

## The search for self-evaluation of knowledge

*Massimo Florio, Paola Galimberti, Silvia Salini*

This special issue of the *Rassegna Italiana di Valutazione* draws from selected papers presented at a workshop on the evaluation of research held at the University of Milan (November 29, 2012). Italian academics are currently experiencing the second nationwide research assessment, a monumental task involving bibliometric analysis and peer review of more than 184,000 publications, i.e. typically three publications per scientist, self-selected as the best research of each scholar. At the same time, universities have implemented an array of self-evaluations, often through the Nucleo di Valutazione's board of experts (appointed by the Rectors) or other bodies. The objectives, scope, and styles of the national assessment and the internal evaluation differ. The workshop in Milan, organised by the Nucleo di Valutazione of University of Milan and the researchers of the UNIMIVAL team (a project self-managed by ten young scientists), focused on the contributions of experts from other Italian universities and from abroad about the specificity of internal evaluation and its relationship with national assessments. The debate was lively, ignited by a set of distinguished speakers, including Remy Barrè (CNAM, former director of the OST), Britt Holbrook (University of North Texas), Luciano Floridi (University of Hertfordshire), Henk Moed (Elsevier), Victoria Ley (Agencia Estatal Antidopaje, former Director of ANEP), Birgit Schmidt (IFQ), Sergio Benedetto (ANVUR), Silverio Bolognani (Università degli Studi di Padova), Dario Braga (Università degli Studi di Bologna), Bruno Dente (Politecnico di Milano), Aldo Geuna (Università degli Studi di Torino), Gilles Pècout (Ecole Normale Supérieure Paris), Paolo Rossi (CUN), and Alberto Silvani (Università degli Studi di Milano). We are extremely grateful to Alberto Silvani, editor of this journal and an expert on these topics, for inviting us as guest editors for this special issue.

We are unable to publish here all the contributions, nevertheless we hope that the readers of RIV will find these papers of interest. We have also included papers

**Rassegna italiana di valutazione, a. XV, n. 52, 2012**

that were specifically commissioned in order to complement the range of topics covered by this special issue.

A striking aspect of this material is that it shows that the perceived existence of an accepted paradigm for evaluation of research is misguided. The more we know about the endeavour of assessing the quality of knowledge in contemporary large research institutions, the more we understand how daunting the task is, how prudent, experimental, and astute evaluation strategies should be. We briefly summarize below some aspects of the included papers (selected with the advice of anonymous reviewers) that may convince the reader of our tenet that self-evaluation of knowledge in large multidisciplinary universities is both important and challenging and that it has a specific role to play with respect to national assessments.

The paper by **Rémi Barrè**, drawing from experience in France, opens the special issue. It offers a systematic analysis of the different levels of the game, the different actors, and the different styles. He suggests that the roles of ministries of higher education and universities are indeed different in terms of timing, the resource allocation process, and policy implications. In this context Barrè suggests that self-evaluation of universities is often related to the scope of quality-assurance mechanisms, somewhat related to the principles of 'new public management'. This complements the coordination and accountability functions of national bodies. The development of performance-based funding in universities is also perhaps related to the new public management trends. Having acknowledged this, Barrè, however, notes wide differences between the experiences of Italy, France, and the United Kingdom, differences that are perhaps not surprising given the very different evolution of government and public education in these countries.

**Paolo Rossi** also discusses the analysis of different evaluation layers. He proposes a useful taxonomy that offers an analytical contribution to the discussion. First, he suggests that each evaluation should be closely related to its objectives – allocation of funds, pay, recruitment, promotions – are some of them, and all need different tools. Second, the subject of evaluation may be individuals, departments, universities, or research bodies external to universities, and there is no one recipe for any and all of these subjects, but different ones for each. Third, objects of evaluation range from books to articles, patents to other output of research, and each needs its own specific assessment framework. Fourth, methodological differences include peer review, bibliometrics, and a combination of these two approaches. The hard sciences, life sciences, humanities, and social research have evolved on different paths, and these differences must be duly considered. But at a deeper level, we still need to think seriously about the social role of science.

**Britt Holbrook**, a philosopher, reminds us that a purely managerial perspective is perhaps not sufficient. He emphasises that universities must be accountable to themselves and to society. The old issue of the tension between autonomy and accountability is not solved by the quest for excellence. After all, 'excellence' is a value judgment from within the academic community, and in the pragmatic and dynamic environment of the United States, according to this paper, the notion that research can always evaluate itself is questionable. What are the benefits to soci-

ety of having so many intellectual energies devoted to knowledge, if knowledge does not accept self-assessment from a social perspective? The typical response from scientists is that at a certain point something good, something totally unexpected ('serendipitous') will happen and ex-post justify academic research and the resources invested in it. This perspective is considered too weak by Britt Holbrook, who suggests that social accountability should be internalised by researchers.

**Ley and Rey-Campos** discuss the experience of Spain and use it as an example of the reasonable conjecture that, in spite of many imperfections, evaluation of research can positively contribute to knowledge; this, in turn, can support a sustainable growth path. In fact, economists and other social scientists would confirm that there is a clear correlation between the quality and quantity of research and productivity growth. This, however, cannot be an argument to support all expenditures in academic research, and governments must be held accountable for funding policies. In turn, deans of faculties, principals of universities, and directors of research bodies have similar concerns and need to manage, sometimes to micro-manage, their organisations that, after all, produce knowledge; this should be measured in some way to make decisions. The controversial aspect of the story begins when inappropriate procedures are implemented. The authors are optimistic about the experience in Spain and highlight some of its positive features. They also point to some misconceptions that may lead to a misunderstanding of the proper scope of evaluation.

**Galimberti** focuses on some critical issues in the current and past experience in Italy. She discusses the difference between the national and the decentralized evaluation in this country, which is a latecomer in the evaluation arena. Italian universities are under financial stress because of the general budgetary situation, and they need to be realistic about what they can achieve with their own resources. Moreover, the national assessment, as designed by the National Agency for the Evaluation of Research<sup>1</sup>, cannot examine the necessary details, sometimes subtle and variable, that only an analysis from within can discover. However, without a proper benchmark, this analysis risks being ineffective, and Galimberti sees some merit in a bottom-up approach (as in the UK project Snowball) – one that can be based on information-sharing agreements and shared indicators between groups of universities. This path would involve the building of common databases and metrics as a process independent (in part) from the national assessment, with all its intrinsic limitations.

In this vein, **Biolcati-Rinaldi, Ferrara, Pinotti, and Salini** recount their experience with the UNIMIVAL bibliometric pilot experiment. They restate the hard fact that bibliometrics is not the same concept as research evaluation. After having noticed that quantitative analysis of science, e.g. through citations, is now a worldwide trend, they support the evidence that there is no general consensus of a simple set of indicators valid for any discipline or context. They are critical of this search for a common algorithm for 'counting' what basically are different objects

1. [www.anvur.org](http://www.anvur.org).

and propose a different approach used at the University of Milan by the UNIMI-VAL group, in agreement with the guidelines of the Evaluation Board of the same university. This is a prototype system for comparative bibliometric analysis (ABC) aimed at comparing different departments, branches, groups, research areas, and single researchers with other benchmark units. These can be selected as actual departments in other universities, specific ‘competing’ research teams, or other entities designed according to specific needs (e.g. discipline or mission). The core concept of ABC is that each group self-selects its own comparators, metrics, and sources of data in a flexible way.

While this approach has been tried only in a limited number of departments at UNIMI, in the last three years a key feature of the Nucleo di Valutazione has been to ask each head of department (and coordinator of each undergraduate course) to start his or her own self-assessment by the identification of a small number of external benchmarks (in other Italian and European Universities) and to build on this comparative approach of similar paths of knowledge creation<sup>2</sup>.

The paper by **Breno and Paganelli** puts the experience of Milan in a wider perspective. Drawing from a survey managed by the CRUI (a body established by the Rectors of Italian universities), it reveals that independently from the national assessment, most Italian universities have started their own evaluation approach. The survey is highly informative, including details about the usage of specific databases for bibliometric analysis, the use of peer review, and the coverage, scope, and actors of the assessments.

We conclude by our own hope that universities will develop evaluation strategies based on experience in Italy and abroad. We believe that this experience is still mixed and unstable. Germany and the USA do not have a systematic national research assessment. The UK, Spain, and Italy have adopted a combination of central and decentralized evaluation. France is re-thinking its own approach. While no one-fits-all paradigm has emerged, there is a wealth of experience, and this points to a role for self-evaluation of research communities, provided that they look ‘outside’, to their peers and to society, to understand to what extent they are actually contributing to knowledge.

2. [http://www.unimi.it/cataloghi/nucleo\\_valutazione/Relazione\\_annuale\\_2011\\_NUV.pdf](http://www.unimi.it/cataloghi/nucleo_valutazione/Relazione_annuale_2011_NUV.pdf).