European Educational Research Quality Indicators (EERQI): An Experiment

Ingrid Gogolin

Abstract 'European Educational Research Quality Indicators (EERQI)' was a research project funded under the EU 7th Framework Programme from 2008 to 2011. The mission of this project was to develop new approaches for the evaluation of quality of educational research publications. Traditional methods of assessing quality of scholarly publications are highly depended on ranking methods according to citation frequency and journal impact factors. Both are based on methodologies that do not reflect adequate coverage of European scientific publications, namely in the social sciences and humanities. Hence, if European science or institutions are exposed to these evaluation methods, not only individual researchers and institutions are widely ignored, but also complete subject domains and language areas. The initiators of the EERQI project, as well as numerous researchers and evaluation bodies within the European Union, recognized the need to remedy the inadequacies of this situation.

According to our hypotheses, educational research served as a model case for research in the social sciences and humanities. EERQI aimed to

- develop a prototype framework for the intelligent combination of new indicators and methodologies for the assessment of quality in educational research texts,
- make this framework operational on a multilingual basis (starting with English, German, French and Swedish),
- test the transferability of the EERQI framework to another field of social sciences and the humanities.

The contribution[1] focuses on the design of the project and its general aims and basic ideas. In brief, the EERQI-prototype framework is sketched: what is it about? How is it composed? What is its scientific and practical value?

---

[1]This article is based on a contribution to the Conference 'Research Quality in the Humanities', Zurich, October 2010. My thanks go to Virginia Moukouli for her support of the presentation.

I. Gogolin (✉)
University of Hamburg, School of Education, Institute for International Comparative and Intercultural Education, Von-Melle-Park 8, 20146 Hamburg, Germany
e-mail: Gogolin@uni-hamburg.de

103

1 General Outline: The EERQI Project

In order to understand the scope and aims of the EERQI project, a brief excursion to the context of the endeavor should be helpful: why was it felt necessary to start the EERQI-project?

1.1 Motivation

All across the world the structures and control mechanisms of publicly funded research projects have changed dramatically in the last decade. There are many widely discussed causes of these developments. The set of causes on which we concentrate here is based on the evocation of the 'ability to compete internationally'—a request that is expressed vis-à-vis national research landscapes in Europe as well as the European research space.

A metaphor that is either explicitly used or implicitly resonates in the existing discourses, in the decisions on new governance mechanisms and in new modes of research funding is *quality*. The discovery, improvement and promotion of research quality are the driving motives for the tendency to re-evaluate and redevelop structures for the research area, for redesigning the funding of research institutions and projects, and for instituting control and legitimization systems that are (or intend, or pretend to be) helpful for decision-makers.

In the framework of these developments the questions of *how quality is interpreted* and *how it is measured* are of fundamental importance. Analyses dealing with this question supplied the starting point for the development of the research project 'European Educational Research Quality Indicators (EERQI)'. The project was developed by a truly interdisciplinary European research consortium, a unique composition of experts from Educational Science, Biblio- and Webometrics, Information and Communication Technology, Computational Linguistics and Publishing Houses. It received funding under the Social Sciences and Humanities Funding Scheme of the European Union's 7th Research Framework until March 2011.[2]

The focus of the analysis prior to the project was on particular questions such as: What constitutes and marks the current quality control systems that are applied in contexts of governance and funding, irrespective of the genre and type of research that is at stake? And what are possible effects of these systems on research that is conducted in the European Research Area, especially in the domains of Social Sciences and the Humanities?

According to our assumptions, educational research is especially privileged for considerations and research on such questions because it can be considered as prototypical for vast areas of the whole field of social Sciences and Humanities.

---

[2]For details have a look at the EERQI website: http://www.eerqi.eu; see also Gogolin (2012) and Gogolin et al. (2014).

This legitimates as follows: Education science and research combine a wide spectrum of theoretical and methodological approaches—from a primarily philosophical-historical methodologies as used in the humanities to psychologically or sociologically based empirical observations of individual development, education, training or *Bildung*; from hermeneutical interpretation, single case studies to the generation and statistical analysis of great amounts of survey data. This manifests relevant characteristics of knowledge production which are also found in other disciplines in the Social Sciences and Humanities.

The EERQI review on the appropriateness of instruments and strategies for quality assessment that are actually applied to educational science resulted in a generic judgment that can briefly be articulated as follows: The existing instruments do not lead to valid results because they do not measure what they claim to measure. An example for the illustration of this statement is quality assessment based on citation indices and journal rankings. This is, at least as yet, the most common approach in vast areas of quality assessment.

The central criterion that is used in these instruments is 'international visibility' of research findings. This is expressed by the placement of the publication, namely in journals with a good reputation, and by the number of citations of a publication. This approach is characteristic of the Social Science Citation Index, a commercial instrument owned by the US-American publishing group Thomson Reuters. Its results often play an important role in reporting systems on research achievement. A closer look at the documentation of the journals represented by this index reveals (for 2009 and the field of educational science according to the 'Journal Citation Report'[3]) the following:

A total of 201 educational research journals were incorporated in the rankings in 2009. Approximately 52 % of these journals were published by US-American publishers. An additional 24 % derived from British publishing houses. The next 'largest' nations in this ranking were the Netherlands (with 4 % of cited journals) and Germany (with 3 % of cited publications). All together 15 nations across the world were represented in the ranking of the Journal Citation Report. A slightly different perspective reveals that 89 % of the publications were in English. The next 'largest' languages with 2.5 % and 2 % respectively were in German, Spanish and Turkish. Eleven languages in total were represented by the index. A language such as French was not included.

We have to admit that the Thomson Reuters-Group itself recently started with a revision of their policies of including journals into the rankings. The Group has incorporated additional journals from other areas of the world into their system—this may be a reaction on international criticism of the instruments, and EERQI may have played a modest role in this. But nevertheless the findings illustrate that these

---

kinds of approaches do not produce valid information in the sense they pretend to do, because the intended international relevancy of the included publications cannot be proven. The rankings are still heavily biased: they essentially refer to US-American or UK-publications and publications in English. International visibility as a quality criterion must be translated here to: the visibility of products from a selection of national research spaces to the rest of the world. The provided information is perfectly suitable to substantiate the powerful dominance of a 'minority' of regional and linguistic research areas.

It is unfortunate that other regional and linguistic research areas, which do not have the benefit of this reinforcement of dominance, participate actively in cementing and safeguarding the existing pattern. This is not least the case in Europe. Prominent research funding institutions affirmatively employ methods that lead to the illustrated result and thus fortify their importance. An example: Calls in the framework of European Research Council's ERC Grant Schemes include the following advice that implies, as we may assume, criteria for the evaluation of proposals. Applicants are asked for 'A list of the top 10 publications, as senior author (or in those fields where alphabetic order of authorship is the norm, joint author), listing all authors, in major international peer-reviewed multidisciplinary scientific journals and/or in the leading international peer-reviewed journals and/or peer-reviewed conferences proceedings of their respective research fields, also indicating the number of citations (excluding auto-citations) they have attracted and possibly the h-index (if applicable)' (European Research Council 2011).[4]

Negotiations about possible alternatives for the assessment of quality in research areas that are not appropriately mirrored in these kinds of methodologies have as yet not been overwhelmingly successful. An example for this is the British Research Excellence Framework—the system for assessing the quality of research in the UK Higher Education system.[5] The 2011 Higher Education Funding Council in England (HEFCE) Report on a pilot exercise to develop bibliometric indicators for the Research Excellence Framework, a review that was used for the preparation of the British Research Excellence Framework, stated: 'The pilot exercise showed that citation information is not sufficiently robust to be used formulaically or as a primary indicator of quality; but there is considerable scope for it to inform and enhance the process of expert review'.[6] Hence, whilst fully aware of the constraints of these methodologies, the respective instruments and data deriving from them, they are extensively in demand and applied by the bodies that conduct processes of research assessment and governance (for the development of this see Oancea 2014).

---

[4]In recent calls, requirements are described in less detail, but still insist on publications in 'the leading international peer-reviewed journals' (see for example ERC-CoG-2015 on http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/9063-erc-cog-2015.html, accessed 15th December 2014).

[5]See http://www.ref.ac.uk, accessed 9th December 2015.

[6]See http://www.ref.ac.uk/about/background/bibliometrics/, accessed 9th December 2015.

1.2 The EERQI-Project

The motivation for the development of the EERQI-project, in a nutshell, was the observation that the strategies of assessment that were developed in 'hard science'-contexts are heavily criticized for their methodological weakness and lack of validity—not only from a social sciences and humanities point of view (Bridges 2009; Bridges et al. 2009; Gogolin and Lenzen 2014; Mocikat 2010).[7] At the same time there is a serious desire to dispose of approaches that can serve better for the aim of detecting research quality. This desire unites the research community as well as relevant stakeholders from other spheres, such as publishing houses, research funding, political decision making.

The initiators of the EERQI-project never had the idea to take up a battle and try to compete with the economically powerful suppliers of approaches like the Thomson-Reuter's, Scopus (Elsevier) or similar players. Our general intention was to develop useful tools that support the *process* of quality detection. An intelligent combination of such tools—that was our assumption—could be able to assist the readers in the task of determining the class and value of a single text or a series of research texts, be it for assessment purposes or for information in a research process. The application of these process-oriented tools should meet two aims:

1. It should raise the transparency and quality of the process of quality detection itself;
2. It should make the task better manageable and less time-consuming.

In order to meet these aims, EERQI's objective was *not* to develop one single method, such as an index. Instead we aimed to develop and test a *set* of tools that can be applied in different stages of an assessment process, as single methods or in intelligent combinations. These tools should be based on explicit criteria that make the assessment process and result more transparent. In other words: EERQI did not aim at replacing the human decision making in evaluation and assessment procedures, but at maintenance for the individual actors in the procedures or for groups of actors, such as assessment boards. The set of tools we developed is what we call the EERQI Prototype Framework.

2 What EERQI Achieved

The EERQI Prototype Framework is based on the products that were developed in the course of the project. It consists of the following

- a content base with educational research texts in the four European languages that were included in the EERQI project as examples for European multilingualism: English, German, French and Swedish.

---

[7]See also http://www.adawis.de/index.php?navigation=1, accessed 22nd May 2011.

- a multilingual search engine that includes query expansion: an effective tool dedicated to educational research in general, capable of finding educational research texts in the web in the four EERQI languages.
- automatic semantic analysis for the detection of key sentences in a text. This method is applicable to educational research publications (for a start) the four EERQI languages.
- a combination of bibliometric/webometric approaches for the detection of 'extrinsic' quality indicators (a tool named aMeasure).
- first tests of a citation analysis method that has the potential to be further developed for the application to educational research (and other SSH) texts.
- a set of text-immanent (intrinsic) indicators for the detection of quality in educational research publications that has been presented to the research community and was positively evaluated.
- an accompanying peer review questionnaire that was tested for reliability and feasibility of the instrument.
- a set of use case-scenarios that advice on how to use which resp. combination of the above mentioned tools.
- first attempts to detect interrelations between 'extrinsic' and 'intrinsic' quality indicators.
- and last not least: a successful test of transferability of the approaches developed in EERQI to political science, another areas of social sciences and humanities.

All products are accessible via the EERQI web site (http://www.eerqi.eu).
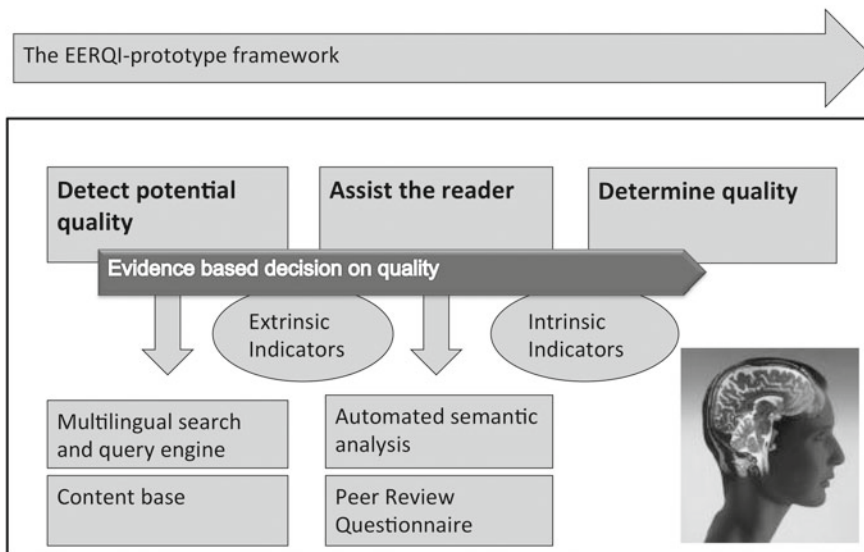Figure 1 illustrates the Prototype Framework and its elements.

Fig. 1 The EERQI prototype framework

The elements of the EERQI Prototype Framework can be used during the process of quality detection. After detecting and identifying relevant texts, different approaches to consolidate a judgment on quality can be applied. The EERQI project distinguished two different types of indicators that are relevant in these approaches: one type that is external to the text, such as bibliometric and webometric features; and another type that is internal in the text—namely the signals that are given within the words, graphs, metaphors of which the text is composed. The application of EERQI tools in the process can be illustrated as follows:

1. detection of potential quality via the identification of relevant educational research texts in different sources. In this step, the EERQI content base (educational research texts provided by the EERQI publisher partners) and the multilingual search and query engine can be applied.
2. gathering information on extrinsic features of a text. For this purpose, an instrument called 'aMeasure' was developed (by EERQI partner Humboldt University). This is a stack of tools and programs which indicate extrinsic characteristics of research publications (such as citations, webmentions) by using different sources (e. g. Google Scholar, Google Web Search, MetaGer, LibraryThing, Connotea, Mendeley and citeulike) and combining their results, thus providing more comprehensive and less (but still!) biased information.
3. supported transverse reading, allowing for quick information on the usefulness or quality potential of a text. For this step, a linguistic technology in order to provide automatic support for evaluating the quality of a text was developed (by EERQI partner XEROX). The method allows for the automatic identification of key sentences to indicate parts of documents to which peer reviewers should pay particular attention (automated semantic analysis). The respective tests in the EERQI-project showed that this method is especially efficient for the identification of the *bad* quality of a text. It can reduce the time that has to be spent on a text in a review process considerably (up to two thirds of reading time).
4. support of a peer review process. For this step, the EERQI project developed a questionnaire containing items that operationalize five generic indicators of research quality (EERQI Peer Review Questionnaire). The indicators as well as their operationalization in the questionnaire have been tested for reliability, practicality and acceptance in the education research community—with very satisfactory results.

The elements of the EERQI Prototype Framework can either be applied as single methods for specific parts of an assessment process; or they can be applied consecutively, leading to a final judgment on the basis of intense reading of selected texts.

3 Conclusion

The approaches that were generated and tested in the EERQI project open up prospects for future developments that can meet the practical needs in accelerating assessment processes and make them better manageable as well as more transparent. Both are necessary, not least because the number and aspiration of such processes are continuously growing. The intelligent combination of qualitative and quantitative approaches, and the multilingual functionalities of the EERQI products, open up the vision that sets of tools can be made available, allowing for well-informed, evidence based judgments on research quality that are supported by technical tools. The application of the tools can accelerate the process and increase transparence—but cannot replace the human judgment. There cannot be any doubt that the EERQI experimental approach had some methodological limitations (see for example the contributions by Mooij (2014) or by Severiens and Hilf (2014) in Gogolin et al. (2014)). Nevertheless, the present empirical outcomes of the project are promising for future EERQI developmental and research activities, which could, for example, also integrate semantic latent factors and indicators. The approaches that were developed and tested in EERQI show encouraging possibilities to appraise Europe's multicultural and multilingual heritage in research, especially in the social sciences and humanities.

References

Bridges, D. (2009). Research quality assessment. Impossible science, possible art? *British Educational Research Journal*, *35*(4), 497–517. doi:10.1080/01411920903111565.

Bridges, D., Smeyers, P., & Smith, R. (Eds.). (2009). *Evidence-based educational policy. What evidence? What basis? Whose policy?*. Oxford: Wiley Blackwell.

European Research Council. (2011). ERC Grant Schemes Guide for Applicants for the Advanced Grant 2011 Call. Retrieved from http://ec.europa.eu/research/participants/portal/doc/call/fp7/erc-2011-adg_20110406/30876-final_gfa_adg2011_en.pdf.

Gogolin, I. (2012). Identificación de la calidad en las Publicaciones de Investigación Educativa. Proyecto Europeo sobre los Indicadores de Calidad en la Investigación Educativa (EERQI). *Revista de Investigación Educativa*, *30*(1), 13–27. doi:10.6018/rie.30.1.140812

Gogolin, I., Åström, F., & Hansen, A. (Eds.). (2014). *Assessing quality in European educational research. Indicators and approaches*. Wiesbaden: Springer VS.

Gogolin, I., & Lenzen, D. (Eds.). (2014). *Qualitätsmessung im Bildungs- und Wissenschaftssystem*. Wiesbaden: Springer VS.

Mocikat, R. (2010). Qualitätsbewertung in den Naturwissenschaften mithilfe quantitative Parameter. *Journal der Sächsischen Akademie der Wissenschaften*, *5*, 90–102. Retrieved from http://repo.saw-leipzig.de:80/pubman/item/escidoc:16291/component/escidoc:16290/denkstroeme-heft5_90-102_mocikat.pdf.

Mooij, T. (2014). A prototype empirical framework of intrinsic and extrinsic EERQI indicators. In I. Gogolin, F. Åström, & A. Hansen (Eds.), Assessing quality in European educational research. Indicators and approaches (pp. 121–138). Wiesbaden: Springer.

Oancea, A. (2014). Research assessment as governance technology in the United Kingdom. Findings from a survey of RAE 2008 impacts. *Zeitschrift für Erziehungswissenschaft*, *17*(6 Supplement), 83–110. doi:10.1007/s11618-014-0575-5

Severiens, T., & Hilf, E. R. (2014). A scientific editors's support tool: Design, analysis and value. In I. Gogolin, F. Åström, & A. Hansen (Eds.), Assessing quality in European educational research. Indicators and approaches (pp. 139–164). Wiesbaden: Springer VS.