Furthermore, it is very difficult to find research groups that score high for any of those

measures mentioned just above, insofar as cultural, social and economic perspectives are not usually consistent with each other. It is well known that a research result can have a high value for industrial users and a low value for scientists, and that efficiency is not linked with effectiveness. On the other hand, the ability to attract resources is very likely to be correlated with the ability to achieve results in the short run. To overcome those difficulties it is necessary to consider those measures to be independent characteristics of the research activity.

Given this situation, it is easy to understand that a "successful equation", containing the reference profile of scores, does not exist. Consequently, each organization is charged with setting up its own reference criteria as part and result of its strategic planning process. We should not forget that the ultimate goal of the evaluation of research activity is not merely the allocation of financial and human resources to research groups, but the preservation and development of key competitive research competency within the University. In keeping with this goal, it is clear that both evaluation criteria and their implementation over time are part of a strategic process.

In this paper only the first one of the above measures is discussed — the measure of *RI*, that is, the acceptance of the research results by the Scientific Community.

2. Evaluation during a period of change within the Italian University system

According to the Italian Constitution, universities should promote progress in science and offer the necessary education for the development of national economic activities through both teaching and research. Governing structures and the general didactic organization are defined by each university's own statutes.

Act 168 of 1989, which established the *Ministero dell'Università e della Ricerca Scientifica e Tecnologica* (MURST), also included provisions for the establishment of a central Office for University Evaluation (art. 12, Section 4). Neither the composition nor the exact tasks of that office were defined in the Act, and it took seven years for the office to be established, thanks to a DM (ministerial decree) issued on February 2, 1996, which created the *Osservatorio Permanente per la Valutazione del Sistema Universitario Italiano* (OPV). Such a delay seems to be the consequence of a lack of awareness of the importance of this issue, which in its turn seems to result from the delay in establishing the autonomous status of the universities themselves (the so-called "autonomia universitaria"). Up to that point, universities depended very heavily on the central Ministry's directives both for budget spending and for internal organization. Nevertheless, some of the concrete aspects of autonomy have developed in most universities, albeit along different lines.

Two factors have fostered this process. The first factor is represented by a change in the mechanisms regulating the public financing of the universities, a change that was enforced by the provisions contained in the national budget legislation for the year 1994 (Act 537 of 1993, art. 5): according to these provisions, Ministry functions related to the juridical and economic treatment of teaching staff are transferred to the universities themselves. Universities are also given major responsibility in the use of public funds. From 1994 on, each would receive a lump sum, and each would be responsible for

sub-dividing it under different categories (staff, equipment, other expenses, etc.). In the previous system, university resources were represented by the teaching and non-teaching staff units and by the funds they received every year, which were strictly preallocated for each balance category. Since 1994, on the contrary, universities have a global budget they can use with great flexibility.

The second factor was the growing demand — especially in some areas of the country — for consultancy services coming from local productive communities; many universities felt the need to better equip themselves with specific provisions in their statutes and in their management functions which could adequately meet the demand coming from society outside the walls of academe.

Act 537 of 1993 also envisaged the constitution of a "Unità di Valutazione Interna" (UVI) within each university. The UVI has the task of assessing, by means of comparative cost-effectiveness analyses, the management of public funds for correctness, teaching and research activities for productivity, and the administrative process for its equitableness and propriety.

Within such a framework, the central university authorities are given reduced management power, but increased governing functions; in order to be carried out in a meaningful and effective manner, these new functions require a substantial information and evaluation base.

The evaluation system's characteristic features therefore lie in the relationship between the decentralized local level (the UVI), and the central level, represented by the OPV.

A particularly significant role in university evaluation has been played by the *Conferenza permanente dei Rettori delle Università Italiane* (CRUI). Since 1992, the CRUI has sponsored a survey aimed at gathering data from Italian universities and at stimulating reflection and dialogue on issues related to the establishment within the universities of periodical evaluation practices.

In particular, the objective of this activity has been to support the creation of an inter-university evaluation system. Within this framework the CRUI itself could provide assistance to the universities in the setting up of their internal evaluation systems, as well as in proposing and testing possible procedures and operations.

As to the UVI's organization, the CRUI suggests some general principles. In the first place, the objective of the UVIs is the global evaluation of the efficiency and effectiveness of teaching, research, and administration activities, as opposed to the evaluation of individual university employees. While carrying out the evaluation process, the UVIs are to keep in mind the context in which these activities take place, the resources available, and the operational procedures. This is the reason why the CRUI also recommends using approaches that take into account both quantitative and qualitative elements: the latter, in fact, make possible a more profound analysis of data and findings obtained through sheer numerical indicators.

The CRUI also proposes a set of indicators with a view to creating a common information base for all universities; the use of the same indicators will make it possible to compare data from different years and institutions.

The indicators proposed by the CRUI amount to 70, and refer to the four areas of university activities: the first concerns the input of resources, the second is represented

by the processes, the third by the output of products, while the fourth area refers to the context. The identification of the indicators chosen was certainly influenced by other industrialized countries' experience in this field. Along with this factor, a major role was also played by the need for an affordable overall description of each university, as well as of its operational procedures and product typology. In order to test these indicators, the CRUI also requested universities to collect all of the necessary data. The elaboration of the information thus gathered was presented in the CRUI Report, which provides a broad description of the activities of the universities, particularly in terms of those related to the area of teaching.

Finally, in order to increase the ability of the overall indicators to interpret the real state of affairs, universities are requested to adopt common basic models for data collection and elaboration. Within this particular sector, the CRUI envisions a pivotal role for itself as the designer and tester of such models.

In this context, institutional autonomy and evaluation activity are strictly connected and seem to be necessary in order to better implement the process of innovation within Italian universities. University evaluation in Italy is still at a rather initial state. The various practices in this matter have developed quite recently and somewhat chaotically. As a consequence, the results of the experiences of official and unofficial university evaluation have been imbalanced.

Therefore, it is difficult to draw conclusions concerning a model which is still within a transition phase and whose implementation is, to all effects, still underway.

The strategic importance of evaluation needs to be adequately acknowledged by institutional policies, even though such evaluation practices do not belong to local university traditions. The assessment activity has to be further developed and enhanced, and should include research, educational and management activities. This perspective requires a change in the functions of the Italian Ministry of Research, which should monitor and coordinate autonomous institutions sharing common strategic objectives and assessment systems.

3. The policy of evaluation at the UFII

Evaluation has been the current practice at the UFII since 1996 when the Statute of the University, which sets the principles and rules of the University organization, was issued. In line with the Statute, three different Commissions were established for the evaluation of management, teaching performance, and research activity. The discussion will focus on the latter issue. The Scientific Commission of the University was established in 1998. From 1996 to 1998, the Scientific Committee of the Academic Senate worked on the theme of research evaluation at the UFII with a more limited scope, i.e., the allocation of internal resources for research activities, in accordance with objective parameters. The evaluation policy of the Committee was mainly influenced by the existence of at least 14 macro-cultural areas in the University, which renders a comparative evaluation of the research performances of individuals or of research units very difficult. The Committee chose departments as units for the evaluation, as they are institutionally devoted to research, according to the Statute. Furthermore, because of the

cultural heterogeneity of the large number of departments at UFII (about 78 in 1977 and more than 100 in 1999), departments were grouped into two large sub-areas, the first including arts and humanities, law, economics, and architecture departments, the second including technical and scientific departments. It is important to note that the research evaluation of departments which has been carried out so far, and which is described in detail in the following section, is to be considered nothing more than a first approach to the problem. Evaluation is, in general, a political, rather than a technical issue. The collection of parameters and indicators of effectiveness depends strictly on the use of their outputs. A political decision was made at the UFII in the past years: the use of objective parameters for resource allocation was considered a priority and introduced by the Academic Senate, even though the evaluation process was considered largely unsatisfactory.

The research evaluation carried out according to the system described in the following section has revealed great limitations in comparing units (departments) which are heterogeneous both qualitatively and quantitatively in terms of outputs (outcomes of the research). Nevertheless, such a trial-and-error approach has achieved a key political goal, that is, it has established a habit within the scientific community of the University. It is now possible to address the problem of research evaluation in terms of a more appropriate approach which is under elaboration by the Scientific Commission of the University. This approach focuses on the comparative evaluation of the research performances of each department at the UFII with respect to its homologues in selected Italian Universities, and to a few representative international institutions.

4. Method of evaluation

The scientific production of the UFII is fairly heterogeneous as a consequence of the already-mentioned cultural heterogeneity of the departments. Because of the large variety of research activities performed, departments were first grouped in two fairly homogeneous sub-areas: one including technical-scientific departments (the SCI sub-area), and the other including arts and humanities, architecture, as well as economics and law departments (the HAEL sub-area). This division was aimed at providing more specific and appropriate assessment criteria, as well as promoting comparisons conforming to a benchmarking logic: it is not always possible to apply similar indicators to social and exact sciences, or the same indicators with identical weights. For example, articles dealing with technical-scientific subjects are often co-authored. On the other hand, this rarely happens in the social sciences. Thus, the value of the "authorship" indicator is different in the sub-areas.

The next step was to select the indicators with their appropriate values which could be used to monitor the scientific performance of each department. Tables 1 and 2 show the indicators selected for each sub-area. It should be noted that these can be grouped in two general categories: those directly related to the scientific production of individual scientists (articles, books, conference contributions, etc.), and those which indirectly relate to research activity of individuals or of the department (coordination of scientific projects, PhD and post-PhD students, participation in editorial boards, etc.).

Table 1
Technical-scientific departments (SCI sub-area)

Factor	Activity	Value
C^1	Articles in journals	Value assigned by the UFII Research Panel
C^2	Participation in Editorial Boards	2 × value of C^1
C^3	Books	40 (author, international publication) 10 (author, national publication)
C^4	Chapters in books	10 (author, international publication) 5 (author, national publication) 15 (editor, international publication) 5 (editor, national publication)
C^5	Conferences	20 (organizer, international conference) 10 (organizer, national level) 10 (participation of scientists as Chairman) 5 (participation in Advisory Board of other conferences) 1 (workshops with invited speakers)
C^6	Conference reports (published proceedings)	4 (international conference) 1 (national conference)
C^7	Coordination of scientific projects	10
C^8	Participation in national steering committees	5
C^9	Journals published by the Department	10
C^{10}	PhD programs	60
C^{11}	Patents	20 (international) 2 (national)

Furthermore, the value of each indicator "weights" the influence of the single indicator with respect to the others, and was established by a trial-and-error procedure on the basis of an a priori ranking list of a small set of departments in each sub-area. As for the values attributed to the indicator C^1, "Articles in journals", these were derived from the values produced by the Institute for Scientific Information (ISI) (Impact Factors and Citation Indexes).

The analysis of research activities was carried out in 1998, and the reference period was 1997. A Research Panel examined the data provided by each department, consisting of publications (papers accepted for publication but not yet published were considered ineligible), and any information on assessable outputs, with reference to the indicators shown in Tables 1 and 2.

An index P was used to represent the value of research activity for each department; it indicates the department's ability to perform research activities, scientific development and dissemination of results. Distinct P indexes were used for different sub-areas: P^{SCI} for the technical–scientific area, and P^{HAEL} for departments of arts and humanities, architecture, economics and law. They were obtained by adding the values of the indicators in Tables 1 and 2.

The first indicator used to describe the quality of articles published in journals (Tables 1 and 2, factor C^1) deserves a special comment. For the technical-scientific sub-area, its values were assigned on the basis of the Impact Factor (IF) values

Table 2
Arts and humanities, architecture, economics and law departments (HAEL sub-area)

Factor	Activity	Assigned value
C^1	Articles or essays published in international journals	12
C^2	Articles or essays published in national journals	10
C^3	Articles or essays published in other journals	5
C^4	Books	30 (author)
C^5	Reviews and various notes	1
C^6	Critical review of texts	12
C^7	International Conference reports (published proceedings)	5
C^8	National Conference reports (published proceedings)	4
C^9	Editor of miscellaneous books	10
C^{10}	Editor of proceedings	5
C^{11}	Chief editor of scientific reviews and collections	10
C^{12}	Participation in Editorial Boards of journals and collections	5
C^{13}	Journal published by the department	8
C^{14}	Coordination of national or international projects	8
C^{15}	Participation in scientific committees	8
C^{16}	PhD students	0.5
C^{17}	Post-PhD students	1
C^{18}	Other (software, ...)	5

reported in the *Journal of Citation Report* (*JCR*).[1] The IF is the ratio between the total number of citations in the international literature in a given journal during the previous two years and the number of articles published in that journal. It should be underlined that, although great debate exists concerning the use of IF for evaluating scientific production, this index reveals the relative importance of journals, especially those in the same or similar fields. Nevertheless, besides other general criticisms which could be leveled at it, IF has an intrinsic restriction: it does not allow the comparison of journals in different fields, because journals which publish papers in areas of more general interest receive more citations than those related to specialized areas. Consequently, areas like botany and mathematics do not generate as many articles or citations as do larger areas such as biotechnology or genetics. Likewise, in some areas, particularly in the arts and humanities, it may take a relatively long time, even several years, for an article to attract a meaningful number of citations, whereas in other areas, such as the life sciences, it is not unusual for citations to peak after only a few years. This intrinsic IF limitation is evidenced by the fact that the *JCR* groups journals by cultural areas. To overcome this restriction, the UFII Research Panel used an internal ranking criterion, which divides journals in each *JCR* ranking list into five sections,

[1] The JCR Science Edition provides citation data on nearly 5000 leading science journals, and ranks journals by the number of times they are cited in a given year, providing an international assessment of journals by scholars who have responded to the items published.

each containing the same number of journals. Journals in the first section scored 10, those in the second 8, and the following 6, 4 and 2 respectively. Thus, homogeneous and comparable values can be assigned to scientific papers in different fields. A value of 1 was given to papers that appeared in Journals not contained in the *JCR*.

For the HAEL sub-area, a value of 12 was always assigned to each paper (C^1 indicator, Table 2), independently of its IF.

Other factors generating P^{SCI} and P^{HAEL} indexes (participation of members of the department on editorial boards, the publication of scientific books or chapters of books, the organization of international and national conferences, etc) scored a value (C^n) which was multiplied by its occurrence during 1997. Moreover, in the case of participation on editorial boards, this score was also multiplied for the value given to the same Journal in the UFII Research Panel ranking list (see above).

P^{SCI} index is obtained using the following formulas:

$$P_i^{SCI} = A_i^1 + A_i^2 + A_i^3 + A_i^4 + A_i^5 + \cdots + A_i^{11}$$

$$= \sum_j C_{ji}^1 n_{ji}^1 + \sum_j C^2 n_{ji}^2 + C^3 n_i^3 + C^4 n_i^4 + C^5 n_i^5 + \cdots + C^{11} n_i^{11}$$

P^{HAEL} is obtained in a similar way, using the C^i values in the HAEL sub-area:

$$P_i^{HAEL} = A_i^1 + A_i^2 + A_i^3 + A_i^4 + A_i^5 + \cdots + A_i^{18}$$

$$= C^1 n_i^1 + C^2 n_i^2 + C^3 n_i^3 + C^4 n_i^4 + C^5 n_i^5 + \cdots + C^{18} n_i^{18}$$

where i is the department ID, j the journal ID, C the factor, n the occurrence of activity in 1997.

5. Results and funding allocation

Results are described in Figs. 2 and 3, which show productivity indexes for the SCI and HAEL sub-areas correlated to current academic staff for each department. P values were used to allocate funding to UFII departments.

The total amount of 1999 research funding expected for the UFII was previously divided into two main parts: one was assigned to departments in proportion to the number of their current academic staff, and its aim was to ensure a minimum funding level for the research activity. The other part refers to typical departmental activities. The UFII Research Panel decided to further split this latter amount into two parts, in accordance with the distinction in sub-areas: 73% was assigned to the SCI area and the remaining 27% to the HAEL area. These amounts were then allocated to single departments in proportion to their P values (see Fig. 4).

6. Further developments

There is no doubt that this method can be further developed, even though it implementation faces many problems; its weak points should be improved in order to take

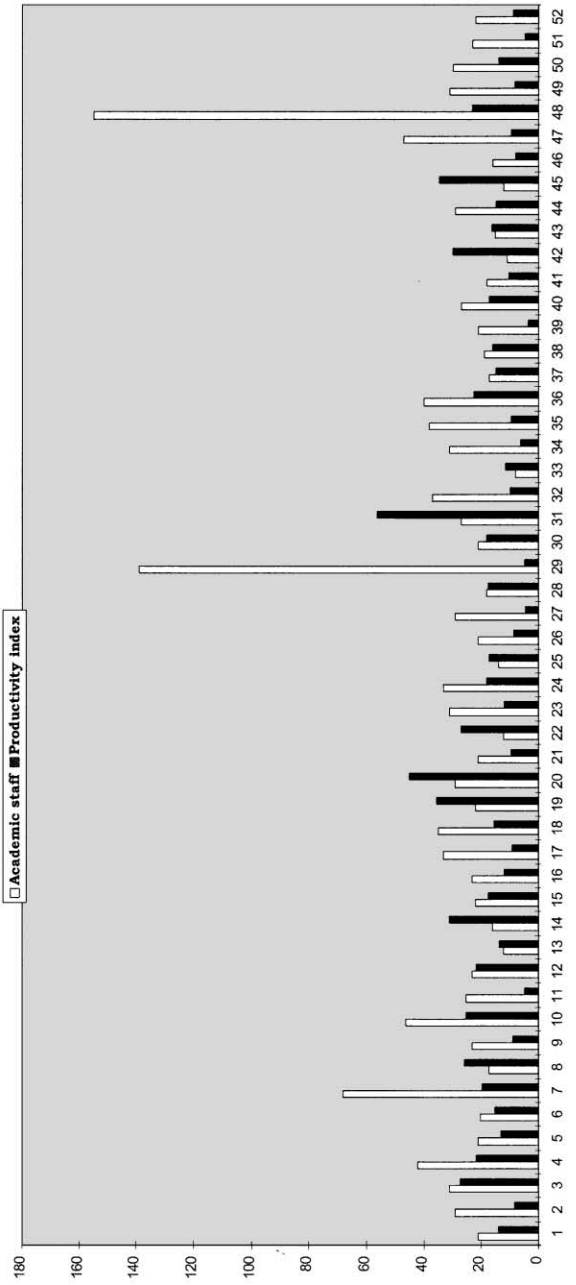
□Academic staff ■Productivity index
180
160
140
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52

Fig. 2

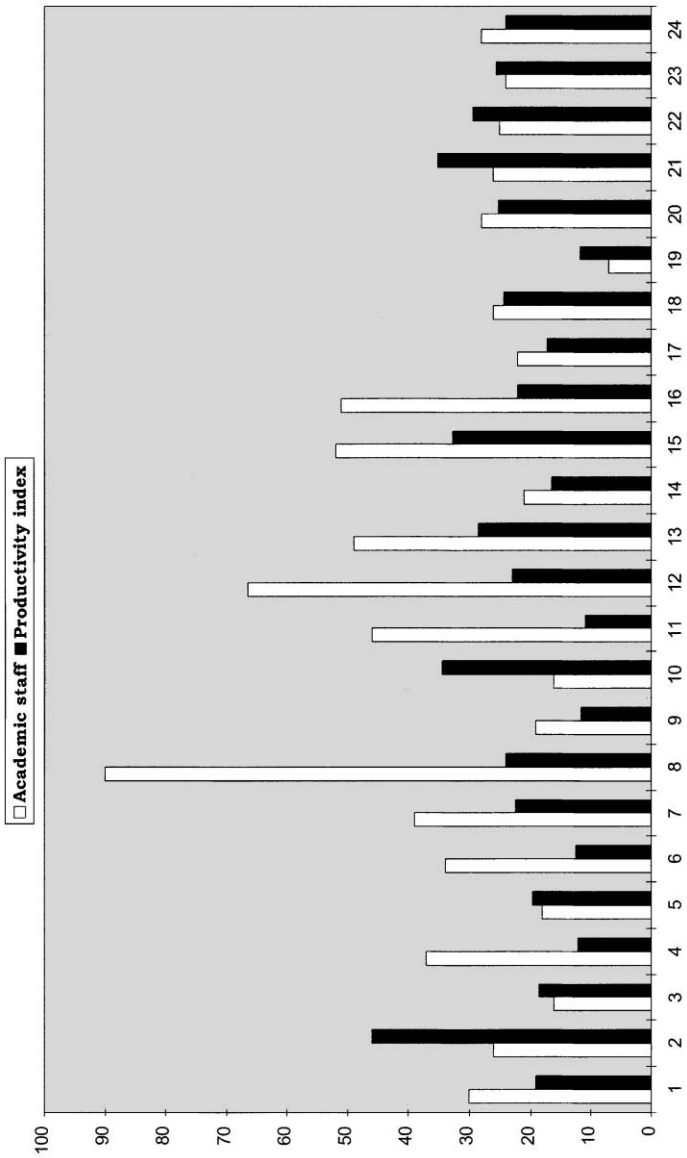
□Academic staff ■Productivity index
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Fig. 3

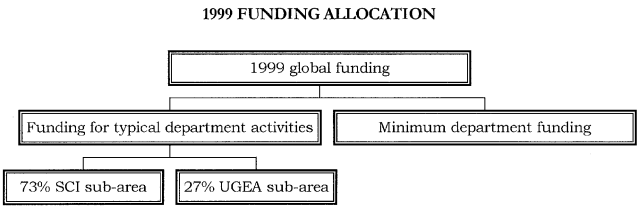

Fig. 4. 1999 funding allocation chart Total amount of 1999 research funding expected for the UFII was previously divided into 2 main parts: one was assigned to departments in proportion to the number of their current academic staff, and the other part refers to typical departmental activities. This latter amount was split into two parts: 73% was assigned to the SCI area and the remaining 27% to the HAEL area. These amounts were then allocated to single departments proportionally to thier P values.

advantage of its potential. In the present work we will simply mention some concepts which could be closely examined in future work:

- There is an important limitation in the use of merely quantitative bibliometric indicators in the evaluation process of departments: the potential number of papers produced by a department during a certain period of time is strictly connected to its specific research area. For example, the writing of synthetic-chemistry articles usually takes much longer than the writing of chemical-characterization articles. Consequently, in order to correctly apply an assessment method which uses bibliometric indicators, a process of "inner normalization" of productivity data would be useful: first, the range of published papers ($n_{max} - n_{min}$) should be stated for each research field on a national or international base. Then the ratio between the number of published papers of department X and this range should be used in the determination of C^1 factor (Table 1).
- Availability of funding and general resources (equipment, academic staff, etc.) affects results and quality of research activity. Consequently, it should be necessary to take into account all these items in the final evaluation of the whole research work.
- Finally, one of the values assigned by UFII Research Panel might be amended:
  - In the framework of university financial autonomy, patents produced by departments (Table 1, Factor C^{11}) should be given greater weight by Assessment Panels, since these could attract funding from industries. As a consequence, researchers would be stimulated to greater effort in considering the economic implications of their activity, and departments would be able to benefit by greater private investment for research. The capability of attracting funding from other corporations or institutional bodies is therefore strictly connected to the objective assessment of research activity, because this would produce the evidence of a successful investment.

The main problem with the above evaluation method is the limited attention to the different scientific practices of research groups, differences which affect the typology of research programs, the length of research projects, the cooperation among researchers

and research groups, publication rates, and the varieties of research results. Furthermore, it does not take into account the points of view of social subjects — companies, institutions — which utilize research results.

Another problem is that the evaluation method pushes people to favor research projects which can achieve results in the short run, and to discard long-term projects. Finally, it privileges large, consolidated research groups and underestimates the efforts of groups working in new scientific fields.

It is clear that is impossible to design an evaluation method that is able to overcome all these problems. The strategy set up by the UFII is to gradually develop an evaluation system which aims at consistency within the strategic plan of the University, rather than to pursue unattainable objective criteria.

Further reading

Ball, C., & Halwachi, J. (1987). Performance indicators in higher education. *Higher Education*, pp. 4–16.

Barnett, R. (1992). *Improving higher education: total quality care*. London: SRHE/Open University Press.

Bauer, M. (1993). *Performance-based funding and institutional quality assurance and development*. Paper presented at the 15th EAIR Forum, Turku, Finland. Utrecht: VSNU.

Beasley, J. E. (1990). Comparing university departments. *Omega*, (18) 171–83.

Berendsen, G. D. J. (1993). Quality assessment of research in The Netherlands, paper presented at the 15th EAIR Forum, Turku, Finland. Utrecht: VSNU.

Biggeri, L., & Scarpitti, L. (1998). *Evaluation in the Italian university system*. Paper presented at International Conference on Evaluation: Profession, Business or Politics? Rome, October 29–31.

Bignami, G. (1992). Academic standards in Italy. *Nature 358*, 364.

Burr, D. (1992). Italian university structure, *Nature 357*, 273.

Cave, M., Dodsworth R., & Thompson, D. (1995). Regulatory Reform in Higher Education in the UK: Incentives for efficiency and product quality. In: M. Bishop, J. Kay & C. Hayer (Eds.). *The regulatory challenge*. Oxford: Oxford University Press.

CHEPS. (1996). In: B. Jongbloed & J. Koelman (Eds.). *Funding mechanisms for higher education: a five-country based survey with special reference to performance based funding approaches.*

CRE. (1995a). *Guidelines for self-evaluation and preparation for audit team visits*. A cura di Frans van Vught e Don Westerheijden. Geneve: CHEPS.

CRE. (1995b). Institutional evaluation and management for quality. *The CRE institutional quality evaluation programme: background, goals and procedures.* Geneve: CRE.

CRE. (1996). *Institutional evaluation programme. guidelines for self-evaluation and preparation for review team visits*. Geneve: CHEPS.

CRE. (1997). *Institutional evaluation as a tool for change – Lévaluation institutionnelle, outil de changement*. Geneve: CRE.

CRE. (1998). *Principles and practice of strategic management in university.* Volume 1, a cura di Pierre Tabatoni and Andris Barblan. Geneve: CRE.

CRUI. (1997). Valutazione delle attività didattiche universitarie, L. Modica & E. Stefani (Eds.), *Le esperienze condotte dalla CRUI*. Rome: CRUI.

CRUI. – Universita'ca'foscari di Venezia. (1998). *Atti del convegno. Strategie ed esperienze innovative di valutazione nelle università italiane*. In: E. Stefani & M. Cavallini (Eds.). Venezia, 9 dicembre 1996. Roma: CRUI.

De Marchi, M., Poti', B., Reale, E., Rocchi, M., & Scarda, A. M. (1998). *Il sistema scientifico pubblico in Italia. Cambiamenti e tendenze*. Milano: Franco Angeli, Istituto di studi sulla ricerca e documentazione scientifica CNR.

EFMD (European Foundation for Management Development). (1997). *The dynamic model for quality development. Guidance notes on the equis quality criteria*. EFMD.

Gaetani, G. F., & Ferraris, A. M. (1991). Academic promotion in Italy. *Nature*, *353*, 10.

Garbisa, S., & Calzà, L. (1995). *Il peso della qualità accademica Cleup*.

Harvey, L., & Green, D. (1988). Defining quality. In: *Assessment and evaluation in higher education*, 1993, 1, 9–34. n. 42, pp. 54–71.

Kells, H. R. (1990). The inadequacy of performance indicators for higher education: the need for a more comprehensive and development construct. *Higher Education Management*, n. (3), p. 258–270.

Kells, H. R. (1992). *Self regulation in higher education. A multi-national perspective on collaborative systems of quality assurance control*. London: Jessica Kingsley Publisher.

Nova Scotia Council on Higher Education, (1997). Discussion paper on the development of a new funding formula for Nova Scotia's universities. Halifax: Nova Scotia Council on Higher Education.

OECD (1994). *Evaluation and the decision making process in higher education*: *French*, *German and Spanish Experiences.* Paris Cedex: OECD.

OECD (1995). *Performance standards in education. In search of quality.* Paris Cedex: OECD.

Persico, F. (1992). Academic standards in Italy. *Nature*, *358*, 364.

Politecnico di torino, CPD. (1996). In: M. Gola (Ed.). *The trend and the methods of quality assessment at the Politecnico of Torino.* Torino: Politecnico di Torino.

Richter, R. (1994). Improving teaching quality in German higher education. *Higher Education Management, 2*.

Roberts, L. (1991). The rush to publish. *Science*, *251*, 260–263.

Romagnoli, M. (1994). L'importanza di essere citati. *Le Scienze*, *312*, 08.

Strata, P. (1995). Citation analysis. *Nature*, *375*, 624.

Taubes, G. (1993). Measures for measure in science. *Science*, *260*, 884–886.

Universita' ca' Foscari di Venezia. (1996). Strategie ed esperienze di valutazione nelle Università italiane. Seminario di studio. *La valutazione CRE dell'Università Ca' Foscari*. Venezia, 9 dicembre 1996. Venezia: Unità Studi, Programmazione e Progetti Speciali.

Università Commerciale Luigi Bocconi. (1998). *Relazione annuale sull'attività di ricerca.* Milano: Università Commerciale Luigi Bocconi, 1997.

Zucco, F. (1994). Valutazioni a peso. *Sistema Ricerca, 35*, 33–35.