Ferenc Kiefer

ERIH's role in the evaluation of research achievements in the Humanities

DRAFT

My talk will be organized as follows. In the first part I am going to say a few words about ERIH, its aims and its possible role in the evalution of research in the humanities. In the second part I will discuss the main differences in publication culture between sciences and humanities, and summarize the main changes in publication practices in the Humanities which have occurred during the last few decades. In the last part I will take up the problem of bibliometrics as an evaluation tool in humanities research.

1. ERIH

ERIH stands for „European Reference Index for the Humanities" and it is an ESF project that was conceived as constructing a database of journals aiming initially at identifying top-quality European research in the Humanities, published in academic journals in all European languages. The background for launching this project in 2001 was the need felt by European researchers for better databases than the existing ones, which are limited in coverage (even some of the best English language journals are missing from the database and there are even bigger gaps in the coverage of journals in other European languages) and tend to be centred on Anglo-American publications. Research conducted in national (especially in so-called lesser used) languages is either not adequately covered or not covered at all. Furthermore there are specificities in the Humanities in terms of cultures of publication and traditions of citations, which make it meaningless to work with the evaluation tools used in science. ERIH's main objective was to remedy this defect by providing a more reliable tool for research assessment in the humanities.

As a first step toward a more adequate database, lists of journals in fourteen disciplines have been compiled and classified into three categories. The international journals include two subcategories, which differ basically with respect to influence and scope. The first subcategory, called A, covers high-ranking international journals with high visibility that are regularly cited all over the world and have a very strong reputation

among researchers in different countries. The second subcategory, called B, includes standard international journals with significant visibility and influence in the various research domains in different countries. Finally, national or regional journals include standard and high ranking national publications with a recognized scholarly significance among researchers in a particular (typically linguistically circumscribed) readership group in Europe. The main target group of national journals is the domestic academic community. In the case of national journals only European journals are considered.

In principle, there should be no qualitative difference between articles published in international and national journals, the difference has to do with the language and with the distribution and the visibility of the journal. That is, the categories A/B and C are not meant to represent any ranking. To be sure, consistently high-quality scholarly content is more typical of A than of B journals and A journals are certainly more visible than B journals. On the other hand, however, a paper published in a B journal need not be qualitatively inferior to A-category publications. Yet the overall higher quality of articles in A journals does provide a ranking.

Aspects of quality become also important in the case of the criteria used to decide whether a journal qualifies for inclusion in an ERIH list or not. For all categories, in order to be included, a journal must fulfil normal academic standards of quality, i.e. it must have a quality control policy that governs the selection of articles and it must also fulfil basic publishing standards such as ISSN, timeliness of publication, active and international editorial board for international journals and an active editorial board for national journals.

The judgement to which category a journal belongs is made by an Expert Panel of peers for each discipline. Great care was taken to select solid scholars with an international reputation as panel members, and to cover as many sub-fields and linguistic areas as possible.

As already mentioned, the lists were meant to serve as a database of journals aiming initially at identifying top-quality European research in the Humanities published in academic journals. The lists as they stand are not a bibliometric tool and should not be used as such. But we believe that it is necessary that peer review in the humanities be informed by better data. One of the great advantages of ERIH is that the database also

includes European research output published in lesser used languages. Such a database did not exist before. It is believed that this approach will enable prompt comparability with other sciences.

To be sure, there exist other quality-based journal databases, such as ISI Thompson's Arts and Humanities citation index or, more recently, Scopus. Not aiming at the broad European Humanities constituency, such databases have not included multilingual European research output in the global context of quality assessment and access. Moreover, their tools tend be parameterized to the rhythms of knowledge decay in most natural sciences. In the Humanities the lifecycle of research is much longer than in hard sciences since earlier findings are often reinterpreted rather than superseded and standard works may never become outdated.

So far we did not mention the problem of monographs, which – at least in some disciplines in the Humanities – still constitute the most important publication genre. Up till now ERIH has concentrated its efforts to survey and classify journals but the need to include monographs has always been acknowledged. Reviewing and classifying journals is already a difficult task but to include monographs into this procedure is even more challenging. No doubt, the categories national/regional and international will play a role in classifying monographs as well. The evaluation of books will have to take into consideration at least the following aspects: the international or national prestige of the publisher, the scholarly standing of the series editor or of the edited volume, and distinction will have to be made between conference proceedings, Festschrifts, books whose target readership is not the scholarly community, etc. Furthermore, it must be taken into account whether the publisher requires anonymous reports of the manuscript submitted for publication. Last but not least the published reviews of the monograph, too, must be part of the evaluation. Classifying books is thus radically different from classifying journals and it requires a very special methodology. The elaboration of such a methodology will be the task of the next ERIH phase.

2. Changes in publication culture

Since the late 1970s and early 1980s, the classical 'von Humboldt' university model characterized by learning through science and unity of research and teaching suffered

gradual transformation under the influence of changes in social, economic and technological conditions. In addition to the classical mission of knowledge creation and conservation for the next generations and the education of tomorrow's elite, society gave new tasks to universities. Nowadays academia is called to play a more active role both in solving new societal problems and in strengthening economic development. Concrete goals are often formulated in management contracts together with performance indicators to measure their realization.

The publication culture has undergone drastic changes since the beginning of last century. Up to mid-20th century Humanities was depicted as disciplines dominated by books, which rely largely upon older literature. In Europe the main publication languages in the humanities were French and German. Moreover, the individual disciplines did not have as yet a 'central journal' or 'central journals'. The situation was quite different two decades later as reported by a number of studies (e.g. A.J.Nederhof and R.A.Zwaan: „Quality judgements of journals as indicators of research performance in the humanities and social and behavioral sciences", *Journal of the American Society for Information Sciences*, 1991, and A.J.Nederhof, „A bibliometric assessment of research council grants in linguistics", *Research Evaluation*, 1996). It has been shown that journals are of prime importance to scholars in various humanities fields. Moreover, many humanities fields are characterized by central or 'core' journals. 'Core journals' were defined as journals that are well known by the scholarly community, that receive high ratings of scholarly quality, and that are found very useful to one's own research. Concerning the presumed reliance on older literature it was shown that 36% of the references in core linguistic journals referred to literature published during the past four years. Although this percentage is below that in some of hard sciences, it is comparable to many others.

The shift from books to journal articles may be different in different disciplines. Linguistics is a good example where monographs have become less important, in literary studies, on the other hand, the monograph is still a major publication genre. A similar difference can be observed between, say, psychology and history. Is it possible to draw a general conclusion from such observations? Is the monograph as a publication genre in the Humanities in danger? This may be the case, for example, in literary studies and in history but much less so in linguistics and psychology where monographs had not played

such a central role before mid-19th century either. Consequently, the changes in publication culture did not affect all humanities disciplines to the same degree. The closer the research methodology in a humanities discipline is to the research methodology in natural sciences, the more importance will be attributed to journals.

Scholars in humanities serve two publics with their products: research products are directed either to a scholarly public, or to a primarily non-scholarly public. In the latter case we speak of ’external knowledge transfer’. As a great number of studies have shown, a considerable (sometimes even a major) part of the scholarly outcome of scholars in various humanities sub-fields is directed towards the ’enlightment’ of the general public. This implies that humanities scholars also publish in magazines and newspapers whose target readership is not the scholarly community. Of course, these publications should not be included under the heading 'scholarly articles' in lists of publications.

A further difference between publication culture in science and in humanities is that in the latter we may encounter ’double’ publications quite often, i.e. a paper written in a lesser used language published in a C category journal may also be published in a major international language. This state of affairs has some important consequences for evaluation since the number of publications will not adequately reflect the scholarly output.

Yet another aspect concerning the difference between science and humanities has to do with the length (in terms of printed pages) of an article. Whereas a mathematical proof may need not more than one page, or a report on an experimental result in neuropsychology may take up just a few pages, it is hardly possible to produce a serious scholarly work in literary studies or linguistics on a few pages. Consequently, size is part of the notion of article in the humanities but not in science. It does not come as a surprise, then, that in lists of publications we often find two categories, shorter notes are listed separately from genuine scholarly articles. (It may be noted that most evaluation studies in the humanities are aware of this problem and it was suggested that only publications with a length of more than five pages should be regarded as ’substantial’ contributions.)

The differences in publication culture between science and humanities reflects a fundamental difference between two types of research activity. Humanities scholars use

research retrospectively from various time periods, representing often conflicting intellectual positions against which scholars define their views and contributions. In the humanities it is not research that cumulates, as in the case of sciences, it is rather the scholarly literature that cumulates with a negligible degree of obsolescence.

The differences between the two publication cultures (and between two types of scholarship) were discussed in considerable detail in a recent study by Stephen Wiberly ("A Methodological Approach to Developing Bibliometric Models of Types of Humanities Scholarship", 2003). Wiberly posits five types of humanities scholarship: (1) descriptive bibliography, (2) editing, (3) historical studies, (4) criticism, and (5) theory. He examines their bibliometric characteristics and demonstrates how one can differentiate these various types of scholarship based on bibliometrics and not on subjective impressions. For example, historical studies contain a high percentage of citation older by 20 years, criticism scholarship utilizes much more recent materials and a high percentage of secondary sources, theoretical scholarship cites a high number of very recent materials most of which are theoretical in nature. The five categories are clearly modelled on literary studies and art scholarship and cannot therefore be taken to be general characteristics of humanities research, but the suty clearly demonstrates the efforts to make judgements about scholarship more reliable..

What has been said about the general characteristics of humanities research need not be true for all humanities disciplines, however. At least two humanities disciplines, theoretical linguistics and cognitive psychology seem to have more features in common with science than with traditional humanities research, as far as methodology is concerned. In these two disciplines articles in journals are more important than monographs, the monographs either summarize earlier research already published or they are textbooks designed for students or interested laymen. The references include mainly recent publications on the topic. Consequently, many articles in theoretical linguistics and cognitive psychology share the fate of articles in some life sciences: their lifespan is not very long. But we have to accept the fact that humanities covers a wide range of disciplines and each discipline may have its specific features which have to be taken into consideration when evaluating the research output.

It has often been claimed that citation index (impact factor) is not very useful in humanities. De Bot („Het gebruik van citatie-indexgegevens bij de kwaliteitsbeoordeling van publicaties in taalkundige tijdschriften", *Gramma,* 1987) found that there is no relation between the impact factor (as provided by ISI), and the impact of individual articles in the humanities journals considered. Nederhof and Zwaam (1991) studied the importance of citations in various humanities disciplines. They came to the conclusion that the citation indexes are potentially useful for citation analysis in largely internationally-oriented humanities fields, whereas for nationally-oriented fields, citation data fail to represent a valid picture of the national importance of research, but still may offer an indication of its potential international impact.

In sum, then, we may conclude that some humanities disciplines may have their own publication practices, and some of these may not differ essentially from hard sciences in this respect.

3. Bibliometrics and evaluation in the humanities

We live in an age of metrics. All around us, things are being standardized, quantified, measured. Scholars concerned with the work of science and technology must regard this as a fascinating and crucial practical, cultural and intellectual phenomenon. In this respect, Humanities are not an exception. Bibliometric indicators are being required for the evaluation of research output in the Humanities as well. There is a general consensus among researchers in the Humanities that bibliometric indicators, complemented with interviews, questionnaires and other qualitative information, provide a good instrument in research management. It has repeatedly been stressed that quality assessment of research performance can only be made by informed peers. Quantitative results can be used as background information to allow such experts to better form their opinion.

It became soon evident that no unanimously accepted methodology was readily available to highlight humanities (and social sciences) research activities. Elaboration of a solid methodology, to take into account the specific characteristics of each discipline, turned out as a matter of concern not only to those in charge with elaborating and implementing research policy, but also to the researchers involved in these disciplines. Discussion focused on research policy and allocation of research funds often led to

acrimonious exchanges between ’hard’ sciences and ’soft’ sciences protagonists. Researchers in natural and life sciences have often shown a tendency to assess work in social sciences and humanities within the framework of their own disciplines.

Only a modest part of bibliometric studies has dealt with humanities. One of the recent studies (Mariela Hristova: Bibliometrics and the Humanities, Texas, 2006) provides an overview of bibliometric studies dealing with the Humanities and discusses their implications for future research. One of the large trends in humanities bibliometrics research was a tendency to focus on the differentiation between scientific and humanities scholarship and trying to draw a distinct and accurate portrayal of humanities research and its characteristics. The differences in publication culture reflect the differences in research activity. The author points out that humanities scholars differ from scientific researchers in that they work individually, instead of collaborating, using materials that span a wide range of years in terms of publication, which demonstrates that humanities research is not as susceptible to obsolescence as scientific research. Also, interpretation is paramount in humanists’ work because they do not report studies done outside of their writing; their publications constitute the research itself due to the focus on interpretation and analysis of primary sources, such as archives or works of art and literature.

A fair and just research evaluation should take into account the diversity of research output across disciplines and include all major forms of research publications. While journal ranking based on impact can help achieve this in many disciplines that have formalized criteria for research and publications, the impact factor for humanities journals is much harder to calculate (if at all) and is less reliable for assessment of research quality. Therefore a more comprehensive bibliometric approach is in order to establish standard criteria for assessing research quality across each discipline.

A study performed by Henk Moed, Marc Luwel and A.J.Nederhof („Towards research performance in the humanities”, *Library Trends*, 2002) attempts to establish the foundation of a comprehensive bibliometric methodology for assessing the research performance of all scholars within any discipline in the humanities. The methodology they propose has been successfully tested in the field of law research. The methodology includes several aspects: detailed classification of publications, ranking of journals, and input from scholars on their practices and perceptions. Based on the finding of their study

of law research they conclude that a comprehensive and successful methodology should be concerned with developing accurate and discipline-specific indicators of research performance. They suggest that this can be achieved by collecting data on publication output directly from the scholars, verifying this data electronically and classifying it in meaningful ways to separate the substantial research contributions from the incidental ones. The classified research can then be further weighted through the use of journal ranking derived partly from scholars' perceptions and partly through bibliometric analysis.

Moed, Luwel and Nederhof's methodology constitutes a significant contribution to bibliometric research in the humanities because it offers a method that utilizes the strengths of quantitative research approaches without ignoring the benefits of qualitative approaches. By collecting information on scholars' output and perceptions, the research methodology involves the community in question and allows them to participate in the formulation of research quality indicators that might not be apparent to an outsider to the discipline. Meanwhile, the actual calculations of research quality remains dependent on a scientific methodology that is both valid and feasible.

It is normally assumed that the international orientation of a discipline is related to the object of research in that discipline. However, we wish to maintain that genuine scholarly research, regardless of the sub-discipline and the object of research, leads to results whose relevance and implications go beyond a purely national viewpoint of interest. Therefore, outcomes of genuine scholarly research, even the ones primarily related to national aspects, deserve to be communicated – in an appropriate form – to scholars outside the country as well.

ERIH, and the wider issues it raises about the definition and identification of excellence in the Humanities, have prompted an ongoing series of reflections about the usefulness of bibliometrics for the Humanities, acknowledging at the same time the need to develop a much wider array of measures of impact. For the time being, however, peer review has been given precedence over a quantitative approach, which remains the standard method used by research communities to identify excellence. Peer review has its advantages and its disadvantages: it can detect originality, but it has also the potential to defend conservative approaches. Generally speaking, however, it is acknowledged that

the peer review introduces a measure of comparability into discussions of different national discourses in Humanities scholarship. ERIH relies on the principle of peer review to identify quality through open scholarly debate, and on a lengthy process of consensus building. It is hoped that this may bring us closer to an objective evaluation of scholarly output.

There is a general consensus among researchers and academic authorities that bibliometric indicators, complemented with interviews, questionnaires and other qualitative information, provide a good instrument in research management. It has, however, to be stressed that quality assessment of research performance can only be made by informed peers. This is the point where ERIH may become important.