

Towards socially robust S&T indicators: indicators as debatable devices, enabling collective learning

Rémi Barré

Since indicators play an increasing role for decisions regarding research policy and management, the issues of their epistemological status and conditions of robustness are important ones. The purpose of this article is to show that the building and interpretation of S&T indicators is a normative process which is thus debatable, such debates having to take the form of collective learning processes through what we call the ‘indicators assessment forum’. One can identify several international working groups and conferences dedicated to indicators discussion, which raises the issue of the extent to which they fulfill the mission of being assessment forums for the indicators. In conclusion, the article first calls for evaluating their capacity for fostering such collective learning and, second, suggests that being debatable is not a limitation, but the essence of the contribution of indicators to decision-making processes.

FEW HAVE DONE MORE than Antony van Raan for the development of S&T indicators: in particular, his contributions — along those of the ‘Leiden group’ — to the scientific and professional status of bibliometrics, have led to their widespread use in evaluative and decision-making processes at large (Van Raan, 2004).

In addition to the advances in bibliometrics, the general adoption of the ‘new public management’ framework along with advances in communication technologies have generalised the use of S&T indicators of all kinds: input, output, interaction, impact, positioning ... S&T indicators now play a central role in monitoring, evaluation and decisions regarding research, higher education and innovation policies and activities at national, institutional and individual researcher levels. In this sense, we witness the triumph of S&T indicators — not only of bibliometric indicators — in the context of the encompassing need for assessments and the striving for evidence-based policies (Lepori *et al.*, 2008).

But, paradoxically, we also witness a continuing and possibly deepening concern about their significance and suitability for assessments and for management as well as policy decisions, as shown for example by the bitter debates on the use of indicators by the AERES¹ in France, on the new RAE in the UK or on their use regarding university rankings. Beyond these debates, surfaces a critique of an allegedly reductionist practice of assessment if research and higher education performance are limited to elements that can be quantified.

Two points can be made in this respect:

- Today, the ‘research enterprise’ governance at large is based on S&T indicators at least as much as on the peer review system;² and
- For a variety of reasons the concern about the relevance and legitimacy of S&T indicators in their relation to scientific, management and policy decisions resurfaces regularly.

In other words, there are possible shortcomings in our S&T indicators handling, and it appears that any misconception, inappropriate use or inadequate interpretation of indicators — however subtle — can have disproportionate and irreversible impacts, not only on particular groups or institutions, but on the

Rémi Barré is Professor at Conservatoire National des Arts-et-Métiers (CNAM), 292, rue Saint-Martin 75003-Paris, France; Email: remi.barre@cnam.fr.

cognitive trajectories of the whole research enterprise if such shortcomings introduce systematic biases. The stakes of the issue of the criticism of S&T indicators are therefore high indeed: if we seriously believe that indicators have a substantial role to play for decisions, we need to assess the articulation of indicators to political and decision-making processes.

We will show that the building and interpretation of S&T indicators is a normative process which is thus debatable (Section 1); such debates having to take the form of collective learning processes to ensure the 'social robustness' of the indicators through what we call the 'indicators assessment forum' (Section 2). We then acknowledge the existence of several working groups and conferences dedicated to indicators discussion, raising the issue of the extent to which they fulfill the mission of being assessment forums for the indicators (Section 3). We conclude by calling for an evaluation of their capacity for fostering collective learning regarding S&T indicators and by suggesting that being debatable is not a limitation, but the essence of the contribution of indicators to the decision-making processes.

1 S&T indicators and their interpretation are normative therefore debatable

1.1 The quantification process: an act of a 'political' nature

Building indicators is about quantifying, that is, presenting in a numerical form what was previously expressed by words. This supposes, first, the set up of:

conventions of quantification ... implying comparisons, negotiations, compromises, codified and replicable procedures, and computations leading to the production of a numerical value ... quantification is made of two moments: conventions set up and measurement. (Desrosières, 2008: 10–14, our translation)

Thus quantification is the set of conventions and measurement operations which creates an expression of the world; it constitutes a 'translation' of the world into a specific language, this translation being based, even implicitly, on a specific conceptual model of how the element under concern 'works'.

In other words the building of indicators:

shapes representations, categories, classifications - an order of what can be seen and expressed, of what can be thought and considered feasible. (Bruno, 2008: 170, our translation)

Indicators 'shape identities', conveying a vision of the world providing norms for individual and collective behaviours, defining performance criteria and what can be compared.

This means that quantifying is translating the reality we want to express in the form of numbers, according to the core hypothesis and vision of the accepted conceptual model. In this respect an indicator is the homologue of a 'fact', a notion which is problematic since:

it doesn't take into account the enormous work of formatting, shaping, classifying, deducing, to provide data with a meaning they never had by themselves. (Latour, 1999: 167, our translation)

After the conventions are codified and routinised, the products of these — the S&T indicators in particular — tend to become 'reality'. It is so until, for some reason, sooner or later, what had become an un-discussed 'black box' is re-opened, at the occasion of a controversy leading to a change in perspective. But meanwhile, the confidence in numbers and in the professional reputation of experts will secure the non-debatable of the indicators.

The point therefore is that the quantification process is of a 'political' nature — in the sense that it is based in a specific shaping and understanding of the way things work — which is not recognised as such, so that indicators are considered, in good faith, as a technical and scientific reality.

1.2 The interpretation process: a value-laden exercise

S&T indicators can be defined as expressions of our knowledge of the system of research and innovation or parts of it, aiming at policy, strategic or management decisions;³ such indicators take the form of the measurement of some parameters considered relevant, and the knowledge they express refers, even implicitly, to science policy studies. Indicators are used as a mediation between science and decision and thus can be considered as a form of scientific expertise.

It follows that a critique of S&T indicators in their relationship to decision-making in the research arena can be expressed along the same arguments as the critique of scientific expertise in its relationship to decisions in general.

Scientific expertise is knowledge provided by scientists which informs a decision-making process, this knowledge being the response of those scientists to questions raised by policy-makers. But in such a policy- or decision-making context: 'usually, scientists do not have ... answers that can be considered as the direct expression of his/her scientific knowledge (Roqueplo, 1997: 18, our translation); the scientific expert:

will answer, on the basis of his/her knowledge, what he/she thinks, what his/her conviction is: he/she will express an opinion ... or conviction of a knowledgeable person ... [S]cientific expertise necessarily transgresses the limits of the

scientific knowledge upon which it is based.
(Roqueplo, 1997: 19–20)

This points to the fact that any interpretation of indicators in terms of policy or management decision is bound to include, one way or another, the personal vision of the expert involved in the process.⁴

1.3 S&T indicators and their interpretation as expressions of representations

These two lines of analysis tell us that policy advice, strategic or management decisions based on S&T indicators are fundamentally dependent on normative statements at two levels:

- The quantification process — choice of parameters and metrics — which define the indicators, is based on a conceptual model of how science works — among several possible models;
- The interpretation of the indicators to give them a meaning for policy matters, strategic or management decisions bears elements of norms and opinion of the analysts or experts involved.

There is here a self-reinforcing mechanism by which the science, the numbers and the experts are aligned to produce advice in accordance with what appears to be normative processes — but which have become considered as scientific knowledge, established expertise and non-disputable rationality.

We suggest therefore the following: S&T indicators are measurements which express a representation, model or theory on the functioning of research. In this view, indicators are neither truth nor arbitrary construct, but the coherent, explicit and rational construct based on a representation or conceptual model of how research functions. In addition, the interpretation of these indicators in terms of a decision is necessarily a mix of opinion and analysis.

It follows that, beyond questions of data and computational accuracy, the definition and the interpretation of any indicator is eminently debatable. This understanding of the status of indicators can be considered as a critique since it reveals what could be considered as a drastic limitation of their usefulness. In the next section we suggest the conditions to overcome this situation.

2 S&T indicators production and use: the function of the 'indicators assessment forum'

In practical terms, what are the implications for S&T indicators production and use? We show in this section it means the necessity to produce 'socially robust' indicators and we show then that they should be built in the context of collective learning through what we call an 'indicators assessment forum'.

2.1 The need for 'socially robust' and contextualised indicators

As we have seen, S&T indicators embody our knowledge about the system of research and this knowledge 'speaks' to society in the sense that it says things regarding S&T policy and strategic or management decisions.⁵ The debatability of indicators can be seen as the necessity for society to 'speak back' to science in order to produce 'contextualised knowledge', that is, knowledge which not only is reliable, but also where validation has involved an extended group of actors. Such knowledge is then 'socially robust', having experienced validation criteria more complete than those set up by scientific peers (Gibbons, 1999).

We suggest that the response to the criticisms on S&T indicators expressed above is to engage in building them as socially robust and contextualised indicators, that is, building them with the relevant stakeholders. This means in a space where the necessary interactive processes can take place — this space having thus the characteristics of an agora or a forum (Barré, 2004).

We need now to be more specific about the functions of this space, which we call the indicators assessment forum.

2.2 The need for collective learning process

From what we have seen, the debates fit for building such socially robust indicators must deal with the representations and mental models the participants have about research; thus the basic function of the indicators assessment forum is to allow for interactions involving explicitly the subjectivity of its participants, that is, for inter-subjective interactions.

To address this requirement, it is relevant to mobilise the notion of learning processes. Codified knowledge is expressed and transmitted in a formal language. For example, the language of numbers in the case of indicators; tacit knowledge, on its side, concerns mental models, representations through which individuals perceive and interpret reality by organising their experiences (Polanyi, 1966). Let us consider the following process, in three steps:

1. The building — conception and computation — of the indicators, which amounts to translating a representation (tacit knowledge) into quantified parameters (codified knowledge).
2. The debates about the indicators, which imply various combinations, comparisons and operations on numbers (codified knowledge treatments), which can be understood by all the actors involved since numbers are a common explicit (universal) language, but which leads to debates about the interpretations and implications of the indicators since the underlying representations of the participants differ.

3. Finally, the insights gotten by each of the participants in the previous step are re-interpreted and translated in terms of their own representation (tacit knowledge), which can be more or less modified by the process.

This three-step process, touching upon the representations, consists in the translation of tacit to codified knowledge and back to tacit through the intermediary step of interactions based on codified knowledge. It is also called a collective learning process and results in a broadening and/or questioning of the representations the actors have (Nonaka, 1991).

It appears that such learning processes have the functional characteristics required for the indicators assessment forum, since they allows for inter-subjective interactions. The indicators assessment forum can thus be conceived as such a tacit-codified knowledge transformation process taking place by interactions among the actors, interactions consisting in expressing and criticising interpretations of indicators.⁶

Other characteristics of the forum are that it must have a high degree of openness to a variety of participants and be accountable to external actors (Barré, 2001).

In summary, these features of collective learning capacity, openness and accountability characterise what we call the indicators assessment forum, fit for the building of socially robust S&T indicators.

3 Indicators assessment forums in practice

All along the rise of the indicators production and use, since the 1970s, several international working groups and conferences dedicated to indicators discussion have been developing. We will identify some of the most important and raise the question of their characteristics.

3.1 Working groups and scientific conferences where indicators are debated

(a) Working groups in international contexts

- OECD-NESTI (National Experts on Science and Technology Indicators), which develops and

These features of collective learning capacity, openness and accountability characterise what we call the indicators assessment forum, fit for the building of socially robust science and technology indicators

discusses statistical methodologies for building indicators of science and technology (R&D, innovation, patents, technological balance of payments, human resources for science and technology, etc.); it acts also as a clearinghouse through which member countries can exchange information and experience on methods of collecting, compiling, analysing and interpreting data and indicators.

- UNESCO Institute for statistics (UIS) and the Ibero-American Network on Science and Technology Indicators (RICYT), which provide opportunities for representatives of member states to discuss the validation of indicators used in the publications of these organisations.
- At the EU level, EUROSTAT has regular working groups not only with the representatives of the national statistical offices, but also with indicators providers and users; focussed on S&T indicators, DG Research supports activities involving technical and interpretative tasks, completed by interactions with member states representatives, but also stakeholders, for example through IPTS⁷ or its science, technology and competitiveness report, which is debated in CREST⁸ and conferences; in addition, there are specific indicators preparation activities, such as the ERA indicators expert group which mandate was to propose indicators for the monitoring of the European Research Area.⁹

(b) Conferences in scientific and professional contexts

- The International Conferences on Science and Technology Indicators organised by the Centre for Science and Technology Studies (CWTS), University of Leiden, every two years since the early 1990's.¹⁰
- The European Network of Indicator Designers,¹¹ created in 2006 as an offspring of the PRIME Network of Excellence, which provides a forum for discussion an advances in STI indicators, and organises training, projects and a conference every two years.

3.2 Are the functions of the indicators assessment forums fulfilled ?

Collective learning capacity, openness and accountability are the characteristics of the indicators assessment forums whose role is to confer social robustness to S&T indicators. The question is then to know to what extent the various working groups and conferences identified have these characteristics.

The assessment forums characteristics are undoubtedly demanding in time, resources and expertise; they require sound working methodologies and a continuing effort. Furthermore, they have to be recognised as relevant and legitimate platforms by the stakeholders, by indicators specialists, by the actors of research and research management and by

decision-makers. These requirements are certainly not easy to achieve.

It appears that, by definition, official working groups in international contexts tend to be somewhat restricted to 'insiders', be they professionals or representatives of ministries — and conferences of scientific nature, also by definition, may lack the involvement of decision-makers and stakeholders.

If it is true that, over the years, such working groups and conferences have made significant efforts towards the objectives of having collective learning capacity, openness and accountability,¹² the hypothesis of a deficit, or partial deficit, of indicators assessment forums is nevertheless quite plausible.

Conclusions

- The implication of such a possible deficit of true indicators assessment forums would be quite worrying. Therefore it would be important indeed to have an independent and systematic review of where such working groups and conferences really stand regarding these objectives. If the deficit hypothesis was confirmed, initiatives for creating indicators assessment forums should be taken, possibly based on one or several of the existing working groups or conferences.
- Coming back to the collective learning process: it happens that indicators have the two remarkable properties of both constituting a common language (numbers characterising a shared issue) and enabling the actors to express their representations (through discussion of the underlying models and interpretations). They are powerful instruments for collective learning leading to robust interpretations. We suggest this is the way to work with S&T indicators not only for their design as advocated in this article but, more generally, for their use, i.e. for their interpretation.¹³

Finally, far from disqualifying indicators, the understanding of indicators as intrinsically debatable makes them highly relevant for evaluation and decision processes. Being debatable is not a limitation, but is the essence of their contribution to decision-making processes, provided a true interaction takes place among a diversity of actors.

This is a way to be reflexive regarding the conception and use of indicators, which has to do with the needed reflexivity in science in general.

Notes

1. Agence d'évaluation de la recherche et de l'enseignement supérieur; established in 2006, it is the agency which has the task of evaluating all public research units, universities, public research institutions and higher education diplomas.

2. With of course relationships between the two, but they are nonetheless two different things.
3. Strategic and management decisions: resources allocation (grants), human resources management (researchers hiring or promotion).
4. In addition to that, the issue of the 'proxies' must be raised: due to measurement problems, usually one specific parameter will be measured and be considered to 'represent' the broader question at stake. For example, the number of patents of a university will 'represent' its interaction with industry, the assumption being that there is an equivalence between the two. Here again, the expert will fill the gaps with his/her own opinion.
5. Section 2 refers to Gibbons (1999) and Nowotny (2003).
6. Note that such processes are also those which are at the basis of foresight exercises.
7. Institute for Prospective and Technological Studies, which is part of the Joint Research Centre.
8. Science and technology research committee, whose members are the representatives of the ministries of research of the member states.
9. <http://ec.europa.eu/research/era/pdf/era_indicators&monitoring.pdf>
10. <<http://www.socialsciences.leiden.edu/cwts/news/stic2010.html>>
11. <<http://www.enid-europe.org/>>
12. For example, representative of industry are observers in various OECD committees, the EU Commission organises conferences open to stakeholders, and indicators conferences also have sessions with stakeholders and often break up in small discussion groups favouring collective learning.
13. The risk is that actors having a poor indicators record would enter a debate which would amount to a negotiation to raise their mark; proper procedures should avoid such deviation from what is meant here.

References

- Barré, Rémi 2001. The Agora model of innovation systems: S&T indicators for a democratic knowledge society. *Research Evaluation*, **10**(1), April, 13–18.
- Barré, Rémi 2004. S&T indicators for policy-making in a changing science–society relationship. In *Handbook of Quantitative Science and Technology Research*, eds H F Moed, W Glänzel and U Schmoch, pp. 115–131. Dordrecht, Boston, London: Kluwer Academic Publishers.
- Bruno, Isabelle 2008. *A vos marques, prêts, cherchez! La stratégie européenne de Lisbonne, vers un marché de la recherche*. Bellecombe, France: Editions du Croquant, Collection Savoir/Agir.
- Desrosières, Alain 2008. *Pour une sociologie historique de la quantification – l'argument statistique I*. Paris: Presses de l'Ecole des Mines.
- Gibbons, Michael 1999. Science: new social contract with society. *Nature*, **402**, C81–C84.
- Latour, Bruno 1999. *Politiques de la nature – comment faire entrer les sciences en démocratie*, La Découverte, Paris
- Lepori, Benedetto, Rémi Barré and Ghislaine Filliatreau 2008. New perspectives and challenges for the design of S&T indicators. *Research Evaluation*, **17**(1), March, 33–44.
- Nonaka, Ikujiro 1991. The knowledge creating company. *Harvard Business Review*, **69**, 96–104.
- Nowotny, Helga 2003. Democratizing expertise and socially robust knowledge. *Science and Public Policy*, **30**(3), June, 151–156.
- Polanyi, Karl 1966. *The Tacit Dimension*. London: Routledge & Kegan Paul.
- Roqueplo, Philippe 1997. *Entre savoir et décision, l'expertise scientifique*. Paris: Collection Sciences en questions, Editions Quae.
- Van Raan, Antony F J 2004. Measuring science. In *Handbook of Quantitative Science and Technology Research*, eds H F Moed, W Glänzel and U Schmoch, pp. 19–50. Dordrecht, Boston, London: Kluwer Academic Publishers.