

# Designing indicators for policy decisions: challenges, tensions and good practices: introduction to a special issue

**Benedetto Lepori, Emanuela Reale and Robert Tijssen**

This special issue collects six papers (out of 80 submissions) from the Science and Indicators Conference organized by the European Network of Indicators Designers (ENID) in March 2010 in Paris, France. They cover quite different aspects of design of indicators for policy decisions, ranging from internationalization policies in research (Edler and Flanagan, 2011), to indicators to support the evaluation processes of funding agencies (Potì and Cerulli, 2011; Neufeld and Von Ins, 2011), to the analysis of publication outputs of higher education institutions (Reale, De Filippo, Gomez, Lepori, Potì, Primeri, Probst and Sanz Casado, 2011) and, finally, to two papers dealing with the highly contested issue of indicators to measure research output in social sciences and humanities (De Jong, Van Arensbergen, Daemen, Van der Meulen and Van den Besselaar, 2011; Probst, Lepori, De Filippo and Ingenhoff, 2011).

**T**HIS SPECIAL ISSUE collects six papers that were presented at the Science and Indicators Conference organized by the European Network of Indicators Designers (ENID) in March 2010 in Paris, France. As such, they represent the final outcome of a selection process which started with about 80 abstract submissions for the conference itself and then went through, first, the selection process for the conference and, second, through an open call for papers for this special issue, and an external peer review process. We gratefully acknowledge the many referees who kindly contributed — often at short notice — their insightful, critical and constructive inputs to help us and the authors to produce the high-quality papers presented in this issue.

Expectedly from the organization of this process, these papers cover quite different topics in the field

of design of indicators for political decisions, ranging from internationalization policies in research (Edler and Flanagan, 2011), to indicators to support the evaluation processes of funding agencies (Potì and Cerulli, 2011; Neufeld and Von Ins, 2011), to the analysis of publication outputs of higher education institutions (Reale, De Filippo, Gomez, Lepori, Potì, Primeri, Probst and Sanz Casado, 2011) and, finally, to two papers dealing with the highly contested issue of indicators to measure research output in social sciences and humanities (De Jong, Van Arensbergen, Daemen, Van der Meulen and Van den Besselaar, 2011; Probst, Lepori, De Filippo and Ingenhoff, 2011).

Underlying this topical diversity are a few general topics and common issues — with different emphases — related to the epistemological and sociological nature of indicators, as well as to the understanding of our profession and of the scholarly and political role of designers and producers of S&T indicators. These common threads are:

- The need for a close relationship between the political and evaluation processes from one side, the design of performance indicators from the other side, implying that vested interests and power structures among actors and stakeholders are

---

Benedetto Lepori is at the Centre for Organizational Research, University of Lugano, via Lambertenghi 10a, CH-6904 Lugano, Switzerland; Email: [blepori@usi.ch](mailto:blepori@usi.ch). Emanuela Reale is at CERIS CNR, Via dei Taurini 19, 00185 Rome, Italy; Email: [e.reale@ceris.cnr.it](mailto:e.reale@ceris.cnr.it). Robert Tijssen is at the Centre for Science and Technology Studies (CWTS), University of Leiden, Postbus 905, 2300 AX Leiden, The Netherlands; Email: [tijssen@cwts.leidenuniv.nl](mailto:tijssen@cwts.leidenuniv.nl).

constitutive of the design process and that indicators designers cannot stay out of the political game. The papers in this special issue take in this respect quite different positions, from designing indicators as derived from policy needs (Edler and Flanagan), to proposing off-the-shelf technical recipes (Potì and Cerulli).

- The shift from the development of general-purpose indicators, allowing broad comparisons between fields, institutions and countries, towards customized indicators closely related to the specificities of their usage context, which raises difficult questions concerning generality and transferability of context-dependent indicators. These issues are, for example, a central concern in the paper of Reale *et al* on university databases, which reflects the tension between local usages and data sources and requirements for broader comparability. Also, the Neufeld and Von Ins paper nicely displays how general-purpose bibliometric indicators can be adapted to fit into local contexts of peer review, by introducing context variables (e.g. the presence of self-selection of applicants based on the same indicators) and by smart combination of different indicators to reflect the multidimensional nature of peer review.
- The conceptual shift from a 'linear' process where indicators proceed from design towards (standardized) production and interpretation towards a more interactive process, where indicators are developed, contextualized and interpreted in an iterative and open way, with the operational risk that there is never a 'final' set of indicators which can be produced regularly (with all the implications for robustness, stability and comparability across time, as well as production and transaction costs). Especially the two papers on social sciences and humanities deal strongly with this process dimension and its potential impact on the produced indicators (De Jong *et al*, Probst *et al*), whereas other papers seem to stick to a more linear conception (e.g. Potì and Cerulli).
- A broadening of data sources from what would traditionally be considered as the main sources for indicators production — including statistical databases and international databases like patents and publications — towards a much more eclectic and diverse set of data sources, raising thus issues of data validation, comparability and reproducibility of the produced indicators, which strongly impact also on the methodological effort required for designing multi-source, multi-context indicators. This is a concern for the use of university databases for the analysis of publications (Reale *et al*), but also for all kind of indicators produced from the ground in social sciences and humanities (De Jong *et al*, Probst *et al*).

Of course, these are well-known conceptual and methodological issues both in the practice of policy evaluation and of the production and design of S&T

indicators (Lepori *et al*, 2008); yet, the concrete examples introduced in these papers might help our community to go beyond the general discussion towards the development of good practices on how to deal with these issues. We thus suggest that, besides their specific topics, the reader might consider the contribution of these papers to these broader issues.

## The papers in a nutshell

One of the papers in the special issue deals with the design and the use of indicators at the policy level; two draw attention to their value to funding agencies for improving the application selection process, one to producing indicators at the institutional level of higher education institutions and, finally, two papers focus on measuring research output for specific fields and research units. These contributions thus cover the main relevant institutional levels where indicators' design and use might take place.

**Edler and Flanagan's** paper provides a broad framework to design indicators to support internationalization policies. Their key argument is thus that the identification of indicators needs has to be built on an analytical model of the policy process. On this basis, they systematically identify the needs at each stage of the process; namely, analyzing the *status quo* of internationalization policies, setting targets and making strategic choices, understanding the international opportunity environment and finally monitoring developments and evaluating policies. They explore examples of existing indicators and discuss the extent to which meeting each of these needs is feasible. They conclude that few dimensions of internationalization are at present well served by existing indicators and that policy-makers will have to work closely with indicator designers to determine innovative approaches that can tackle this gap.

The two papers on funding agencies address two very different sides of the proposal selection process; namely, how to select the best applicant for academic grants and how to select companies where the impact of public subsidies on private R&D is larger. The paper by **Potì and Cerulli** develops a proposal on how to build on econometric evaluations of additionality of public R&D subsidies to public companies in order to improve the selection process of beneficiaries. They point to the fact that, if additionality is one of the main goals of public R&D subsidies, commonly used indicators to evaluate the financial viability of private companies should be complemented by indicators measuring this potential; however, forecasting additionality is a complex task which cannot be addressed by simple financial indicators. To overcome this problem, they propose a procedure, led by an econometric/counterfactual approach, able to get a synthetic measure of future (*ex-post*) additionality-potential of firms demanding an RDI subsidy.

**Neufeld and Von Ins'** paper continues a long-standing debate concerning the selection criteria of academic grants and the functioning of peer review processes (Van den Besselaar and Leydesdorff, 2009), by investigating the correlation between funding decisions and the applicant's past bibliometric performance, using the example of applicants of the German Research Foundation *Emmy Noether Program*. Their findings point, first, to the importance of self-selection effects among applicants; as high-level international publications are explicitly indicated as a requirement, researchers with low numbers of publications are virtually absent from the pool of applicants and this weakens the predictive power of bibliometric indicators. At the same time, they show evidence that, while no single bibliometric indicator provides good predictions of funding decisions, combinations of several indicators result in significantly better predictions; this points to the complex nature of review processes, where different criteria are considered and combined to come to funding decisions.

The next paper, by **Reale and co-authors**, is the result of a comparative international evaluation of the usability of institutional publication databases for the evaluation of the publication output of universities. Their results support the idea that institutional databases are social constructs, displaying a self-representation of the research performance of universities, which is strongly affected by the interests of the different communities and the intended usages. The study also displays a number of general problems in using these databases for international comparisons, including differences in coverage, definitions, as well as in the structure of the data themselves. Accordingly, institutional databases of university publications are still a 'work in progress', which need further development and standardization, but which have considerable potential for the future development of indicators; matching information from these databases with international databases (e.g. to overcome the problem of author affiliations) is particularly promising perspective.

The two final papers centre on one the most debated topics in S&T indicators: how to measure research outputs in social sciences and humanities (Nederhof, 2006). Both papers rely on a broad and multidimensional conceptualization of research output, and on the notion that considering the context and working together with actors in the field are central requirements. The paper by **De Jong and co-authors** presents key results from the Dutch 'research in context' project. Their fundamental thesis is that proper evaluation of research needs to account for the specific nature and context of research fields and groups. Accordingly, in the case of two

strongly practice-oriented fields (law and architecture), they analyse the nature of the field and the characteristics of the knowledge dynamics, as well as of the most relevant involved audiences, and translate these findings in directions to develop indicators for evaluation purposes.

Finally, **Probst and colleagues** discuss the findings of a large-scale empirical study on research output in Swiss communication sciences, where an attempt was made to study the activity profile of research groups (Larédo and Mustar, 2000) and to operationalize it through quantitative indicators. The paper points particularly to two findings:

- The fact that constructing profiles is essentially a social process where consensus on measures has to be negotiated between actors (the research community in the field in this case) and where indicator designers play a central role as social mediators.
- The close relationship between technical aspects — the choice of indicators, normalization, construction of benchmarks — to one side, value choices and actor's interest on the other side and thus the need of a process design where close interaction is established between these two dimensions.

## References

- De Jong, S, P van Arensbergen, F Daemen, B van der Meulen and P van den Besselaar 2011. New uses of the institutional databases of universities: indicators of research activities. *Research Evaluation*, **20**(1).
- Edler, J and K Flanagan 2011. Indicator needs for the internalisation of science policies. *Research Evaluation*, **20**(1).
- Larédo, P and P Mustar 2000. Laboratory activity profiles: an exploratory approach. *Scientometrics*, **47**(3), 515–539.
- Lepori, B, R Barré and G Filliatreau 2008. New perspectives and challenges for the design and production of S&T indicators. *Research Evaluation*, **17**, March, 33–44.
- Nederhof, A J 2006. Bibliometric monitoring of research performance in the social sciences and the humanities: a review. *Scientometrics*, **66**(1), 81–100.
- Neufeld, J and M von Ins 2011. Informed peer review and uninformed bibliometrics? *Research Evaluation*, **20**(1).
- Poti, B and G Cerulli 2011. Evaluation of firm R&D and innovation support: new indicators and the *ex-ante* prediction of *ex-post* additionality-potential. *Research Evaluation*, **20**(1).
- Probst, C, B Lepori, D A de Filippo and D Ingenhoff 2011. Profiles and beyond: constructing consensus on measuring research output in communication sciences. *Research Evaluation*, **20**(1).
- Reale, E, D de Filippo, I Gomez, B Lepori, B Poti, E Primeri, C Probst and E Sanz Casado 2011. New uses of the institutional databases of universities: indicators of research activities. *Research Evaluation*, **20**(1).
- Van den Besselaar, P and L Leydesdorff 2009. Past performance, peer review and project selection: a case study in the social and behavioral sciences. *Research Evaluation*, **18**(4), October, 273–288.