Innovative Scientometric Methods for a Continuous Monitoring of Research Activities in Educational Science

Werner Dees

17 October 2008

Abstract

This paper reports on a project sponsored by the German Research Foundation (DFG) aimed at developing new methodological approaches to evaluate the relevance and quality of educational research publications and at examining the potential of the German Education Index (FIS Bildung) as a tool to monitor and map educational research. The project is exploratory and its objective is to develop a methodology for constructing research performance indicators, not to assess or compare research performance of individual scholars, research groups or university departments. The paper gives an outline of the project, presents results of its first part, the analysis of the publication behaviour of exemplary educational research institutions, and finally illustrates the development of a new multidimensional quality indicator.

1 Introduction

Evaluations of research performance become increasingly common and frequent in all academic fields including the social sciences and humanities and consequently educational science is also affected by this trend. In this context there is a growing need to design and operationalize standardized quality indicators which are adequate to the respective discipline.

The development of such indicators is complicated by the fact that international instruments for analysing scientific publications include German educational science only marginally, and German databases inadequately support the demands of scientometric analysis. But, as Nederhof et al. (2001: 243), have pointed out, “bibliometrics is much more than conducting citation analyses based on the ISI citation indices.” Consequently, the growing need on the part of research managers and administrators to assess scholarly work in the social sciences and humanities in a similar form as it is usual in the natural and life sciences has led to an increase of bibliometric analyses of these fields (cf. e.g. Moed et al. 2002). To conduct an evaluation study it is useful to consider the results of descriptive studies on the communication practice in social sciences and humanities which have shown some specific features, like the predominant use of national languages, the importance of books and edited volumes and the less prominent role of journals (Nederhof et al. 2001: 241, Nederhof 2006). However, there is also some evidence that there are relatively large differences in the publication behaviour among the various disciplines of social sciences and humanities and that there is a change in publication patterns over time (Hicks 1999; Hicks 2004, Nederhof 2006). So, to construct an adequate methodology for the educational science in Germany, it is imperative...
what Luwel et al. (1999: 13) pointed out in a similar context: one should take into account the present characteristics of this field of scholarship, the nature of the scholarly research object, the communication practices among scholars and the structure of its communication system.

Therefore the starting point of the project and the main topic of this paper is an empirical analysis of the actual publication practice of German educational researchers to answer the question: How does the scholarly communication in this discipline with regard to certain characteristics, such as publication types and languages, look like? Thus, this study is the first application of an empirical publication activity analysis to the discipline of educational science in Germany and forms the basis for the development of a new bibliometric indicator.

2 Background and outline of the study

The study presented in this paper is part of an ongoing exploratory project sponsored by the German Research Foundation (DFG) and supported by the German Corporation for Educational Science (DfGfE). The project is a result of a continuous debate within the community of educational researchers on the possibilities of evaluating the research performance of German educational science (cf. Hornbostel and Keiner 2002; Hornbostel 2005; Böllert 2004; Tippelt 2004). It is carried out by the Information Centre for Education at the German Institute for International Educational Research (DIPF). This centre, in cooperation with 30 partners, produces the bibliographic database German Education Index (FIS Bildung) which is the basis of previous evaluation efforts in this research area, e.g. by the Centre for Higher Education Development (CHE). The project team is assisted by an advisory board composed of educational researchers, bibliometricians and information scientists.

The project is designed as an exploratory and methodological study. Its objective is to develop a methodology for constructing research performance indicators, not to assess or compare research performance of individual scholars, research groups or university departments. The first, already completed, step of the project’s work plan was the analysis of the publication activities of selected educational research institutes and university departments with regard to the prevalent publication types and languages. The motive for this step is to gather more specific information about the current communication practices of the scholars in the respective field. For that purpose the publication lists of selected, pertinent educational research institutes and university departments were collected from the institute’s homepages or university bibliographies and analysed statistically. Research questions in this part of the project were e.g.: How important is the role of journals, especially ISI-indexed journals, in the communication among scholars? How important is the role of (edited) books? Are there differences between university departments and research institutes? This step is supplemented by an analysis of the ‘topography’ of educational science publications, i.e. the important ‘places’ (publishers of monographs and edited books, journals, scientific series, digital repositories) educational researchers use as their publishing outlets are identified. This extended analysis is based on the data of the collected publication lists and data of the German Education Index. In subsequent phases the project intends to develop a new multi-attributive indicator of scientific relevance.

The project resembles in some respects approaches by the Centre for Science and Technology Studies (CWTS) aiming at the development of research performance indicators for the social sciences and humanities (Luwel et al. 1999). Their study was also exploratory, but concerned with linguistics and law, not educational science. Yet, it provides an important point of reference for our project.

---

2 Although there are comparable studies, like Baumert and Roeder (1989) or MWFK (2004) they are limited to certain types of publication or to only one federal state of Germany.

---

3 At the same time this part of the project aims at the examination of the representativeness of the German Education Index by matching the collected documents with those in the database and thereby determining its goodness of coverage. One could label this step as evaluating the evaluation instrument.
3 Data

This section outlines briefly the data collection process and describes the data used as the basis of the publication behaviour analysis. The selection of the research institutes was specified in the project proposal. These are nine renowned institutes in the field, each of them covering mainly a specific subfield of educational research, like adult education, early childhood research or higher education research (cf. Weishaupt 2002). Starting point for the selection of university departments was the CHE university ranking which identified ten universities as being especially strong in educational research (Berghoff et al. 2006). Since the goal of the project was not to evaluate the research performance of educational science departments, but to describe the publishing behaviour of some exemplary institutions, a pragmatic approach was chosen to decide which of these universities should be included in the study. The decisive factor was the accessibility of the bibliographic data. This criterion reduced the possible candidates to four which were complemented by two other universities with large educational research departments and comparatively easy access to their publication lists4. From these altogether fifteen institutions (preferably) complete publication lists were collected. The time period considered was 2004–2006.

Different options were used to collect the data depending on the institution’s availability of data. In most cases the lists were provided by the institutions themselves or institutional publication databases and university bibliographies were accessed and searched for relevant documents. In few other cases the personal homepages of each staff member of an institution had to be scanned manually for his/her publications. Another approach was chosen by Luwel et al. (1999) who collected the bibliographic data via questionnaires in which the scholars had to list their publications themselves. Apart from the problem of incorrect information sometimes given by the surveyed scholars, the problem of non-response seems to be a major drawback of this procedure. But also because of time restraints of our project this way would not have been possible.

The consequence of our pragmatic approach was that the collected information (metadata) about the publications was rather inconsistent and heterogeneous between the institution’s bibliographies and even within some of them. The minimal information for each publication was publication year, author(s) or editor(s), publication title and, in case of articles or book chapters, the title of the source. In addition most lists provided information on the publisher, i.e. the name and location of the publisher, and the publication length, at least for articles and book chapters.

The publication type and language was recorded only rarely however. Thus these variables had to be supplemented automatically or by inspection in many cases. The resulting sample of this data gathering process differentiated by publication year and type of institution is presented in Table 1.

Table 1: Collected publication data

<table>
<thead>
<tr>
<th></th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research</td>
<td>1,000</td>
<td>985</td>
<td>1,009</td>
<td>2,994</td>
</tr>
<tr>
<td>institutes</td>
<td>33,4%</td>
<td>32,9%</td>
<td>33,7%</td>
<td>100%</td>
</tr>
<tr>
<td>University</td>
<td>800</td>
<td>558</td>
<td>342</td>
<td>1,700</td>
</tr>
<tr>
<td>departments</td>
<td>47,1%</td>
<td>32,8%</td>
<td>20,1%</td>
<td>100%</td>
</tr>
<tr>
<td>total</td>
<td>1,800</td>
<td>1,543</td>
<td>1,351</td>
<td>4,694</td>
</tr>
<tr>
<td></td>
<td>38,3%</td>
<td>32,9%</td>
<td>28,8%</td>
<td>100%</td>
</tr>
</tbody>
</table>

The table shows that the resulting collections of publications from universities are apparently not as complete as those from research institutes, at least for the last two years of investigation. As a consequence the sample is biased towards non-university research institutions. Since such institutions are quite often already affected by evaluations of their publication output and productivity, they probably attach greater importance to the completeness of their bibliographies than university departments.

4 Easy access means the availability of university bibliographies which make it possible to gather all (recorded) publications of a department in one process step. In a comparable study Nederhof et al. (1989) used annual reports of Dutch universities to assess their output. In Germany comparable research reports of universities are often published with large delays and not annually. The aim of our project was, however, to describe the publication behaviour as current as possible especially with regard to an assumed change towards an increasing use of digital publication outlets.
4 Results

There exist a number of studies that look at the communication practice in the social sciences (for an overview see Nederhof 2006). Most of them analyse the referencing behaviour and only a few focus on the publication behaviour of social scientists in general and educational researchers in particular. These studies conclude that journal articles are not as common for the communication of research results in the social sciences and humanities as they are in the natural sciences while monographs and edited books are still very important publication media, at least in the humanities and most fields of the social sciences. Is this description also true for German educational science?

4.1 Publication types

According to Clemens et al (1995: 433), the type of publication or genre (book or journal article) can be taken as a statement of identity: “What we write and where we publish may be taken as signals of who we are and how we think”. It can also say something about the character of a discipline. So, how does the distribution of genres in educational science look like?

In order to answer this question all collected publications were arranged into document types. Since the bibliographic information on the documents was, as described before, not standardized and very limited in many cases, only a quite simple and undifferentiated classification scheme consisting of four categories could be applied. The categories are: book, chapter in edited book, journal article and other (like working papers, reports etc., i.e. mainly grey literature). Figure 1 shows the distribution of these publication types in the analysed sample of documents.

Figure 1: Distribution of documents among document types

Very much unlike natural sciences book chapters are the most important publication type accounting for 46.7% of all documents. Journal articles account for 33.4%, books for 14.8% and other for 5.1%. This result corresponds to the distribution of publication types found in the evaluation of educational science at universities and colleges of education in Baden-Wuerttemberg (MFWK 2004: 33). A former study on German social sciences, using data from a bibliographic database, also points to the relative importance of chapters in edited volumes (Winterhager 1994: 545) and the prominence of this publication type seems to last to date. So, a restriction of evaluation efforts to journal articles would miss about two thirds of the publication output of German educational researchers, as it is represented in our sample.

Even though the German Education Index (GEI) covers all these types of documents this empirically found distribution of document types is not in line with the respective proportions in the GEI which supports the suspicion that its selection practice is unbalanced.

In our results it is furthermore noticeable that the publication culture regarding the used document types for publication is rather diverse among the institutions included in our survey. The share of journal articles, for instance, ranges from 22.1% to 44.5%. Depending on an institution’s scientific orientation and research topics the preferred channel for publishing results of the work seems to differ considerably.

---

5 For an exception that deals with the communication practice of German educational science see Keiner (1999)
6 Contrary to the approach of Nederhof et al. (1989) we could not differentiate between scientific and popularizing publications with our data. Such a distinction would surely also be beneficial for the assessment German educational science. The project tries to find characteristics of documents or journals on which an analogous distinction could be based.

7 The selection practice of this database leads to a disproportional coverage of document types with book chapters being relatively under-represented.
4.2 Importance of journals

As seen in section 4.1 the role of journals is less prominent in communicating research results in educational research than it is in many fields in the natural and life sciences. In the analysed time period of three years the included institutions published only one third of their total output in journals. The total number of 1,567 journal articles in our sample is distributed among 505 different journals. In 289 of these journals only a single contribution was published by the scholars of the included institutions within the three-year period of investigation. On the other hand the 30 journals most frequently used as publication outlets account for 39.4% of all articles. Comparable tendencies of a relatively high diffusion of publications over a wide range of journals and on the other hand only a low degree of concentration on certain journals are also found in the aforementioned evaluation report (MWFK 2004). On the basis of their results this study concluded that educational science addresses very heterogeneous audiences and is characterised by a huge diversity of research topics. A specific feature of the most frequently used journals in our sample is the relevance of ‘local’ journals. These are communication organs of some of the included institutions that publish exclusively or predominantly research results of their own staff. On the other hand there are only two journals (Zeitschrift für Erziehungswissenschaft and Zeitschrift für Paedagogik) which are used from at least about two thirds of the institutions as publication outlets. So a journals dispersion of authors and their institutional affiliations could be seen as an indicator of its relevance for the discipline. This idea was proposed by Holsapple (2008) who introduced the concept of ‘publication power’ in his approach to identify the most important information system (IS) journals. This concept is composed of two dimensions: ‘publishing intensity’ and ‘publishing breadth’. Assessing the importance of a journal only on its publishing intensity, i.e. the sum of the number of times the journal has been the publication outlet across a certain group of researchers, would be a too narrow view. If one also accounts for the ‘publishing breadth’ of a journal, i.e. the number of, in our case, institutions which have published at least one article in this journal, the relevance of some journals is put into perspective.

The use of publication power, which is the product of publishing intensity and publishing breadth, as an indicator of a journals importance results in a more realistic picture of the relevance of some journals in our sample. While the importance of the two most frequently used journals for publication becomes even more salient the relevance of some journals characterised by high publishing intensity but very low publishing breadth is decreasing.

In the context of bibliometric projects one question is of particular interest: how many of the journals in which German educational researchers publish are indexed in the databases of ISI? The two most frequently used journals in our sample are at the same time the only German educational science journals in these indexes (of which one, however, is only included since 2006 and with no impact factor yet). But beyond these two journals a considerable number of other ISI-indexed journals are used for publication. About 12% of all 505 journals are covered by the SSCI or SCI. Among these periodicals are German journals from neighbouring fields like educational psychology, but also a wide range of international journals, including e.g. Child Development, Journal of Curriculum Studies or Learning and Instruction. However, in about two thirds of these journals only one article was published by the scholars of the analysed institutions. Overall 13.5% of all journal articles and merely 4.5% of the total number of approx. 4,700 collected documents were published in the ISI-indexed journals. There are also data available on this aspect from Australian educational science. According to this data only 9.7% of the total publication output of Australian universities between 1999 and 2001 in the field of education was published in these universities. Since our object was only to describe the publication practice of exemplary institutions in the field we do not have comparable data to draw corresponding conclusions about the significance of individual journals. Hence, we only borrow the basic idea behind his approach.
4.3 Publication language

The research agenda of the social sciences is strongly influenced by the national context and oriented towards national audiences (cf. e.g. Hicks, 1999: 202). It can be assumed that this also applies to educational science, maybe even more because of the national or even (regional) character of educational systems. Can this national orientation thus also be found in our sample?

The main publication language of the selected documents in educational research is German. 88.1% of all publications were written in German, and 10.7% in English, i.e. the lion’s share of publications is directed to a national audience. The target groups of educational researchers include in addition to their peers also students, policy makers, professionals (teachers, social workers etc.), and the general public. The last-mentioned audiences, but often also the community of researchers, are normally addressed in the national language.

There were, however, rather big differences in the percentage of publications in English between the institutions. In some institutes which have a tendency towards (educational) psychology, like the Leibniz Institute for Science Education or the educational research section of the Max-Planck Institute for Human Development, the share of publications in English is greater than 15 percent, and in one exceptional case, the Knowledge Media Research Center, even above 50 percent.

On the other hand, there were more than a few university departments and also some of the institutes with less than 5 % publications in English. With regard to the publication language the field of educational science in Germany seems to be characterised by different publication cultures. Depending on the object of research and the used research methods some institutes show a tendency towards an internationalization of their scholarly communication while most institutions have a rather strong focus on the national context in terms of their preferred publication language.

With regard to the document type only small differences were found. The share of English documents is slightly higher for book chapters than for the other documents whereas books are written in German more often than the other publication types.

4.4 Publication length

Page length is a relatively easily extractable information if one has complete bibliographic data. It does not say anything about the content of a document, but Luwel et al. (1999) analysed the relation between page length and what the surveyed authors characterised as ‘substantial contributions’ and suggested that a minimum publication length could be introduced as a requirement to mark a document as substantial.

Since our data does not contain any explicit hint on the substantiability of a document we could only analyse the distribution of page length for each publication type in our sample.

Data on the page numbers were available for a relatively high proportion of documents in our sample so that this variable could be calculated. The page length distribution for each document type is shown in Table 2.

<table>
<thead>
<tr>
<th>Doc. type</th>
<th>Nr. Publ</th>
<th>SUM Pages</th>
<th>$Q_{25}$</th>
<th>Median</th>
<th>$Q_{75}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book</td>
<td>296</td>
<td>86,145</td>
<td>144.5</td>
<td>225.5</td>
<td>317.75</td>
</tr>
<tr>
<td>Other</td>
<td>59</td>
<td>2,812</td>
<td>15</td>
<td>23</td>
<td>71</td>
</tr>
<tr>
<td>Book chapter</td>
<td>1,815</td>
<td>27,182</td>
<td>6</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>Journal article</td>
<td>1,396</td>
<td>11,612</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
</tbody>
</table>

10 The page length was calculated as the difference between the ending and the starting page number, added up by one.

9 The remaining 1.2% are distributed among French and other languages.
Considering the median number of pages per type it is striking that the typical length of contributions to edited books (13 pages) corresponds to the results of Luwel et al. (1999: 161) for this publication type in the field of linguistics at Flemish universities. Journal articles, in contrast, seem to be considerably shorter in our sample with a median of six pages compared to 16 pages for refereed journal articles and 11 pages for non-refereed articles in Luwel et al.’s study. The attribute refereed or not could not be extracted from our data but, as described before, it was looked up which of the journals are indexed in ISI databases (and therefore are necessarily refereed). For articles in these journals the median length of 13 pages is considerably longer than that for other journal articles.

4.5 Co-authorships

In the natural and life sciences single-authored documents are rather unusual and the number of co-authors per paper is increasing. Is this trend also observable in German educational science? More than half of the publications in our sample (56.8%) were written by one author, another quarter (24.9%) by two and the remaining part by three or more. The average number of authors per publication is 1.79. In an older study by Keiner (1999) the average number of authors per article in three selected German educational research journals was just 1.2 for the time period from 1970 to 1990. So there seems to be a clear trend towards more collaboration, although it is still far from the situation in the natural and life sciences. Consistent with the results on Flemish linguistics by Luwel et al. (1999), institutes in which English plays a bigger role as publication language tend to show a higher average number of authors per document. For the three above mentioned institutes with relatively high proportions of English language publications the number is well beyond 2 whereas it is smaller than 1.4 for some of the institutions in which English plays only a minor role as publication language. But there are also some exceptions from this general picture.

5 Discussion and outlook

The analysis of the publication behaviour of German educational researchers has shown that the communication practice with regard to publication types, languages or collaborative authorships is relatively heterogeneous. While some institutes, at least partially, seem to follow the trend towards internationalization and multiple authors per publication, these tendencies are weaker for others. Beyond that, it became evident that in terms of numbers only a small part of the publication output is found in the international journals indexed by ISI. This corroborates the necessity to develop other evaluative criteria than the usually applied citation impact based on ISI data and raises the question how new indicators which are suitable for the described situation could look like. Consequently, the project intends to develop a new indicator for quality assessment based on alternative attributes of literature production and reception. This new multi-attributive indicator of scientific relevance refers to two different levels of a document (Botte 2007: 306):

- the measurement of well-defined characteristics of a publication itself as well as of the way it is used and
- the measurement of the characteristics of the editorial process it has undergone, i.e. to base the relevance assessment on the affiliation of a publication to a certain journal, publisher or editor.

The first way of measuring relevance aims at the assessment of the usage of documents. This measure of usage consists of the number a document indexed in the German Education Index is downloaded by users of this database (this attribute is labelled web-ranking). It takes into consideration that usage-based metrics are not unproblematic since they indicate strictly speaking merely interest, not user’s reception and they are susceptible to manipulation. To what extent this download impact correlates with citation impact in the field of educational science is an open question (for studies on usage-based indicators see e.g. Chu and Krichel 2007; Shepherd 2007; Duy and Vaughan 2006; Mayr, 2006). In addition to the usage of a document, its characteristics such as content, target group or purpose will be used as signals
of relevance (document genre). There seems to be an increasing interest in the usage of document genres for the improvement of the retrieval of scientific documents (cf. e.g. Montesi/Mackenzie Owen 2008) but, up to now, hardly for the assessment of their relevance.

The second approach to rate scientific publications is based on the reliability of the source of a document respectively on the evaluation of the editorial process the document passed through. A key element in this approach is the rating of journals and publishers and their reputation by experts in the field (peer rating). The pros and cons of a journal rating by a limited number of selected experts or a survey of scholars are discussed by Nederhof et al. (2001: 261-262). Surveys in which scholars rate the journals of their discipline are certainly the most widely used alternative to impact quality measures.

Also on the level of the source of a document, quality will be assessed by evaluating the compliance with basic publishing standards, such as timeliness of publication and English-language bibliographic information, the existence of a peer review process and certain other characteristics like the internationality of the editorial board of a journal or serial (editorial quality). Some related approaches already exist even for the field of educational research like that of Lu (2004) who mapped the editorial policies of over 600 education journals or that of Fernández-Cano and Bueno (2002) who suggested four evaluative modalities for the assessment of Spanish educational research journals: reputation survey data, adjustment to edition norms, adjustment to scientometric patterns and impact citation data.

From these and other studies (like Forgionne et al. 2002; Lopez-Cozar and Perez 1995; Martin-Sempere et al. 2000; Miller and Serzan 1984; Wellington and Torgerson 2005) feasible variables to rate journals and serials are extracted. The future goal is to establish a database that contains the information on editorial quality for educational science journals. Such a database could also support prospective authors in their decision where to submit a manuscript. Data sources for the collection of the necessary information will be other databases like Ulrich’s but also interviews with journal editors (cf. e.g. Schloegl and Petschnig 2005).

The combination of the above-mentioned methods of assessing the research relevance of publications shall constitute an alternative to citation impact indicators that is appropriate to German educational science and transferable to other social sciences. The idea of this new multi-factor approach is to balance the weak points of each individual attribute. To assess the validity of this new indicator its relation with citation-based quality indicators shall be analysed. To this end it is considered to use data from Google Scholar since this data is not restricted to the small fraction of academic literature covered by ISI (Harzing and van der Wal 2007). Because of the problems of Google Scholar concerning reliability, coverage etc. (cf. Jacso 2008) this assessment can only be exploratory however.

The bottom line of this effort should then be a weight calculated from the four described attributes which will be attached to every document in the German Education Index.

References


Berghoff et al. (2006), Das CHE Forschungs-Ranking deutscher Universitäten 2006. Arbeitspapier Nr. 79. Centrum für Hochschulentwicklung.

http://www.che.de/downloads/CHE_ForschungsRanking_2006.pdf [01.04.2008]


A continuing collection of publications relevant to the project can be found at:

http://www.bibsonomy.org/user/wdees


Hicks, D. (1999), 'The difficulty of achieving full coverage of international social science literature and the bibliometric consequences', Scientometrics 44(2), 193-215.


MWFK, Ministerium für Wissenschaft Forschung und Kunst Baden-Württemberg,


