Science-Metrix

Final Report

August 2004

The Use of Bibliometrics in the Social Sciences and Humanities

Prepared for the

Social Sciences and Humanities Research Council of Canada (SSHRCC)

Science-Metrix

Final Report

The Use of Bibliometrics in the Social Sciences and Humanities

Éric Archambault and
Étienne Vignola Gagné

August 2004

Prepared for the
Social Sciences and Humanities Research
Council of Canada (SSHRCC)

Science-Metrix specializes in the measurement and evaluation of science, technology and innovation. Our data collection and assessment methods include bibliometrics, scientometrics, technometrics, surveys and interviews, environmnetal scans, monitoring and intelligence gathering. We perform program and policy evaluations, benchmarking and sector analyses, market studies and strategic planning. Science-Metrix has a robust knowledge of life and environmental sciences.

514.495.6505 • 4572 De Lorimier Avenue • Montreal • Quebec • Canada • H2H 2B5
info@science-metrix.com • www.science-metrix.com

Summary

The Social Sciences and Humanities Research Council (SSHRC) asked Science-Metrix to identify current practices in bibliometric evaluation of research in the social sciences and humanities (SSH). The resulting study involves a critical review of the literature in order to identify the specific characteristics of the SSH and their effects on the use of bibliometrics for evaluating and mapping research. In addition, this report presents an overview of methods of research benchmarking and mapping and identification of emerging SSH fields. This part of the report is particularly relevant because of the need to exercise considerable caution when using bibliometrics to evaluate and map SSH research.

This report shows that bibliometrics must be used with care and caution in a number of SSH disciplines. Knowledge dissemination media in the SSH are different from those in the natural sciences and engineering (NSE), particularly because of the much greater role of books in the SSH. Articles account for 45% to 70% of research output in the social sciences and for 20% to 35% in the humanities, depending on the discipline. Bibliometric analyses that focus solely on research published in journals may not give an accurate representation of SSH research output.

In addition, bibliometric analyses reflect the biases of the databases used. For example, the Social Science Citation Index (SSCI) and the Arts and Humanities Citation Index (AHCI) of Thomson ISI over-represent research output published in English. Original findings produced by this study show that the bias results in an estimated 20-25% over-representation of English material in the two databases. Findings from the scientific literature support those of Science-Metrix.

In order to benchmark national performances and identify Canada's strengths in SSH, it is possible to use research articles published in journals representing disciplines where this medium of communication is popular, such as economics. For other disciplines, journal-based bibliometric analysis may be used with due caution and databases can be built in order to factor in other knowledge dissemination media. However, one must be wary of conducting comparative analyses of SSH disciplines without taking into account the effects of the knowledge dissemination media of each discipline on the bibliometric tools being used.

Bibliometric methods have not yet been refined to the point where they can serve to identify emerging fields. In this regard, the methods with the greatest potential are co-citation analysis, co-word analysis and bibliographic coupling. However, their usefulness for policy development has been challenged. It is therefore preferable to combine bibliometrics with research monitoring and even peer review for identifying emerging fields. Another approach is to track the development of bibliometric methods, which nonetheless show promise on many fronts.

In short, bibliometrics must be used carefully for SSH research evaluation. Furthermore, each discipline has its own specific characteristics, so bibliometrics is to be applied differently in each case. This report presents original findings that will help in determining how bibliometric analysis should be applied to the various SSH disciplines.

It is possible to adopt at least two possible attitudes toward the challenge of offsetting the limitations of bibliometrics: a passive one (laissez faire) or a proactive one (interventionism). Given current trends such as the increased publication of articles and open access, the laissez faire approach may be the most effective way of enhancing the validity of SSH bibliometric analysis. The interventionist approach focuses on creating and optimizing databases such as the Common CV System.

Acknowledgements

Science-Metrix would like to thank the Observatoire des sciences et des technologies (OST) and, in particular, Vincent Larivière for his generous contribution in the analysis of citations made to journal papers among various fields of the social sciences and humanities. We would like to thank Yves Gingras of OST for his critical comments on a draft version of this report. We would also like to thank Malcolm Williams for the English translation of this report. Finally, we would like to thank SSHRC staff for their support during the course of this project. In particular, we would like to thank Christian Sylvain for his comments on a previous version of this report and we would like to thank Sylvie Paquette for her support, constructive comments and encouragement during this project.

Contents

Figures

Tables

1 Introduction

Bibliometrics and scientometrics are a set of methods for measuring the production and dissemination of scientific knowledge. Derek de Solla Price and Vasilij Vasilevich Nalimov were the originators of the discipline, which they developed for the purpose of providing research tools to historians and sociologists of science.

However, it was only with the advent of the tools developed by the Institute for Scientific Information (now Thomson ISI) and the research conducted by its founder, Eugene Garfield, that the use of bibliometrics became widespread. With their systematic archiving of articles from a selection of some of the most prestigious and most often cited scholarly journals, the Thomson ISI databases considerably reduce the effort required to carry out bibliometric analysis. The field grew out the sociology of science, information science and library science, but it quickly carved out a place for itself in quantitative research evaluation.

Whether used for historical or sociological research or for evaluation in a science policy context, bibliometrics is used most of all in the natural sciences and engineering, because these fields have appropriate bibliographical characteristics for bibliometric analysis as currently practised. The same cannot be said for the social sciences and humanities. In these fields, research results are disseminated through a much broader range of media than in the NSE. Furthermore, SSH research often focuses more on local issues, with the result that its cognitive structure is sometimes quite fragmented. Thus the application of bibliometrics to SSH faces specific problems. The purpose of this report is to identify the limitations of bibliometrics in the context of SSH research and then propose methods to offset those limitations, particularly with a view to benchmarking research output in a variety of countries and identifying emerging research fields.

The report content is as follows:

- Section 2 presents the basic principles and main methods and tools of bibliometrics;
- Section 3 presents the limitations of bibliometrics, particularly in terms of its application to SSH;
- Section 4 presents and proposes methods for benchmarking national performances in SSH;
- Section 5 presents and proposes methods of identifying emerging research fields;
- Section 6 presents conclusions and recommendations on the use of bibliometrics in the SSH and the development of tools to offset the limitations referred to in the report.

2 Bibliometrics

Many articles have been written on the use of bibliometrics in the social sciences and humanities. Before presenting the various viewpoints on the strengths and limitations of bibliometric tools for evaluation, it is necessary to define bibliometrics and outline its main indicators. Bibliometrics is made up of methods for conducting quantitative analysis of science. Some of the methods serve to measure sociological aspects of one of the researcher's most important activities — dissemination of research results in published form. Bibliometrics is based on two assumptions: (1) the goal of researchers is to advance knowledge, and this means disseminating the results of their research and studies through a variety of communication media, including writing, which lies at the core of the academic tradition; (2) scholars have to publish in order to build a reputation and advance their careers.

A publication count is one means of measuring and comparing the production of various aggregates such as institutions, regions and countries. It can also be used to evaluate output in individual disciplines, such as philosophy and economics, and to track trends in research fields, collaborative research and many other aspects of research output.

Bibliometrics uses three main types of indicator:

- **Publication count**

The number of articles published in learned journals during a specific time frame is an indicator of the output of a set or subset within the science system. It is also possible to compare numbers in order to gauge output intensity in specific fields (specialization index).

- **Citations and impact factor**

Number of citations can be used to evaluate the scientific impact of research. The number of citations received by learned journals is systematically compiled by Thomson ISI and sold under the trademark *Journal Citation Reports* (JCI). This product includes a number of indicators related to citations received by journals, and the *impact factor* is probably the one most commonly applied.

- **Co-citation and co-word analysis**

Many co-citation-based indicators are used to map research activity: co-citation analysis, co-word analysis, and bibliographic coupling. Mapping is a means of studying the development of emerging fields using time as a variable. Co-citation and co-word indicators can be combined with publication and citation counts to build multifaceted representations of research fields, linkages among them, and the actors who are shaping them.

2.1 Use of databases in bibliometrics

While bibliometric analyses can be conducted on manually compiled data, they are usually based on databases that in many cases exist primarily for bibliographic purposes. A number of databases are available in all scientific fields: Medline, Sociological Abstracts and Francis are just a few in a long list. Databases are essential because they contain data that otherwise would have to be compiled manually. Note, however, that they have been built more for finding articles than for performing complex mathematical counts. In other words, they were designed for bibliographic rather than bibliometric purposes. Therefore, the first step in any bibliometrics project is to "condition" bibliographic data in order to build bibliometric databases. This essentially involves standardizing data — mainly data pertaining to authors' addresses — to facilitate counting.

The most commonly used databases in bibliometric analyses are the ones produced by Thomson ISI. Established scholars select the journals to be covered in cooperation with users, publishers and members of editorial boards. Selection criteria include publication frequency, compliance with international presentation conventions, existence of an English abstract, and a peer review committee. According to Nederhof and Zwaan (1991), however, some of the journals do not meet the peer review committee criterion.

In addition to using the Web of Knowledge (WoK), a meta-database available on the Internet, most bibliometric studies are based on one or more of the Thomson ISI databases: the Science Citation Index (SCI), the Social Science Citation Index (SSCI), and the Arts and Humanities Citation Index (AHCI). They are particularly useful for a number of reasons (see Katz and Hicks 1998, among others):

- Their primary advantage is coverage. They cover all research fields, giving ready access to aggregated data. Unlike a number of other databases, which may cover some journals only in part (depending on the relevance of an article to single-discipline databases, for example), these databases systematically index all articles and other items in the selected journals. Nearly 10,000 existing learned journals are covered (Katz and Hicks 1998).
- The criterion for including a journal in SCI and SSCI is the number of citations it receives. Because citations are perceived to be an indicator of interest and recognition on the part of scholars, the citation count is considered as evidence of the usefulness, quality and/or impact of a journal. According to Eugene Garfield, 90–95% of the most frequently cited articles in the natural sciences are published in a core group of about 2000 journals (CNER 2002, citing Garfield 1996). AHCI inclusion criteria are more subjective. According to Thomson ISI,

  > Publishing standards, including timeliness, are also important in the evaluation of Arts and Humanities journals. Citations in the Arts and Humanities, however, do not necessarily follow this same predictable pattern as citations to Social Sciences and Natural Sciences articles. Citations to an article on the 19th Century Romantic novel, for example, may accrue slowly at first, and then slacken, fluctuating over time in cycles consistent with scholars' varying interest in the topic. In addition, Arts and Humanities journal articles reference non-journal sources heavily (e.g., books, musical compositions, works of art and literature). Consequently, citation data, while sometimes useful, are frequently much less so in journal evaluations in the Arts and Humanities.
  >
  > Arts & Humanities journals are selected by the primary editor with the support of the subject experts from the ISI Arts & Humanities indexing staff. The goal is the collection of superb Arts and Humanities content that reflects the complex nature of cultural phenomena across a broad range of fields. http://www.isinet.com/essays/selectionofmaterialforcoverage/199701.html/

- The Thomson ISI databases contain the institutional addresses for all authors of a given article. Other databases usually contain only the address of the first-named author. Without a complete address list, collaborative research analysis is impossible.

- Only the Thomson ISI databases contain citation information, which makes it possible to measure research impact. In the opinion of Katz and Hicks, this attribute alone justifies using these databases as science policy and research management tools.

At the same time, the Thomson ISI databases have some disadvantages. First, their coverage is not as deep as some of the specialized databases. Second, they are relatively costly to use because they are produced by a private company (Katz and Hicks 1998). Third, they are not as suitable for SSH research as they are for research in the natural sciences. This third disadvantage will be examined in Section 2.

2.2 SSH bibliographic databases

Table I presents information on some databases of potential use for SSH evaluation. Appendix I gives a more exhaustive list of such databases, including ones that are smaller than those in the Table.

Operational since 1972, Francis contains about 2.5 million SSH bibliographic references and is particularly useful in that it covers 4335 journals in addition to monographs and theses. It serves as a very helpful complement to other databases, which do not cover monographs and do not always provide adequate coverage for articles from outside the Anglo-Saxon world. Unfortunately, only one author address per publication was included in Francis up to the year 2000, but since then it has included all addresses. It follows that, up to the year 2000, it could be used only for a few core bibliometric indicators. Furthermore, it is not a very stable database in terms of selection criteria for journals and other bibliographic sources, and the criteria are not properly documented. This explains why even the Observatoire des sciences et des techniques (OST) in France does not use Francis to produce bibliometric statistics.

Table I Databases with the greatest potential for SSH bibliometrics

Database	Field	No. of journals	Years covered	Types of literature	Update frequency	Author address(es)
America: History and Life	History and Culture[1]	2,100	1964-	A,B,C	Quarterly	None
ABELL Online	Literature, Language and Culture	800	1920-	A, B, C, F	Quarterly	None
ABI/INFORM Global	Business and Management	1,600	1971-	A, D, F	n.a.	None
AHCI	Arts and Humanities	1,150	1975-	A	Weekly	All
CSA Worldwide Political Science Abstracts	Political Science	1,432	1975-	A	Monthly	First author
Econlit	Economics	500	1969-	A, B, C, F	Quarterly	All
Francis	Humanities and Social Sciences	4,335	1984-	A, B, C, F	Monthly	First author
Historical Abstracts	History[2]	2,100	1955-	A, B	Monthly	None
International Bibliography of the Social Sciences	Economics, Political Science, Sociology and Anthropology	2,600	1951-	A, B	Weekly	None
PsycINFO	Psychology	1,932	1872-	A, B	Monthly	First author
Sociological Abstracts	Sociology	2,000	1963-	A, B, C, F	Bi-monthly	First author
SSCI	Social Sciences	1,700	1956-	A	Weekly	All
Wilson Humanities Abstracts	Humanities in general	400	1984-	A, B	Monthly	None
Wilson Social Sciences Abstracts	Social Sciences in general	420	1983-	A, B	Monthly	None

Source: Compiled by Science-Metrix
1. History of Canada and U.S. only; 2. History, excluding Canada and U.S.
A. Articles in learned journals; B. Monographs; C. Doctoral theses; D. Articles in mass-circulation newspapers; F. other

Worldwide Political Science Abstracts is another interesting database. It has started to archive article-to-article citations, in the same way that Thomson ISI does. However, the citations are from only a few articles, so the database cannot yet be used for citation analysis.

Apart from the Thomson ISI databases, only Econlit lists the addresses of all authors. Thus it is the only one, apart from AHCI, SSCI and Francis post-2000, that can be used to study international collaborative research and benchmark national performance with any degree of refinement.

All but one of the databases in Table I contain complete lists of authors. The exception is PsycINFO, which includes only the first-named author. On the other hand, most of them do not list author addresses or list only that of the first. Since there is basically no point in doing bibliometrics if you cannot assign articles to institutions or places, this limitation automatically eliminates most of these databases as tools.

Some databases may be of interest because they are multidisciplinary and thus cover fields outside the traditional breakdown of disciplines. They include the International Bibliography of Social Sciences, Francis, Wilson Humanities Abstracts and Wilson Social Sciences Abstracts, along with the Thomson ISI databases.

In general, these databases may be an interesting alternative to the conventional Thomson ISI databases, but they cannot be used to conduct bibliometric analyses of the same breadth or depth.

2.3 Open access and impact of ICT developments on SSH bibliometric evaluation

Fairly recent developments in information and communication technologies (ICTs) can influence SSH bibliometric evaluation methods.The impact of two aspects will be examined here: methods for measuring Web-based research dynamics and the trend toward open access to learned journals.

2.3.1 Methods for measuring Web-based research dynamics (webometrics)

In the wake of Internet/Web developments, some bibliometricians drew analogies between Web-based and research documents and came up with the idea that the scientific content of the Web could be analysed in the same way as the science journal system. Hyperlinks between websites were compared to article-to-article citations, and this spawned a number of exploratory studies on ascertaining whether conventional bibliometric methods could be applied in whole or in part to the Web (Almind and Ingwersen 1997; Ingwersen 1998; Larson 1996; Rousseau 1997). Some methods developed in this context are outlined below.

It is possible to measure a Web impact factor (WIF). For example, the number of hyperlinks attributed to a field can be divided by the number of pages in a website (Bjorneborn and Ingwersen 2001; Li 2003). A more interesting alternative is to measure the number of relevant hyperlinks to an academic institution and then divide that number by the number of scholars affiliated with it (Li 2003 citing Smith and Thelwall 2001; Li, Thelwall, Musgrove and Wilkinson 2003; Thelwall 2001). The resulting departmental WIFs parallel the rankings established by the Research Assessment Exercise (RAE) in the U.K. (Li, Thelwall, Musgrove and Wilkinson 2003; Thelwall and Harries 2004). By determining WIFs, it is possible to measure universities' prestige rather than the quality of their research work, just as citation analysis measures impact rather than quality (Li 2003).

It is possible to build clusters of sites based on the degree of proximity generated by hyperlink analysis. The results are similar to those produced by co-citation and co-word analysis. As in the case of bibliometrics, website measurement could serve to highlight specific linkages within the field or science system. For example, Thelwall and Wilkinson (2004) have tried to apply the bibliographic methods of bibliographic coupling and "co-link" (instead of co-citation) analysis to the Web. The results do not yield any firm conclusions, but they do point in a new direction for testing the emerging field of webometrics.

Notwithstanding all their potential, hyperlinks suffer from one major drawback compared with citations: they are not related to a time variable. There is always a date associated with citations, which is not the case with hyperlinks (Prime, Zitt and Bassecoulard 2002 citing Egghe 2000). The time variable is very useful for citation analysis and can produce very good results; hyperlink analysis may not be as informative

A second, more practical obstacle to the development of website measurements resides in the many technical problems (Li 2003), particularly in the context of collecting and cleaning up data.

In addition, some authors say that hyperlinks should not be expected to replace citations (Prime, Zitt and Bassecoulard 2002; van Raan 2001). Their main rationale for this is that hyperlinks will not generate reliable evaluations of research impact (Egghe 2000).

In short, the present state of webometrics methods is such that that they cannot really be used for SSH research evaluation. At this stage, their application to mapping seems more reliable. Developments in webometrics are worth tracking, however, because they may lead to some innovative research evaluation methods. In particular, the work of the Statistical Cybermetrics Research Group at Wolverhampton University in the U.K. looks very promising.

2.3.2 Open access

Universities cannot give their scholars access to every academic journal in the world. Subscriptions are quite expensive, meaning research libraries have to make a selection. It is fair to assume that the science system as a whole would perform much more effectively if researchers had access to all research. The proponents of open access have developed a number of systems to provide it. The idea is to give scholars free access to a specific set of publications and eventually extend that access to all research output.

The idea has caught the attention of many scholars, as shown by the number of editorials on it in learned journals. It also deserves to be considered by decision makers because it could have a major effect on the science system in its present state. The European Commission recently launched a study on open access and optimization of the research publication system as part of its efforts to create a European Research Area and raise the international profile of European research (CORDIS 2004).

A variety of systems

Open access can take a number of forms. Under a recent initiative of the Public Library of Science (PLoS), researchers have to use their own research funds to publish in the academic journal *PLoS Biology*, which is available free of charge. A similar model had already been launched by the *Journal of Clinical Investigation* (JCI), which has been available online free of charge since 1996 (Varki 2003). Harnad et al. (2004a) refer to free publications operating on this principle as “gold” journals. Another approach is for journals to continue to charge for subscriptions but to authorize posting of the article on the Internet. Harnad et al. (2004a) refer to these as “green” journals. This approach depends to some extent on the honesty and good will of research libraries, which will be relied on to continue subscribing to academic journals even though the articles are posted.

A method that seems to hold even more promise is self-archiving, or the downloading of articles into institutional archives, which may be bound by the Open Archive Initiative (OAI) protocol (Harnad et al. 2004a). Specific research engines can scan the archives as if they formed a single archive. In addition, the archives can index citations or downstream downloads and thus add other functions that are particularly useful for bibliometrics. Reportedly, more than 100 universities now have their own archives (Harnad et al. 2004b).

Current situation and impact

A relatively recent concept, open access has quickly created an impact and gained currency. According to Harnad et al. (2004a), 80% of academic journals are "green": they authorize authors to archive their articles subject to restrictions varying in number and scope. Another 5% of journals are "gold": access to them is complete and free of charge (*PLoS Biology* is an example). Science-Metrix used the Directory of Open Access Journals (DOAJ) to ascertain the number of open-access ("gold") journals in the SSH and found 529, compared with 436 in the NSE.

What impact does open access exert on the use of research? There is a divergence of findings and opinions on this question. A number of studies have compared the number of citations attributed to open-access articles with the number attributed to non-open-access articles. Lawrence (2001) covered 119,924 conference papers in computer science produced from 1989 to 2000. The mean number of citations attributed to articles not online was 2.74, compared with 7.03 for online articles. Harnad et al. (2004a) report that Kurtz (2004) and Odlyzko (2002) came up with similar estimates for astronomy and mathematics. A consortium made up of the OST, Southampton University and Oldenburg University has undertaken a study on the subject, and preliminary findings show that physics articles available free of charge on the ArXiv online archive are cited 2.5 to 5 times more frequently (Brody et al. 2004; Harnad et al. 2004a). Thomson ISI has examined the issue too and has produced its own statistics (Testa and McVeigh 2004). Testa and McVeigh looked at the impact factors of academic journals in the NSE and found that the factor for the 148 open-access journals covered was generally comparable to that for subscription-based journals. The authors conclude that open-access journals enjoy only a slight advantage over their priced counterparts. Thus their findings militate against the more positive ones outlined earlier. Thomson ISI does point out, however, that its results are preliminary. It is interesting to note that the company gives a good performance rating to *PLoS Biology* for its first four issues.

It will be important to track the development of the open-access approach. For now, its use is not widespread and its impacts are unclear, but it could have a major effect on research publishing. The ways in which it could affect the application of bibliometrics are presented in Section 6.3.

3 Limitations of bibliometrics

Bibliometric methods are very useful for measuring the dissemination of knowledge in the natural sciences, but they are less effective in some applied fields, such as engineering (van Raan 2003). Their use in SSH evaluation poses specific problems; some solutions are available, but they need to be applied with considerable care. According to the research, this is due to structural differences in knowledge production and dissemination in the various fields. A number of scholars have highlighted fundamental differences between the scientific communication practices of scholars in the NSE and those in the SSH (Glänzel and Schoepflin 1999; Hicks 1999 and 2004; van Raan 2003). These structural differences explain why the usefulness of bibliometric methods varies with the SSH discipline.

We will now outline the overall limitations of bibliometrics and look in greater detail at the limitations derived from the specific characteristics of SSH disciplines.

3.1 Limitations of bibliometrics

Specific problems affect the validity of bibliometric evaluation in both the NSE and the SSH. According to a French organization, the Conseil national d'évaluation de la recherche (CNER 2002), the main problems with research publication databases are as follows:

- limited coverage
- exclusion of certain types of documents
- classification of journals by discipline
- changes in journal titles
- names spelled the same way
- number of authors (and distribution of work)
- excessive, selective, secondary, negative and erroneous citations, self-citation and personal strategies

Since bibliometrics is generally based on use of bibliographic databases, all the above factors reduce the reliability of bibliometric research evaluation. The problems are common to all the databases and reflect (1) the limitations of compiling and indexing academic journals and articles and (2) practices and situations specific to the field concerned.

3.2 Bibliometrics and SSH

Clearly, it is difficult to perform bibliometric analysis in the SSH with a single, all-purpose method because of the broad range of disciplines. Each discipline has its own specific characteristics and practices, which exert different effects on conditions for applying bibliometrics. Accordingly, the reader should always keep in mind that the general findings presented here may not obtain in certain disciplines or areas of specialization.

The SSH differ from the NSE not only in terms of their research subjects but also because of their scientific communication practices and methods (Moed, Luwel and Nederhof 2002). Each difference affects conditions for applying bibliometrics.

First, knowledge dissemination media and, by extension, communication media in general are more varied in the SSH than in the NSE. This is reflected in the greater role played by monographs, conference papers and proceedings, and non-scientific literature in the SSH. Depending on the discipline, articles may be a relatively minor publishing medium compared with others, such as books. Unfortunately, no database covers these other forms of publication as systematically and exhaustively as Thomson ISI does for journal articles.

Second, SSH research subjects are sometimes more local in orientation and, as a result, the target readership is more often limited to a country or region (Glänzel 1996; Hicks 1999; Hicks 2004; Ingwersen 1997; Nederhof et al. 1989; Nederhof and Zwaan 1991; Webster 1998; Winclawska 1996). SSH scholars reportedly publish more often in their mother tongue and in journals with a more limited distribution (Gingras 1984; Line 1999).

Furthermore, according to Hicks (1999), a number of SSH disciplines have more paradigms competing with one another than do those in the NSE, and as a result SSH literature is more fragmented — a situation that hinders the formation of a solid "core" of scientific journals (Hicks 1999), thereby making article-based bibliometric analysis more difficult to conduct successfully. The main differences between the SSH and NSE concerning the use of bibliographic methods are analysed in greater detail below.

3.2.1 SSH communication media

Scientific communications indexed as bibliographic entries lie at the core of bibliometrics. For example, scientometrics can be used to analyse verbal communications as long as they are systematically indexed. Accordingly, while scientometric methods do not target any one form of communication in particular, bibliometrics is used primarily for analysing academic articles. The bibliometrics/article connection goes back many years. Bibliometrics really began to take off with the advent of SCI in the 1960s. In fact, it is now hard to dissociate bibliometrics from journal article measurement, yet it has to be done because the learned journal is not as pervasive in some SSH disciplines as it is in the NSE.

What reasons can be adduced for the broad range of knowledge dissemination media in the SSH compared with the NSE? According to Line (1999), most social sciences can be described as "relatively young, and scarcely organized as coherent disciplines." He suggests that the SSH are fragmented because they do not have international standards for rigorously defining concepts. Terms used vary markedly between regions and over time. The publication of articles in learned journals is seen as an indication of consensus and as a practice that brings scholars in a discipline together. Some researchers claim that SSH scholars belong to a large number of competing paradigms, which creates an impetus to publish in book form (Hicks 1999 citing Pierce 1987). Since books are not subject to peer review to the same extent as journal articles, SSH scholars see book publishing potentially as a means of taking a position outside the prevailing paradigms. According to Hicks (1999), the best NSE research is published in articles, but the best SSH research may well be published in books.

Hicks bases these conclusions on an article by Clemens et al. (1995), who combined the arguments of Merton and Kuhn to establish a connection between epistemological dynamics and the publication of books and articles. Clemens et al. advance the hypothesis that the prevalence of articles in the natural sciences reflects a publishing system operating as a social control vehicle and thereby generating original, reliable, cumulative knowledge. In its turn, this system reflects the paradigmatic status of the individual discipline as well as a consensus on questions and methods specific to it. Thus a discipline in which the book is a major publishing medium is probably still in its pre-paradigmatic phase.

Clemens et al. attribute different scientific roles to books and articles, at least with respect to the sociology of science. Books have greater impact outside the discipline and are generally read by more people. Articles cover more recent problems and are a better way of achieving standing as an authority within a discipline. Preference for a given publishing medium over another may depend on the scholar's gender, rank, institution and academic background.

Note that the risk of having one's idea or invention stolen is much lower in the SSH, so scholars have more time to develop and flesh out their arguments before publishing them (Hicks 1999).

Articles and books are the two most commonly used scientific publication and communication media, but several other publishing media play an important role in some disciplines. They include official (government) publications and grey literature — organizations' internal reports that are usually distributed more informally and less systematically than commercially available journals and periodicals. Their value for knowledge production is gaining increasing currency (Grayson and Gomersall 2003; Larivière and Godin 2001). Some of these reports get ISBN numbers. Moreover, they are enjoying increasingly wide distribution, mainly because of the Internet. Working papers posted on the websites of some research centres are playing a growing role because they are sometimes more exhaustive and provide more data than the final versions published as articles. Note, however, that the purpose of some official and grey literature is to convince and inform rather than contribute to the scientific dynamics specific to peer-reviewed journals. In short, caution must be the watchword here, because this type of literature should be considered at least in part as an SSH dissemination medium outside the academic community instead of a scientific communication medium.

Books, articles and, to a lesser degree, grey literature and government reports are all considered to be scientific communication media. As publications, they are a response to factors (incentives, behaviours, criteria) within the disciplines concerned. Yet scholars are not writing only for their colleagues. They are also publishing articles and books for the purpose of disseminating scientific knowledge within the community at large. While most bibliometric analyses focus exclusively on scientific literature, it may be worthwhile measuring activities designed to disseminate knowledge to the general public.

A number of authors claim that publication media other than journal articles play a significant role in SSH literature, but the fact is that there is no precise information on the prevalence of the various publication media. Nederhof et al. (1989) studied the distribution of six scientific communication media in eight SSH disciplines in the Netherlands. In all cases, the article was the primary medium, accounting for 35–57% of publications. Next came book chapters, with 21–34%. Even though these findings may challenge certain established views on the pre-eminence of monographs in the SSH, they should not lead to the conclusion that evaluations based only on articles will suffice. The statistics given by Nederhof et al. also show that, while 35% of the publications in general linguistics are articles, 34% are book chapters. In short, monographs play a significant role in knowledge dissemination.

Their findings cannot, however, be applied to all the social sciences and humanities in all countries. For example, Nederhof et al. note that the article percentage may be exceptionally high because of the specific characteristics of the scientific field in the Netherlands, where there is an unusually high number of locally published learned scientific journals. It is therefore important to consider statistics from other countries. In Denmark, Andersen (2000) conducted a survey suggesting that only one quarter of the publications by Danish scholars in the social sciences were journal articles. Hicks (1999) estimates that books make up 40–60% of the literature in the social sciences and that they have a very high research impact because they account for 40% of citations.

Note also that the importance scholars place on mass distribution varies widely from discipline to discipline. Nederhof et al (1989) examined the proportion of SSH enlightenment publications for the period 1980–1985. In experimental psychology, Dutch scholars devote 3–12% of their publishing work to enlightenment of a lay readership, and in Dutch literature the figure is 30–43%.

In order to shed light on the importance of journal articles, the Observatoire des sciences et des technologies (OST) produced a set of statistics that make it possible to gauge the importance of SSH communication media other than articles. Figure 1 provides statistics on citations from journal articles covered in the SCI, SSCI and AHCI databases. Its shows that, whereas 85% of citations in the natural sciences are to journal articles, less than 50% of SSH citations are to this publication medium.

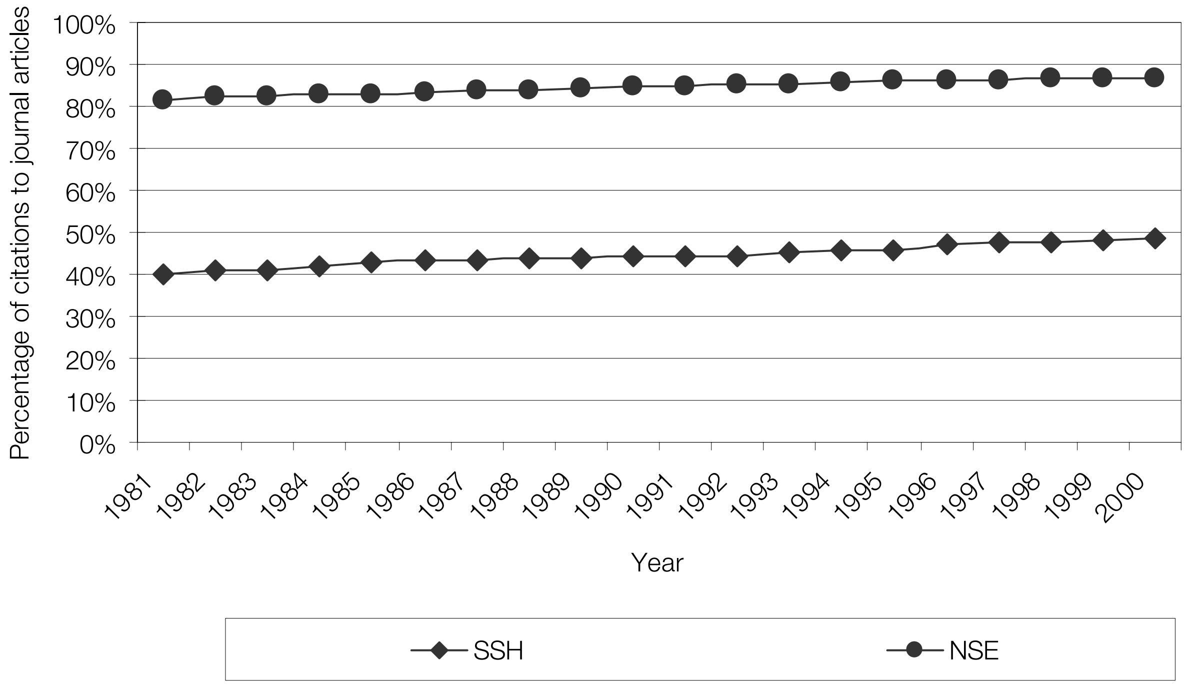

Figure 1 Percentage of citations to journal articles in science
Source: Larivière, Archambault, Gingras and Vignola-Gagné 2004

Admittedly, the SSH are changing quickly, for the number of citations to articles has risen almost 22% over the last 20 years, compared with 6% in the natural sciences. This does not mean, however, that the trend is happening everywhere. Over the last 10 years, for example, economics and administration, psychology and “other” social sciences are the SSH fields where the role of journals in knowledge production has grown at a rate similar to that of the SSH as a whole (Figure 2). On the other hand, the role of journals in advancing scientific knowledge is diminishing in history, literature and the other humanities. In other words, SSH disciplines do not all progress at the same rate and in the same way.

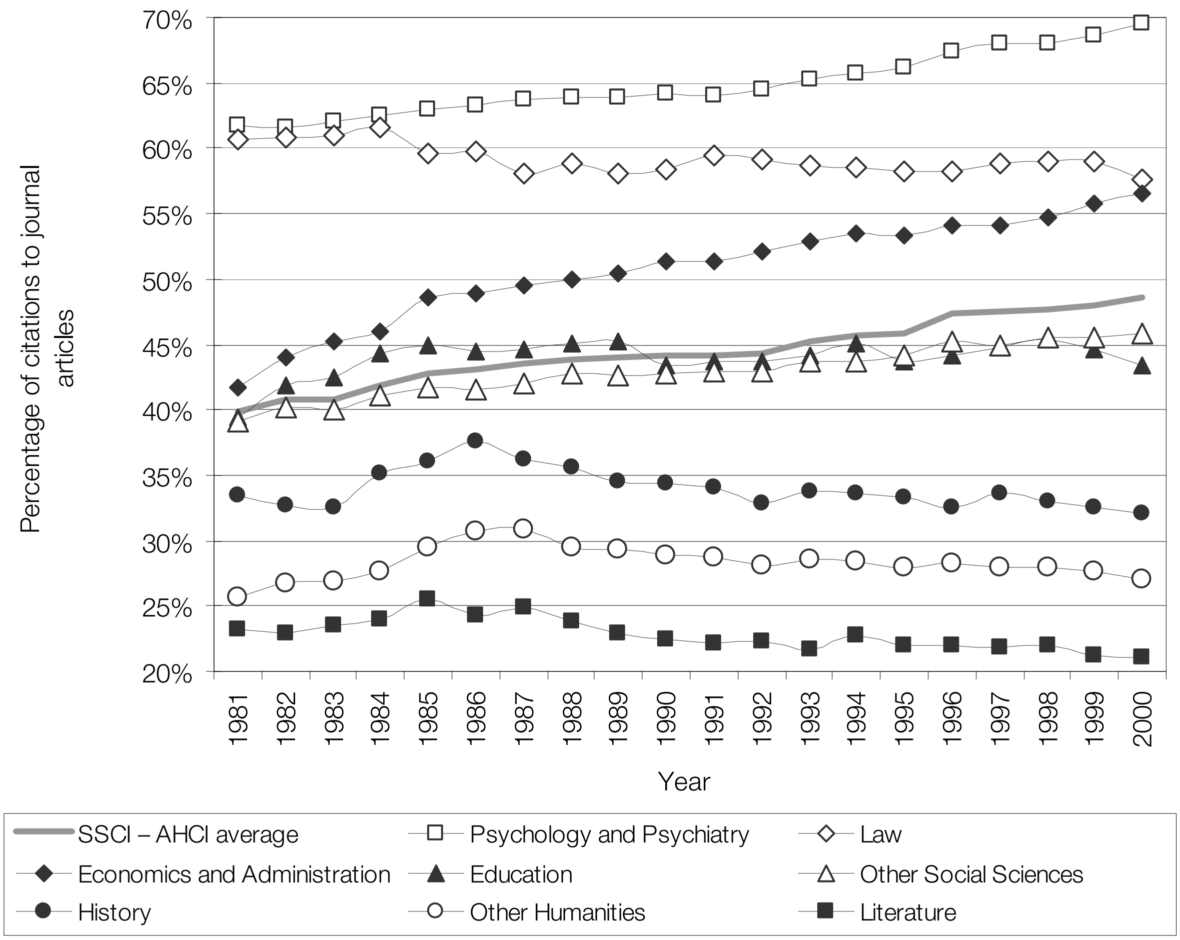

Figure 2 Percentage of citations to journal articles in science
Source: Larivière, Archambault, Gingras and Vignola-Gagné 2004

Thus the research article clearly does not play as central a role in the SSH as it does in the NSE. All things being equal, it could be claimed that there is a correlation between article and monograph publishing – that an evaluation based solely on journal articles is valid because article publishing corresponds closely enough to monograph publishing. However, all things are *not* equal.

As mentioned earlier, Clemens et al. (1995) have shown that scholars publish articles and books for different reasons. Their preference for one form over the other is often the result of socialization and habitus. In their analysis, Clemens et al. conclude that there is a divide between the two "worlds" of the article and the book. The findings of Line (1981) and Cronin, Snyder and Atkins (1997) support and quantify that conclusion. These authors show that books and articles form two more or less independent pools of citations. Line manually compiled 11,000 citations in 300 monographs and 48,000 citations in 140 academic journals. Line noted that, in 80% of cases, the monographs cited by the journals in the sample were cited only once. Only 27 books received more than 10 citations from the journals under analysis. In all, the journals assigned 47% of their references to other journals and

39% to books. Books assigned 51% of their references to other books and 25% to journals (note that references may represent more than one citation).

Cronin, Snyder and Atkins compiled 30,000 references from 90 monographs in sociology. A total of 26 authors are considered to be heavily cited, exceeding the team's threshold of 27 citations. They went on to compare the list of 26 authors and the number of citations to them with a list of the 26 authors most frequently cited in high-level sociology journals. The names of only 10 authors are featured on both lists. The citation-based ranking also changes, in some cases significantly. In their sample, Freud ranks fourth among authors most frequently cited in books but only sixteenth with respect to journals. Robert K. Merton, on the other hand, ends up with a 22.5 ranking (the 0.5 indicates a tie) among the authors most frequently cited in books but is ranked sixth for journals. Cronin, Snyder and Atkins conclude that there must be two parallel populations of heavily cited authors in sociology.

Example of differences between types of publication

Nederhof and van Raan (1993) compared the performances of six British economics departments that had received funding from the Economic and Social Research Council (ESRC). They counted the publications and the number of citations associated with them, including articles, books, book chapters and other types of publication. The results were then combined with a peer-review-based analysis. The study covered the period 1980–88.

All six departments combined produced 524 publications during the target period, including 140 articles listed by Thomson ISI, 53 articles not listed by Thomson ISI, 58 book chapters and 13 books. More than 260 documents did not belong to any of these categories: they included government reports and unpublished documents. The 13 books obtained an average of 3.15 citations, compared with an average of 0.95 for the articles covered by Thomson ISI. Thus the impact of books would seem to be the factor contributing the most to strong performance in the study. The results also show that strong performance by a group is often related to the strong performance of a single author whose citation rate is particularly high.

The work of Line and Cronin, Snyder and Atkins is certainly not without its shortcomings. It is not systematic and the samples are relatively small. Yet if their findings are combined with those of Clemens et al., the hypothesis that article-based measurement gives an adequate picture of research disseminated in monographs has to be rejected. Any evaluation that covered only one of the two media would probably be incomplete to varying degrees, depending on the discipline. This is why it is important to underscore the fact that the results of measurement of article-based research output cannot be generalized to total output.

Bibliometric indicators that are solely article-based are less representative of SSH research. Line (1999) is amazed at how confidently some bibliometric studies can present evaluation results in the social sciences on the basis of journal article analysis alone, given the importance of monographs in these disciplines. Ideally, SSH bibliometric research evaluation should include data on articles and books, and indeed on other scientific communication media as appropriate to the field in question. Unfortunately, the Thomson ISI databases do not provide this type of coverage, and no other database provides it adequately.

In cases where data is compiled from books and conference and other reports and abstracts, another set of problems arises. To be complete, these compilations should include citations between monographs, conference proceedings and articles. One of the advantages of such an analysis would be a larger citation pool, since analyses limited to one publication medium exclude citations from others. That being said, compilation of all this data would require considerable effort and significant financial resources. In addition, Lewison (2001) refers to problems with book analysis in bibliometrics — the fact that bibliographic data does not systematically include all authors' addresses, the fact that book content varies widely, and the lack of criteria for drawing comparisons between books. Villagra Rubio (1992) also did work on monographs and books and identified similar problems.

This section has shown that the SSH do not use only the article as a communication medium; they use other types of publication as well. However, fairly serious problems of coverage arise even in the case of articles, mainly because of the local orientation of the subjects examined and, by extension, the language in which articles are written.

3.2.2 Language and local interest

Whereas the problems identified in the NSE tend to be universal in nature, those of concern to the SSH are sometimes more localized. The literature suggests that the readership in the natural sciences is the international scientific community. Research in physics or biology can be carried out and reused by interested experts anywhere in the world, with the same results; the properties of an electron are the same everywhere on Earth. In the natural sciences, the corollary of this is dissemination media with an international focus, and the preferred medium is the scientific article, very often written in English and published in a journal with international distribution.

In the opinion of a number of authors, SSH research deals with local and regional problems more often than is the case in the NSE (Glänzel 1996; Hicks 1999; Hicks 2004; Ingwersen 1997; Nederhof et al. 1989; Nederhof and Zwaan 1991; Webster 1998; Winclawska 1996). It follows that SSH research subjects would be specific to a particular culture more often than would the problems tackled in the NSE. Hicks (1999 and 2004) suggests that theoretical concepts in the social sciences are more subtle and cannot be expressed in the universal language of mathematics as much as theoretical concepts can be in the natural sciences. In many cases, the concepts and subjects covered in the SSH can be expressed and understood only in the language of the culture that is shaping them. Accordingly, SSH scholars publish somewhat more often in their own language and in journals with national distribution.

Example of differences between NSE and SSH regarding local publications

Table II presents data on Finnish research output. It includes data on the number of publications written in Finnish, irrespective of type of publication (article, book, conference proceedings), as well as data on Finnish articles in peer-reviewed journals with an international distribution (this data necessarily covers only research articles). The data supports the thesis that the SSH are more locally oriented, since the number of publications in the first category is much lower than in the second. It also supports the thesis that the SSH is becoming increasingly internationalized, because the number of articles published in internationally distributed journals has grown significantly since 1994.

Table II Annual output of Finnish scholars

Field	Type of publication	1994	1998	2002
Natural Sciences and Engineering	Local orientation, written in Finnish	3,787	3,032	2,828
	International journal with peer review	6,419	6,702	7,857
	Ratio	0.6	0.5	0.4
Social Sciences and Humanities	Local orientation, written in Finnish	2,871	4,001	3,570
	International journal with peer review	685	984	1,265
	Ratio	4.2	4.1	2.8

Source: Academy of Finland 2003

If this epistemological conclusion is correct, do the Thomson ISI databases provide adequate coverage of literature that is more national in scope with a more local distribution? Several researchers have tried to answer this question, and their studies focus on the extent of SSCI and AHCI coverage according to country and publication language.

According to Royle and Over (1994), 73% of Australian articles in the natural sciences are covered by SCI, but only 27% of Australian articles in the social sciences are covered by SSCI. Nederhof and Zwaan (1991) noted that the coverage provided by the two databases varied considerably by field and journal importance and language. For example, only 3% of Dutch articles in public administration were covered by SSCI, compared with 58% of articles in experimental psychology. With respect to the humanities, coverage varied from 10% of articles in Dutch language studies to 39% of articles in general literature. It is interesting to note that, if articles from journals not considered to be scholarly (based on an international survey or experts) are excluded, SSCI's and AHCI's percentage coverage goes up, in some cases significantly.

According to Kyvik (1988, cited by Nederhof and Zwaan 1991), only one third of Norwegian publications in the SSH during the period 1979–81 were written in a language other than Norwegian, compared with 74% of the publications in the natural sciences. This clearly shows that language variables are important. It is a well-established fact that France, Spain and Germany are under-represented in SSCI (Ingwersen 2000). Even though all three countries enjoy a strong SSH tradition, this fact is not reflected in SSCI. For example, a study covering the periods 1989-93 and 1994-98 reports that Germany is last out of 17 countries ranked according research impact in the social sciences (Ingwersen 2000). More specifically, Schoepflin (1992) gives the results of surveys of German scholars, who were asked to identify the academic journals with the highest profile and the greatest value for their respective disciplines. Based on their responses, SSCI covers 94% of German journals in developmental psychology but only 26% in sociology and 8% in education. Given the strong German tradition in the social sciences, these figures are very surprising.

Andersen 2000 (citing Andersen 1996) suggests that the authors of 60% of the articles indexed in SSCI have U.S. addresses and the authors of 20% of them have U.K. addresses. Since Thomson ISI selects journals according to the number of citations they receive, the citation habits of the various language communities play an important role in the actual compilation of the Thomson ISI databases. For example, U.S. and U.K. sociologists cite articles written in English in 99% of cases. At the same time, those articles account for approximately 70% of international literature in sociology (Yitzhaki 1998). This explains in large part the Anglo-Saxon over-representation in the Thomson ISI databases.

A comparison of the UNESCO list of social sciences periodicals with that of ISI reveals some quite significant differences (Schoepflin 1992). The UNESCO list contains about 2.5 times more academic journals than the SSCI list. U.S. journals account for about 60% of SSCI coverage, yet they represent only 17% of all journals according to the UNESCO list. The comparison shows that SSCI includes more U.S. journals than the number included in the UNESCO list.

In light of statistics like these, a number of bibliometricians claim that the SSCI and AHCI databases have a bias in favour of English-language journals from Anglo-Saxon countries — specifically, the U.S., the U.K. and, to a lesser extent, Canada (Andersen 2000; Glänzel 1996; Nederhof and Noyons 1992b; Schoepflin 1992; Webster 1998). Because there was no solid data to back up their claim, Science-Metrix decided to check whether or not the Thomson ISI databases were sufficiently representative of the countries where and the languages in which knowledge is produced.

Table III presents statistics on proportion of journals by country of editor (not publishing company). It gives, in sequence, data on NSE and SSH journals covered in Thomson ISI's databases and in Ulrich's journal database, which is recognized as one of the most exhaustive directories available. The Table shows that journals with U.K. editors are very heavily over-represented in the Thomson ISI databases, especially in the SSH. According to Ulrich, 18% of journals have a U.K. editor, whereas Thomson ISI indexes 27% of reviews with an editor in that country — an over-representation factor of 55%. SSH journals with editors located in the Russian Federation, the U.S., Switzerland and the Netherlands are over-represented too, while virtually all other countries are under-represented.

Table III Journal coverage rates by country of editor

Country	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
United Kingdom	23%	17%	36%	27%	18%	55%
Russian Federation	1.6%	1.4%	12%	0.3%	0.3%	36%
United States	36%	31%	19%	50%	37%	35%
Switzerland	2.7%	2.1%	26%	0.6%	0.5%	8%
Netherlands	9.4%	8.3%	14%	7.7%	7.4%	5%
Canada	1.3%	1.3%	1%	2.5%	3.2%	-21%
France	2.4%	2.6%	-6%	1.0%	1.4%	-24%
Germany	7.7%	6.2%	25%	3.9%	5.9%	-34%
Japan	2.3%	3.7%	-39%	0.5%	1.0%	-55%
Australia	1.2%	2.1%	-42%	1.1%	3.6%	-71%
Spain	0.4%	1.3%	-72%	0.3%	1.0%	-75%
Belgium	0.2%	0.4%	-52%	0.5%	2.1%	-75%
India	0.9%	2.2%	-61%	0.2%	1.6%	-86%
Poland	0.7%	1.6%	-58%	0.2%	1.3%	-87%
Italy	1.1%	1.7%	-38%	0.1%	1.2%	-89%
China	0.9%	2.9%	-69%	0.1%	0.9%	-91%
Brazil	0.3%	1.1%	-72%	0.04%	1.0%	-96%
Other	7.5%	14%	-45%	3.5%	13%	-73%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

Table IV gives the distribution of Thomson ISI coverage and Ulrich's listing by principal language of country of journal editor. It shows that only journals with editors in countries where the language spoken is English or Russian are over-represented. For example, SSH journals with editors in French-speaking countries are under-represented by 27% in the Thomson ISI databases.

Table IV Journal coverage rates by language spoken in editor's country

Language	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
Russian	1.6%	1.4%	12%	0.3%	0.3%	36%
English	64%	55%	17%	82%	65%	26%
Dutch	10%	8.6%	11%	8.2%	9.4%	-13%
Czech	0.3%	0.7%	-55%	0.3%	0.4%	-15%
French	2.4%	2.6%	-6%	1.0%	1.4%	-27%
German	11%	9.0%	19%	4.6%	7.0%	-34%
Japanese	2.3%	3.7%	-39%	0.5%	1.0%	-55%
Danish	1.2%	0.9%	40%	0.3%	0.7%	-61%
Afrikaans	0.3%	0.6%	-57%	0.4%	1.1%	-66%
Swedish	0.3%	0.4%	-10%	0.2%	0.7%	-67%
Chinese	1.7%	3.9%	-56%	0.4%	1.7%	-76%
Spanish	1.0%	3.2%	-69%	0.6%	3.0%	-82%
Polish	0.7%	1.6%	-58%	0.2%	1.3%	-87%
Italian	1.1%	1.7%	-38%	0.1%	1.2%	-89%
Portuguese	0.3%	1.1%	-74%	0.1%	1.1%	-92%
Arabic	0.1%	0.6%	-90%	0.0%	0.5%	-100%
Other	2.3%	5%	-57%	1.1%	4%	-76%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

To determine the role of language factors in journal coverage rates in the Thomson ISI databases, Science-Metrix also considered the actual language of journals. Table V shows a clear selection bias in favour of journals in which the articles are written in English. Whereas 75% of peer-reviewed journals indexed in Ulrich are in English, almost 90% of those selected by Thomson ISI are in English, yielding an over-selection rate of about 20%. The only other over-represented language in the Thomson ISI databases is Czech, a language that plays a marginal role in the global science system. French is under-represented by 26%.

Table V Journal coverage rates by language of article

Language	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
English	89%	78%	13%	90%	75%	20%
Czech	0.04%	0.3%	-85%	0.2%	0.2%	8%
Russian	0.5%	0.9%	-48%	0.3%	0.4%	-24%
French	3.3%	3.4%	-3%	3.2%	4.4%	-26%
Multiple languages	0.2%	0.2%	-14%	0.3%	0.5%	-45%
Dutch	2.2%	2.2%	0%	1.3%	2.6%	-48%
German	3.2%	3.9%	-18%	3.0%	5.8%	-50%
Japanese	0.4%	1.7%	-74%	0.2%	0.6%	-64%
Swedish	-	0.1%	-100%	0.1%	0.4%	-69%
Spanish	0.6%	2.6%	-75%	0.9%	3.0%	-69%
Italian	0.1%	0.8%	-83%	0.2%	1.1%	-80%
Danish	0.04%	0.1%	-50%	0.1%	0.3%	-83%
Portuguese	0.1%	0.7%	-85%	0.1%	1.0%	-86%
Chinese	0.3%	2.4%	-88%	0.04%	1.2%	-96%
Polish	0.05%	0.7%	-92%	-	0.9%	-100%
Arabic	-	0.1%	-100%	-	0.3%	-100%
Turkish	0.01%	0.2%	-95%	-	0.1%	-100%
Other	0.3%	1.6%	-80%	0.3%	2.4%	-87%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

The data clearly shows that Thomson ISI's SSH journal selection favours English. This may be due to differences in the quality of research output according to the language of articles. According to Hodgson and Rothman (1999), 388, or 84%, of the 463 editors of the 30 most prestigious economics journals are affiliated with U.S. institutions. That being said, it is questionable whether research articles written in all languages other than English are of lower quality in all cases. On the contrary, the situation is attributable to Thomson ISI's inability to analyse the content of journals in languages other than English – a fact stated on the Thomson ISI website:

> English language article titles, abstracts, and keywords are essential. English language cited references are also recommended. Although, important scientific information is published in all languages, authors must provide English translations of article titles, author keywords, and abstracts if they hope to reach the widest possible audience. Likewise, as a purely practical matter, it would not be feasible for ISI to take on the task of translating this material[1].

[1] http://www.isinet.com/essays/selectionofmaterialforcoverage/199701.html/

It is important to factor this bias into any international comparative analysis. Indeed, any benchmarking based on SSCI and AHCI may well over-evaluate the U.S., the U.K. and Canada and under-evaluate Germany, Spain, France and other non-English-speaking countries, and this bias will affect both publication counts and citation analyses.

According to some authors, the problem arises only where bibliometric evaluation has to cover locally oriented research output. In the opinion of Moed, Luwel and Nederhof (2002), truly academic research should be relevant internationally and local orientation should not be a factor in bibliometric evaluations. From this perspective, research not covered by the Thomson ISI databases simply fails to reach the relevance threshold that would warrant closer evaluation. This is a weak argument, because it turns Thomson ISI into an impartial judge of what is and is not quality research output. In fact, because its selection criteria require journal bibliographic information to be in English, Thomson ISI may fail to index the content of an excellent journal in philosophy, for example, because its overall content is in German.

In light of the above statistics, there is cause for concern about the influence of language factors on bibliometric measurement of Canadian research. Because Thomson ISI has a bias in favour of English-language articles, Quebec and French-language research literature could be under-evaluated compared with English-language articles and journals. Science-Metrix examined this question in the context of its comparative analysis of the Ulrich and Thomson ISI databases. The results are presented in Table VI, which shows that Quebec is, in fact, somewhat over-represented in the SSH in Canada (17% of Canadian editors in the Thomson ISI databases, and 15% in Ulrich). In other words, based on the place of origin of SSH editors, Quebec is not under-represented in the Thomson ISI databases.

Table VI Journal coverage rates by editor's place of origin (Quebec vs. rest of Canada)

	NSE			SSH		
	Thomson ISI	Ulrich	Difference	Thomson ISI	Ulrich	Difference
Quebec	11%	13%	-18%	17%	15%	12%
Other provinces	89%	87%	3%	83%	85%	-2%

Source: Compiled by Science-Metrix from Thomson ISI's and Ulrich's data.

In the end, it is up to bibliometric evaluation practitioners to determine whether more locally oriented research can be factored into evaluation adequately or whether only international contributions can be considered. It is also up to them to advise their readers — science policy analysts and decision makers — that their tools have serious flaws, particularly in the humanities. At present, SSCI and AHCI coverage is adequate only in the case of journals of international scope published in English. This situation could soon change, however. A recent research paper suggests that the validity of SSCI data for benchmarking the SSH performances of non-Anglo-Saxon countries has been on the rise in recent years (Ingwersen 2000). This conclusion is supported by Hicks (1999 and 2004), who claims, among other things, that the social sciences are becoming increasingly international and less fragmented as scholars increasingly adopt common concepts and paradigms. As a result, a number of research fields are developing a core group of academic journals than can form a basis for solid bibliometric analysis.

The wider range of knowledge dissemination media in the SSH than in the NSE and the fact that SSH research output is sometimes more local in orientation are not the only factors making it difficult to use bibliometrics to measure SSH research. Consideration also has to be given to the distinctive citation practices of SSH disciplines.

3.2.3 Citation practices

Solid, valid bibliometric analysis requires large quantities of data (van Raan 2003): it must be possible to apply the law of large numbers. Given this requirement, some statistics on the Thomson ISI databases serve to explain why article counts and citation analysis can be problematic in the SSH. Every week, about 17,000 new articles and 300,000 new citations are added to the SCI database, 2800 articles and 50,000 citations to SSCI, and 2200 articles and 15,000 citations to AHCI (CNER 2002). SSCI articles make reference to about 72,000 discrete items; those in SCI, to about 195,000. In 1999, SSCI contained 2.3 million citations, while SCI contained over 20 million (Leydesdorff 2003).

Journal Citation Reports (JCR) is a Thomson ISI publication that gives figures for the journal impact factor. In 1999, the JCR for the social sciences covered about 1700 journals whereas the JCR for the NSE covered more than 5500. In the case of AHCI, there is no publication giving figures on citations. There is thus a major disparity among SCI, SSCI and AHCI with respect to number of citations. The impact factors presented in the social sciences JCR are less reliable than the NSE factors because of the fixed two-year citation window that Thomson ISI uses to calculate the impact factor. In the SSH, an article does not start to have a real impact until a few years after publication. Evidence of this lies in the fact that only 32 of the 100 most heavily cited authors in the humanities were born in the twentieth century (Nederhof, Luwel and Moed 2001 citing Garfield 1979).

There are other SSH citation characteristics that can limit the application of bibliometrics. Glänzel and Schoepflin (1999) described the specific bibliometric properties of a number of disciplines in the social and natural sciences (see Table VII). The proportion of research articles without references is only 3% in the natural sciences but between 3% and 30% in the various social science disciplines. The numbers are relatively low in the history of science, psychology, sociology and economics, but the proportion is 26% in information science and library science and 28% in business.

Table VII Bibliometric characteristics of literature in various disciplines, SCI and SSCI 1993

Discipline	No. of articles	Percentage of journals	Mean No. of references	Mean reference age
Immunology	23,396	94.3	29.6	6.9
Research Medicine	24,369	92.1	25.9	7.9
Solid State Physics	28,466	85.2	23.6	10.1
Analytical Chemistry	9,605	83.8	20.9	9.4
Mathematics	11,987	64.4	16.2	11.3
Psychology and Psychiatry	11,886	64.0	31.0	11.4
Electronic Engineering	19,222	62.2	15.0	8.6
Business	3,663	56.0	20.8	10.9
Economics	7,959	48.7	21.6	10.6
Library and Information Science	2,128	47.6	14.9	9.1
Sociology	3,675	40.4	32.7	12.5
History and Philosophy of Science and Social Sciences	658	34.7	48.7	38.8

Source: Glänzel and Schoepflin 1999

In 1993, SCI listed about 28,500 articles in solid state physics, while SSCI listed about 3700 articles in sociology. However, sociology has a higher mean number of references per article (about 33) than physics (about 24), while its mean reference age is only slightly higher (12 and 10 years respectively) (Glänzel and Schoepflin 1999). Based on these statistics, the number of articles and citations that can be analysed in the SSH is very small compared with the situation in the NSE.

As stated earlier, there is a major disparity between the rate of citations to academic journal articles and the rate of citations to other knowledge dissemination media. The article rate is only 40% in sociology, but it is twice as high (85%) in physics. Other statistics can de adduced in support of the need to place greater importance on media that are covered little, if at all, by SSCI and AHCI. In 1999, approximately 2.4 million citations were compiled in SSCI. However, only about one million of the cited references, or 45%, are included in SSCI. The figure is 79% for SCI (Leydesdorff 2003).

The time required to accumulate citations in the SSH makes analysis more difficult, particularly when the goal is to assist in decision making and policy setting. The citation pool varies by discipline. This means that analysis can be done in some fields but should be avoided for methodological reasons in others, where the law of large numbers does not apply and the citations compiled do not adequately reflect the dominant dissemination medium. Monographs and related citations simply cannot be ignored.

Co-citation analysis is a method that can be used to characterize the structure of science. It is potentially a means of identifying emerging research themes and fields. Hicks (1987) identified a major problem with the use of co-citation analysis for science policy purposes. The method requires a fairly high number of articles and citations so that researchers have a critical mass to work on. However, a new research field may emerge well before the critical mass required for co-citation analysis has been built up. Hicks gives the example of spin glass research, where the critical mass developed more than ten years after the field itself started to emerge. The time lag between emergence of the field and development of the critical mass apparently varies and is influenced in part by external factors (sociological factors, etc.). This is why it is especially difficult to apply co-citation analysis in a science policy context, and the difficulty is compounded by the fact that the bibliometrician is identifying emerging research fields in the SSH, where the number of citations is built up more slowly.

Hicks identifies another obstacle to co-citation analysis for science policy purposes — the process of delineating research fields. This step is required for all bibliometric analyses. According to Hicks, the delineation process should always meet the following requirements: it should be explicit, it should seem reasonable to a number of scholars in the field (even though they probably do not agree on where the boundaries of their field lie), and it should be stable over time so that trends can be established with an adequate degree of reliability. Unfortunately, a number of co-citation analyses do not meet these requirements.

Hicks identifies even more problems: citation errors, under-representation of certain types of articles, and some undeclared methodological subjectivity. Finally, Hicks highlights the prohibitive cost of developing a solid co-citation-based analysis. In many cases, the required resources include experts in bibliometrics and science policy or research management. Van Raan (1998) challenges Hicks's conclusion, claiming that bibliometric analyses are an inexpensive complement to peer review. In fact, the two authors are referring to two different bibliometric methods. The one discussed by van Raan is simpler, and that may explain the difference of opinion. Note also that Hicks's analysis dates back to 1987. With the development of new software, costs may have gone down.

3.3 Potential for SSH bibliometric evaluation

In Section 3.2, it was shown that the SSCI and AHCI databases cannot be used to produce SSH analyses that are as solid as the SCI-based NSE analyses. Since there are no alternatives that are as complete as the Thomson ISI databases, SSH bibliometrics researchers have no choice but to conduct SSCI- and AHCI-based analyses of more limited scope and with a potentially significant bias or invest substantial resources in each evaluation so as to compile data on books and literature with a more local orientation.

Given the methodological limitations outlined above, some researchers might be tempted to give up on bibliometrics as an SSH research evaluation tool. However, it is important to consider the position taken by van Raan (2003): bibliometric evaluation must always complement peer review. Moreover, the shortcomings of bibliometrics and the SSCI and AHCI databases can be offset in many instances. In spite of their limitations, the results of bibliometric analysis, when presented clearly and transparently, bring out dynamics that cannot easily be identified on the basis of expert judgment.

Other factors militate against dropping SSCI and AHCI as evaluation and mapping tools. For example, notwithstanding all its limitations, Hicks (1999) concludes that SSCI is still the best available database on SSH publications and gives the following reasons why:

- complete indexing of authors' addresses
- complete indexing of a known proportion of academic journals
- multidisciplinary coverage
- indexing of citations
- international coverage

In addition, some of the shortcomings of these databases are reportedly becoming less serious, as has already been mentioned, and the SSH are going through a transformation and adopting forms more suitable for bibliometric evaluation. Hicks (1999) claims that the social sciences are going through a homogenization process to some extent. According to her, globalization and the internationalization of economies are possibly playing a major role here. In Europe, the European Union has initiated funding for social science research, which could enhance the international dimension of research. The fact that a number of East European countries have joined the Union could help to disseminate their research output internationally.

Lastly, specific concepts and paradigms are apparently gaining currency in all social sciences. With the adoption of common analytical frameworks, the number of divergent paradigms could be dropping, and this would help to create a solid core of high-level academic journals that are recognized as such by scholars in each discipline and receive a large proportion of citations in it. The existence of these cores explains in part why the NSE are more suitable for bibliometric evaluation.

Note, however, that the above observations seem to be based on Hicks's intuition and are therefore at least partly speculative instead of being derived from empirically verified facts. That being said, they seem to be correct to some extent, because studies have confirmed the growing validity of SSCI- and AHCI-based SSH bibliometric evaluations. For example, 36% of citations in the humanities are assigned to articles published less than four years previously (Luwel et al. 1999 citing Zwaan and Nederhof 1990). The number of citations for articles in linguistics and literature peaks in the third and fourth years (Nederhof and Noyons 1992b). These two statistics indicate that some humanities disciplines possess similar characteristics to those in the natural sciences. Furthermore, So (1998), in a performance analysis of scholars in communication, has shown that citation-based evaluation is valid in that discipline.

However, the changes referred to by Hicks and others may not occur or may not be extensive enough for SSCI and AHCI to become reliable tools. Some authors have looked at an alternative — building databases themselves. In a report submitted to the French government's minister with responsibility for research and new technologies, Philippe Jeannin (2002 and 2003) examines research evaluation in the social sciences. Starting from the premise that SSCI cannot be used for bibliometric evaluation because its coverage is too limited (particularly for France), Jeannin proposes a method of establishing a list of journals on which evaluations would be based. Yet all the important work still has to be done once the list has been drawn up, and building up a database is a laborious, long-term, expensive undertaking that would not be suitable for most evaluation projects. Furthermore, Jeannin's method does not include books and monographs.

The validity of a bibliometric method depends on the discipline to which it is applied. Today, there are publication media in fields such as linguistics, experimental psychology and economics that are similar to those in many of the natural sciences, with a core of international journals playing a predominant role and significant growth in citation numbers for recent works and articles (Nederhof and Zwaan 1991). So at this stage the publication characteristics of the various SSH disciplines need to be identified so as to gauge the bibliometric evaluation potential for each one.

It is also worth noting that bibliometrics can be used for science policy purposes other than evaluation. Mapping methods can be used to map international collaborative research activities and expertise, the large-scale coverage of which would otherwise be too time-consuming and labour-intensive. These methods are based on Thomson ISI databases too but face fewer problems. The ties between scholars in a variety of institutions are becoming closer all the time, in the social sciences just as much as in the NSE (the same cannot said for the humanities – see Gingras 2002). Maps of research fields can be very useful tools for science policy making.

Example of research field mapping

McCain (1986) mapped the work of 41 authors in macroeconomics for the years 1978-82, using the co-citation method of analysis. The results are shown in Figure 3. The horizontal axis represents a continuum of knowledge and research traditions. The vertical axis represents a continuum of the empirical and/or econometric content of cited documents. The higher on the vertical axis authors are cited, the greater the empirical and mathematical content of their work is. This is a good example of mapping being used to establish the morphology of a discipline.

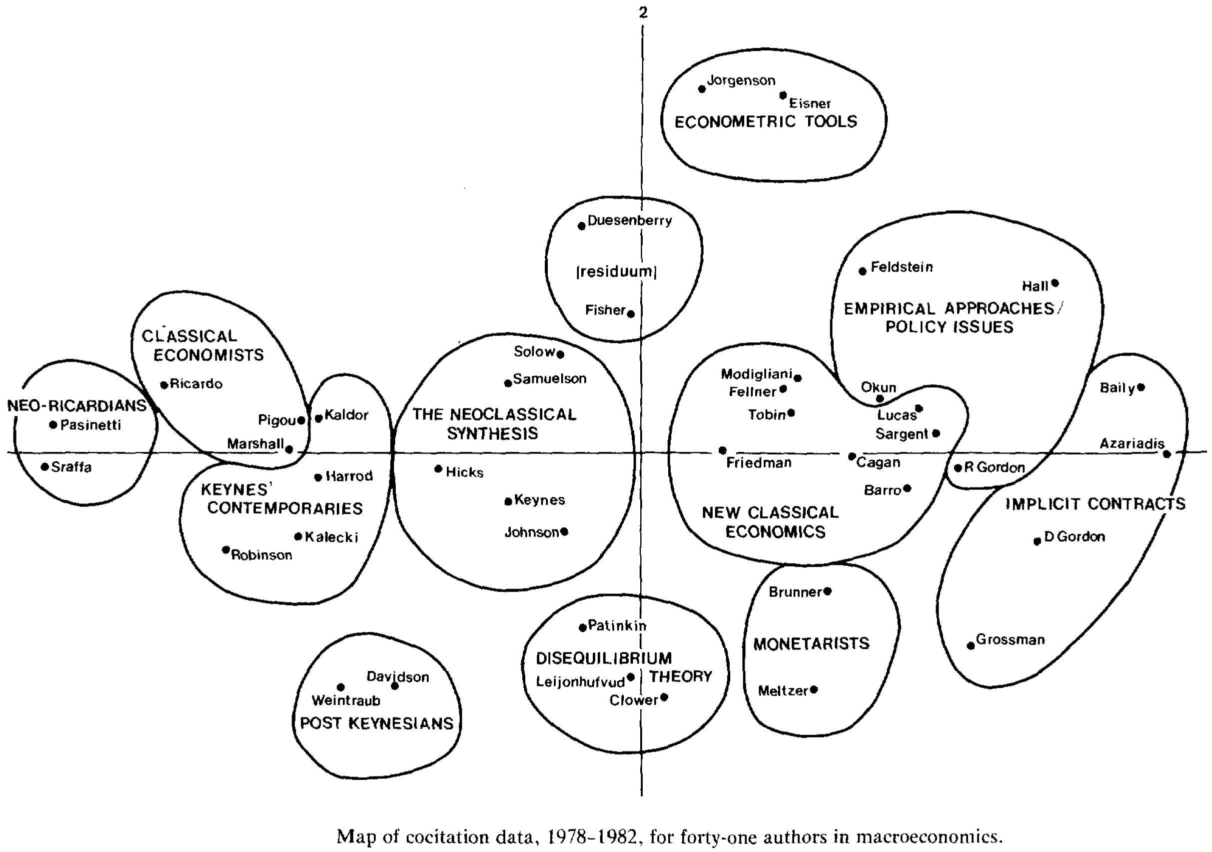

Figure 3 Authors in macroeconomics, 1978-82
Source: McCain 1986

3.4 Building databases

As mentioned above, the shortcomings of SSH bibliometric analyses based on the Thomson ISI databases may become less serious over time. In addition, the development of open access promises to make research articles available on a market dynamics basis and may well reduce the selection bias inherent in the Thomson ISI tools. All this is hypothetical, however, and it is therefore worthwhile trying to devise solutions to offset database limitations. There are at least two ways of solving the problems identified: (1) try to remove the limitations at the source, and (2) work around them. The first option is examined here. The second will be examined in greater detail in sections 4 and 5, which describe tools to measure national performance and identify emerging fields.

Removing at the source problems in applying bibliometrics to the SSH requires building new databases or upgrading, or perhaps expanding, existing ones. There are two types of databases: temporary databases built for specific projects; and databases that are built up systematically and updated continuously.

For some evaluation projects, it may be useful to build ad hoc databases that include indicators for generating reliable measurements of research output. They could also include data on all addresses for collaborative research calculations and data on citations for research output impact evaluation. They could be exhaustive, containing all types of relevant research outputs

For example, the OST has established the Banque de données sur les revues savantes québécoises (BRSQ) [Database on Quebec learned journals] for the purpose of analysing the research dynamics reflected in Quebec journals (Godin, Archambault and Vallières 2000). In addition, a new information service may be set up on a permanent basis. In both instances, an investment commensurate with the size of the database is required. The BRSQ does not index citations or monographs and has required a lot of time and effort. The analysis of Cronin, Snyder and Atkins (1997) yields similar findings, although the scope of their study was quite narrow. Yet with the increasingly widespread use of electronic data interchange (EDI), the cost of building databases of this kind should go down steadily.

Another approach is to use and upgrade existing tools. There is a Canadian repository of curricula vitæ of university researchers, including information on their publications, known as the Common CV System. If the bibliographic information in the System were standardized and supplemented as required, evaluators would then have the best possible database for conducting detailed, accurate, in-depth bibliometric analyses. The CVs include articles, books, conference proceedings and other media. Another solution would be to merge this information with the data in existing databases.

Example of ad hoc database

Nederhof and Noyons (1992b) evaluated the performance of Dutch general linguistics and literature departments. They compared the selected departments with foreign departments that were performing well. Their comparison was based on citation counts and covered the period 1980–85.

The results in linguistics show wide year-over-year variations for each of the four departments under consideration. For three of them, books and book chapters are much more heavily cited than journal articles. The mean annual impact of the departments varies from 1.0 to 3.1. The authors also determined the countries of origin of citations assigned to departmental publications so that they could gauge how "international" they were. One department received half of its citations from abroad, while the others received 31–37%. The performances of the three departments of comparative literature were very different. The Dutch department achieved a mean annual impact of 0.9-1.0 for four years out of five, and then improved to 1.6. The impact of the U.S. department varied from 2.2 to 5.0; that of the German department, from 1.5 to 2.3. The U.S. department's performance was attributable to books, because its articles had the same impact as those of the Dutch department.

4 Bibliometric methods for benchmarking national performances in the SSH

This section provides information on a number of techniques and indicators that can be used to benchmark national performances in the SSH, including the properties of the main indicators and key methodological considerations.

4.1 Article analysis

Any benchmarking of national performance is based on a relatively simple unit of measurement: the number of publications produced by an actor in the science system. The methods described here use journal articles as core data to generate indicators measuring various facets of the scientific knowledge production system. One indicator, the specialization index, is calculated by means of simple ratios derived from the number of articles. It can be used to gauge the expertise of a country or institution in various disciplines or areas of specialization. The publication count and the specialization index can be used to analyse research output at a number of different levels, but they will not yield any information on output quality. Citation counts can be used to determine the impact of publications, and it is possible to calculate a mean impact factor for research output on the basis of statistics in the JCRs.

4.1.1 Publication counts

Publication counts are the basis for any performance benchmarking exercise. The number of publications by a scholar or country is an indicator of their level of *production* of new knowledge. The number of publications can be used as is, but it can also be expressed in relative terms: for example, expressed in relation to number of scholars and level of funding, it can generate an indicator of *productivity*.

A number of authors have analysed SSH performance using article counts. Katz (1999) studied the production of economics and psychology articles by U.K. scholars, using SSCI data. He evaluated performance by calculating U.K. articles' percentage share of world output in the two disciplines. Godin (2002) and Uzun (1998) conducted a similar analysis of output in the social sciences in general, focusing on Canada and Turkey respectively.

In an effort to offset the limitations of SSCI with respect to Polish sociology, Webster (1998; see also Winclawska 1996) combined SSCI data with information from a database on Polish sociology for the purpose of establishing publication and citation counts. The approach increases coverage, but its scope is still limited because it cannot generate international comparisons.

Example of publication count

In a study for the Economic and Social Research Council (ESRC), a U.K. organization, Katz (1999) evaluated the research output of a number of countries in the social sciences. The results are presented in Table VIII. Katz used Thomson ISI data. Articles written by several authors from different countries are included in the total of each country involved; this is why total counts exceed 100%.

Table VIII Countries of origin of research output in seven disciplines, 1998

Country	Communication	Law	Economics	Education	Management	Psychology	Sociology and Anthropology
United States	77%	89%	57%	61%	54%	57%	56%
European Union	16%	8%	32%	22%	33%	28%	25%
United Kingdom	9%	5%	13%	15%	19%	10%	11%
Canada	3%	2%	6%	6%	6%	8%	5%
Netherlands	3%	1%	4%	2%	4%	3%	2%
Australia	2%	1%	4%	5%	4%	4%	3%
Germany	1%	1%	4%	3%	2%	7%	3%
France	0.1%	0.2%	3%	0%	2%	2%	5%
Total No. of articles	866	1,749	7,055	2,560	2,611	14,711	4,263

Source: Katz 1999

Glänzel (1996) used SSCI to compare national performances from 1990 to 1992 in six disciplines: business and management, economics, psychology and psychiatry, sociology, information and library science, and history and philosophy of science and social sciences. For each discipline, the author established publication counts by country and percentage share of output, by discipline, for articles from each country.

The above studies base benchmarking on absolute numbers of publications. In the opinion of the CNER (2002), however, it does not suffice to use only this indicator when comparing national performances; the count should be weighted by the number of scholars in the country, the R&D budget and other input indicators. Ingwesen uses weights of these kinds (Ingwersen 1997; Ingwersen and Wormell 1999), calculating number of publications per 1000 inhabitants.

Article count is a very useful indicator and it is the first step in bibliometric analysis. However, the information it provides is limited. This explains why most bibliometric studies go further, combining article count with indicators like the specialization index.

4.1.2 Specialization index

Using the specialization index, researchers can quickly identify the disciplines in which a country, region, institution or any other aggregate achieves greater research output than in all other disciplines. The specialization index can be represented in two ways. First, as stated above, an aggregate is said to be specialized when it produces more in a specific discipline than in all other disciplines. The second approach is to consider an aggregate as specialized when its percentage of output in a given discipline is higher than the other aggregates contributing to a system. In other words, the second approach indicates more precisely the specialties in which an aggregate's output is larger or smaller than the average of a group. These two ways of representing the specialization index are equivalent and produce the same result, as shown below.

Each approach has a corresponding formula for calculating the specialization index. Thus the specialization index of a group *x* relative to a reference group *y* ($IS_{x/y}$) can be calculated in two ways:

$IS_{x/y} = (x_a / x_t) / (y_a / y_t) = P_{xa} / P_{ya}$ or

$IS_{x/y} = (x_a / y_a) / (x_t / y_t) = P_{ax} / P_{ay}$

where

- x_a = number of articles published by group *x* in discipline *a*
- y_a = number of articles published by group *y* in discipline *a*
- x_t = total number of articles published by group *x*
- y_t = total number of articles published by reference group *y*
- P_{xa} = percentage of articles of group *x* belonging to discipline *a*
- P_{ya} = percentage of articles of group y belonging to discipline *a*
- P_{ax} = percentage of articles in discipline *a* produced by group *x*
- P_{ay} = percentage of articles in discipline *a* produced by group *y*

Group *x* is always a subset of group *y*. An index higher than 1.0 indicates that *x is specialized* in relation to *y* and an index lower than 1 indicates that group *x* is not specialized in discipline *a*.

For example, suppose that one wants to find out the specialization index for Nova Scotia (*x*) relative to Canada as a whole (*y*) in the discipline of sociology (*a*) relative to output in all the SSH (*t*). Suppose that Nova Scotia produces 7 articles in sociology and a total of 10 articles in the SSH, and that Canada produces 15 articles in sociology and a grand total of 100 articles in the SSH. The table below shows that the two ways of representing and calculating the specialization index produce the same result:

	x	*y*	
a	**7**	**15**	$x_a / y_a = 46.67\%$
t	**10**	**100**	$x_t / y_t = 10\%$
	$x_a / x_t = 70\%$	$y_a / y_t = 15\%$	$IS = x_a / x_t / y_a / y_t$

$= x_a / y_a / x_t / y_t$

$= 4.67$

The specialization index is in very widespread use internationally but is often given a different name. It is often called the "revealed scientific advantage" in the Anglo-Saxon world (Soete and Wyatt 1983), while French researchers may use the term "indice d'effort spécifique" (see, for example, Filiatreau et al. 2003). Science-Metrix and the OST use the term "specialization index" in all their reports.

The specialization index is a relative indicator providing relatively complex, highly synthesized data. It is one of the best indicators for determining the areas where research output of one group differs from that of the others. The advantage of using it is that it is relatively easy to calculate; the disadvantage is that it requires data on both specific disciplines and total output of a reference population. The bibliometrician therefore needs complete databases in order to be able to produce data by using this indicator.

4.1.3 Citation analysis

Once the number of published articles has been determined, it is possible to assign a value to each one of them as a reflection of their impact on the science community or even of their "quality." A number of methods and indicators are available for this purpose. First, the mean number of citations received by articles can be calculated (observed citations) and then compared with that of comparable groups. Second, an indirect quality indicator can also be assigned to articles by using the impact factor of the journals in which articles are published (presumed impact). Third, the number of observed citations can be compared with presumed impact or other international averages to gauge the performance of one group in relation to those of reference populations.

The most frequently used indicator of the research impact of publications is the total number of citations received by articles of a scholar, institution, country or any other relevant unit of analysis for a given period (van Leeuwen, van der Wurff and van Raan 2001; van Raan 2003). The citation count and any other citation-based measure can be determined with self-citations being included or excluded.

Example of publication and citation counts

Glänzel (1996) conducted a bibliometric analysis of six SSH disciplines, using SSCI data for the period 1990–92. Among the methods applied were publication and citation counts. Table IX presents the results for psychology and psychiatry:

Table IX International performances in psychology and psychiatry, 1990-92

Country	Publications		Citations	
	Count	Percentage (%)	Count	Percentage (%)
United States	22,121	63.4	45,196	72.3
Canada	2,845	8.2	5,589	8.9
United Kingdom	2,601	7.5	5,496	8.8
Germany	1,481	4.2	1,892	3.0
Australia	1,015	2.9	1,673	2.7
Netherlands	726	2.1	1,601	2.6
Japan	600	1.7	436	0.7
Israel	557	1.6	759	1.2
France	498	1.4	787	1.3
Italy	336	1.0	452	0.7
Total	**32,780**	**94.0**	**63,881**	**102.2**

Source: Glänzel 1996
Note: Glänzel did not deal with duplications due to articles written by two or more authors from different countries; this explains the totals in the Table.

It is possible to calculate a self-citation percentage. Another indicator is the number or percentage of articles receiving no citations (excluding self-citations). Dividing the number of citations by the number of articles in the initial total population gives the mean number of citations per article (van Leeuwen, van der Wurff and van Raan 2001; van Raan 2003), or observed research impact (Katz 1999).

The impacts of different countries and institutions can be compared on the basis of absolute numbers, or further calculations can be made to gauge relative impact by comparing observed and presumed impacts. Presumed impact can be calculated on the basis of the impact factor — as defined by Thomson ISI in the JCRs — of the journals in which articles are published. It is also possible to calculate presumed impact using the mean number of citations received by all the journals in which the actor has published, the total number of journals in the actor's field, or both. The various methods used by bibliometrics practitioners to calculate presumed impact are outlined below.

Uzun (1998) determined the characteristics of Turkish research in the social sciences by assigning to the articles analysed the impact factor established by Thomson ISI for the journals in which the articles were published. Thomson ISI obtains the impact factor by calculating the mean number of citations received by each journal's articles two years after publication. Uzun compared the impact factors of Turkish articles with maximums attained in a given field.

Ingwersen, Larsen and Noyons (2001) applied a method developed by Noyons (2001) to the social sciences. The observed number of citations for a country in a specific field is divided by the mean number of citations received by articles published in that field to yield what is known as the Tuned Citation Impact Index (TCII):

$$TCII = \frac{\sum C_f}{\sum (P * WIF)_f}$$

where C_f is the number of citations received by the country, P is the publications for the country, and WIF the World Impact Factor, or mean international impact indicator, for a field f. The WIF of field f is calculated by dividing the world total of citations received by the field by the number of publications in the field.

Glänzel (1996) uses a similar method to benchmark national performances in business, economics, psychology, sociology, information and library science, and history and philosophy of science. The researcher calculates the mean observed citation rate and the mean expected citation rate. The latter indicator is based on data on all journals in a particular field that have been fully covered by SSCI. Dividing the first indicator by the second yields the relative citation rate of a country for a field.

According to Nederhof and Noyons (1992b), it is preferable to calculate expected impact from data on all the journals in which research institutions publish because data on an entire field is not sufficiently refined and does not take into account field subspecialties with lower citation numbers, etc. The authors developed their method to evaluate institutional rather than national performance, and using it to gauge national performance can indeed be problematic. It is possible to calculate the performances of all institutions active in a field for a given country and then aggregate them, but it can require a major investment in resources, especially when comparing many countries. Another way of offsetting the limitations identified by Nederhof and Noyons is to apply the "normalized and weighed impact factor" (NWIF) suggested by Gingras (1995). Science-Metrix and the OST use this indicator under the French name "moyenne des facteurs d'impact relatifs" (MFIR), for calculations based on the journal impact factor and the relative citation average (RCA).

Example of relative impact measurement

The Academy of Finland used a number of bibliometric indicators in its report on the state of research in Finland in 2003. Table X presents the results of relative impact calculations for the 15 countries with the greatest impact in the SSH. Relative impact is calculated by dividing the mean citation number for publications of a given country by the mean number of citations received by all OECD countries. A score above 1 indicates that the country receives more citations than the mean score for OECD countries. The data is taken from the Thomson ISI databases, so the limitations identified earlier apply.

Table X Relative research impact of OECD countries, 1998-2002

	Social Sciences		Humanities	
	Country	Relative impact	Country	Relative impact
1	United States	1.15	Greece	2.16
2	Netherlands	1.04	Netherland	1.56
3	Belgium	1.00	New Zealand	1.35
4	Canada	0.99	Finland	1.33
5	Italy	0.99	Portugal	1.26
6	Finland	0.97	United Kingdom	1.23
7	United Kingdom	0.96	United States	1.19
8	Switzerland	0.93	Australia	1.17
9	Sweden	0.92	Japan	1.16
10	France	0.85	Turkey	1.14
11	New Zealand	0.82	Denmark	1.00
12	Norway	0.81	Hungary	0.94
13	Australia	0.80	Canada	0.92
14	Hungary	0.80	Belgium	0.82
15	Germany	0.80	Germany	0.77

Source: Academy of Finland 2003

Van Leeuwen, van der Wurff and van Raan (2001; see also van Raan 2003) developed a method combining the indicators described above with a new one. Their method includes calculating the usual core indicators: number of publications, number of citations and mean number of citations per article. They also calculate self-citation and non-cited article percentages. The mean number of citations to publications is given a relative value in terms of both the mean number of citations for the group of journals in which the articles are published and the mean number of citations for the field. Since it is possible to publish in more than one field, they use a weighting system (the problem applies only to individuals and institutions, not to countries, since national performance is normally measured by research field). Then the mean number of citations for the group of journals in which the actor publishes is divided by the mean number of citations for the research field to determine whether the individual or institution concerned generally publishes in prestigious journals in the field (without giving any indication of intrinsic quality of articles).

Lastly, Katz (1999) gives a few words of caution on the use of impact indicators, claiming that existing calculations of presumed impact have been slightly distorted by a bias in favour of countries producing the greatest number of articles (primarily the U.S.) in terms of citation attribution. Katz proposes a modified expected impact indicator, the "relative international citation impact," to take into account these scale effects.

4.1.4 Peer-review-based impact indicators

The methods outlined in the previous section use citation counts to evaluate the quality or relevance of a population of research articles. They are purely statistical, and subjectivity has no role to play in them (except in the selection of parameters, such as analysis time frame and delineation of research fields). Lack of subjectivity is generally considered one of the main advantages of bibliometrics. Some authors suggest, however, that basing indicators on statistical methods alone means losing other analytical dimensions. This section outlines the efforts of two scholars to build a peer-review-based indicator. While the results of their efforts cannot, in the final analysis, be considered conclusive and their method is not really a bibliometric one, it is nonetheless worth examining it as an alternative.

The originality of the approach adopted by Nederhof and Zwaan (1991) lies in their method for evaluating the expected impact of a journal. They endeavour to weight article quality by determining the impact of journals on experts' judgment rather than on citations received. Nederhof and Zwaan conducted an international survey of 267 experts in six SSH disciplines, asking them to give academic journals a quality rating from 1 to 10. They then calculated scores for each journal from the survey results. This method makes it possible to weight research output on the basis of criteria that are potentially more representative of research quality than the number of citations received. In theory, it may also be very useful for evaluating articles in journals that are not covered by AHCI and SSCI. Furthermore, an indicator based on peer judgment may enjoy greater legitimacy in the eyes of the research community than a statistical one.

However, it turned out that this approach to journal evaluation was not as useful as expected. First of all, it is not very accurate or precise, because experts tend to give similar ratings to the vast majority of academic journals (between 5.7 and 8 in Nederhof and Zwaan's case study). Second, it has limited potential for evaluation of small local and specialized journals, which are precisely the ones that the evaluation should target because they are not covered by SSCI and AHCI. The fact is that few foreign experts are sufficiently familiar with the smaller journals to be able to evaluate them properly. At the same time, scholars who are familiar with some of them may have a more favourable opinion of journals connected with their field of interest. In short, this method, which lies in between peer review and bibliometric evaluation, offers none of the advantage of these two approaches.

4.2 Variety of publication media

The methods outlined above are used primarily for bibliometric evaluation based on journal articles, and the data required for analysis comes mainly from the Thomson ISI databases. Yet, as stated in Section 3, bibliometric evaluation of a number of SSH research fields cannot be based on articles alone. To compensate for this limitation, some scholars have developed methods that factor in the publication of monographs, conference proceedings and other types of output. They are described below.

4.2.1 Counts for publications of different types

As in the case of articles, analyses of more than one type of publication can start with a count indicator. A country that has published more articles and books than another one would have a higher performance rating.

Van der Meulen and Leydesdorff (1991) used publication counts to evaluate changes in the productivity of Dutch departments of philosophy in the wake of institutional restructuring. They counted articles, contributions to books and monographs.

Nederhof and Noyons (1992b) identify a major obstacle to counts for publications of different types. Counts require data on the publications. The data sources can be departmental publication lists and databases, which are not always complete. In other words, it may be difficult to obtain exhaustive, accurate publication lists. Nederhof and Noyons developed a method for "reconstituting" lists, including articles and books. They started with AHCI and then scanned other databases. They completed their data gathering with an examination of reference lists in specific articles and books.

In an evaluation of the performance of law faculties, Moed, Luwel and Nederhof (2002) use the results of a survey of faculty members to establish publication lists. The publication count procedure is basically the same for articles and for other types of publication. Data on publications can be gleaned from databases and/or lists of publications published by a number of departments and institutions and then aggregated to produce national profiles.

4.2.2 Weighting publication media

Once drawn up, the publication lists can be compared. At this stage, the problem is to determine whether some types of publication should be considered more important than others. Should a 20-page article have the same weight as a 300-page book? A number of scholars have developed schemes for weighting research publications and thereby refining the benchmarking procedure. Their work is described below.

Following up on the work of Luwel et al. (1999) on the establishment of SSH performance indicators for the government of Flanders, Moed, Luwel and Nederhof (2002) tried to codify the value of publications against a variety of parameters. First, they ascertained whether there were any correlations between the number of publications by a scholar and recognition by his or her peers, between the number of pages in a publication and its quality, etc. They did this by combining a survey of university scholars with bibliometric measurements. In light of the results, they drew a distinction between "substantial" research contributions, "small" contributions, publications for the general public and other types of publication.

One interesting discovery was that 84% of substantial publications are more than 5 pages long, compared with 17–23% for the three other types mentioned above. They therefore suggest that it is possible to identify a substantial contribution by the number of pages in the publication. Researchers then conduct their evaluation by comparing numbers of substantial research publications produced. Publication lists are drawn up from data generated by surveys of the scholars under study. This leads to the calculation of a weighted number of substantial publications. Article weights are based on the journal in which they are published; weights of contributions to books and monographs, on the number of pages. Moed, Luwel and Nederhof emphasize the importance of keeping the evaluated scholars involved in the evaluation process so that it is as accurate as possible.

They do, however, identify certain problems with their method. While number of pages as an indicator of publication quality is based on results confirming its validity, it seems arbitrary. In addition, the method is very time-consuming. The survey data has to be cleaned up and standardized, because scholars make a number of errors in classifying their own publications. For example, they may count two editions of the same book as two separate publications. This makes the process of evaluating just one group or department very long and expensive. Therefore the method does not appear to be appropriate for the evaluation of national performance.

In evaluating the performance of six Hong Kong universities in the SSH, Ho (1998) uses a similar method to that of Moed, Luwel and Nederhof (2002). Ho draws up a list of publications from the universities' annual reports and then weights each publication on the basis of a fairly complex set of rules established in cooperation with experts in the field. For example, he assigns a score of 0.5 to an article in a mass-circulation newspaper but 20 to a journal article or a book with an international impact. He also applies rules such as a 50% reduction in score for a book with fewer than 100 pages and point splitting according to number of authors. The scores of the various actors are then compared. While the system may compensate to some extent for the lack of refinement in publication counts, the amount of work required to assign scores to each publication may become excessive, and the weights are highly subjective and therefore difficult to reproduce.

Finkenstaedt (1990) also uses a method of weighting publication counts, but it covers only academic publications (unlike Ho 1998), and the scope of the analysis is too limited to provide useful answers to the problems considered in this report.

Economists are one of the groups that use bibliographic data the most often to evaluate their discipline. A number of economists have developed methods for weighting research output. After values have been assigned to economics journals, economics departments are rated by calculating the number of pages published and then applying to it the value of the corresponding journal (Baltagi 1999; Coupé 2003; Dusansky and Vernon 1998; Lucas 1995). The value is often based on citation analysis. In other words, each page of an article is weighted by the presumed impact of the journal in which it is published. Some authors go as far as calculating a mean number of characters per journal article (Kalaitzidakis, Mamuneas and Stengos 2002) and even mean number of citations by character (Laband and Piette 1994).

The above methods, along with that of Nederhof and Zwaan (1991) described in Section 3, propose a codification of peer review rather than a full-fledged bibliometric approach, with "self-organized" data. This in itself does not affect the validity of these methods, but some of the epistemological foundations of the evaluation process are different. For example, it would be inappropriate to attribute to them the objectivity commonly associated with bibliometrics. It would perhaps be more correct to talk of an intersubjectivity underlying rules that, once set, do not change. Beyond the fact that their foundations to some extent echo the principles of peer review, these rules face another problem: they cannot capture the individual characteristics of, and subtle differences between, the various types of publication being evaluated. For example, the article by Watson and Crick that revolutionized the natural sciences was only one page long. The methods used by some economists are too rigid to be able to factor in this type of situation.

4.2.3 Citation analysis

Researchers have mastered the process of citation analysis of articles published in academic journals covered by the Thomson ISI databases, although it needs a little refining for SSH purposes. It does, however, present a number of problems when applied to a variety of types of publication. This section outlines the methods that some authors propose for dealing with this and the difficulties involved.

Lewison (2001) uses a citation count to evaluate the performance of the U.K. in the history of medicine, a discipline in which equal value is placed on books and articles, with a large proportion of articles in fact being book reviews. A book review obviously involves a citation to the book and is therefore an indication of interest in that book. On this basis, Lewison evaluated the U.K.'s performance by counting the number of book reviews and the number of citations received by books in SSCI. The results were compared with those for the U.S. and validated by a survey of experts. By way of complementing Lewison's method, Nicolaisen (2002) developed tools for measuring the scholarliness of book reviews. Another study by Lewison (2004) focuses on citations to books.

Two studies have endeavoured to approach the problem more holistically by archiving and analysing monograph-to-monograph and article-to-article citations in the SSH. The first was conducted by the library scientist Line (1981), who processed 59,000 citations in the social sciences, including 11,000 from 300 monographs and 48,000 from 140 journals of U.S. and U.K. origin. He was unable to cross-tabulate citations between the two types of documents owing to significant differences that would have compromised the validity of the results. In a more recent study, Cronin, Snyder and Atkins (1997) followed the same procedure for sociology publications, processing 30,000 citations from monographs. Here too, the authors did not cross-tabulate data on citations between the two types of publication.

Both experiments were nonetheless useful in that they led to a better understanding of existing obstacles to the analysis of citations from a set of documents including at the very least monographs and academic journals. According to Line, reliable analysis is impossible with manually processed data. Cronin, Snyder and Atkins draw a similar conclusion, stating that, in the absence of a database containing large-scale information on monograph citations, their data cannot be considered very reliable. The process of building a database with 30,000 monograph citations proved to be very difficult and time-consuming for the authors.

Ideally, to produce a complete inventory facilitating in-depth citation analysis in the SSH, the researcher must have access to databases containing information on citations from both journals and monographs. These databases should include books-to-book, journal-to-book, book-to-journal and journal-to-journal citations. Of course, such a database does not exist. Given the great amount of work that was required to build a database with only 30,000 citations in just one discipline, it is hard to imagine the level of resources and the amount of time that would be needed to build a complete, multidisciplinary database in the SSH. Another consideration is that monograph citations are not standardized in the way that article citations are. So even more effort would be required.

In short, large-scale citation analysis cannot yet be applied effectively to bibliographic populations made up of several types of documents.

4.3 Co-citation analysis, co-word analysis and bibliographic coupling

Research maps produced by means of co-citation analysis, bibliographic coupling and co-word analysis are another means of evaluating the strengths and weaknesses of actors at various levels (Noyons 2001). The CWTS group in the Netherlands is currently working on the development of strategic maps for visualizing the structure of a field, the institutions and organizations in it, and their performances.

Noyons, Moed and Luwel (1999) propose a method for measuring the performance of countries, universities and departments that combines co-citation-based mapping, analysis of output (articles), and analysis of impact (citations). The first step is to count the number of articles by the scholar, institution or country concerned. The publications are divided into specialties using a “self-organized” method. The specific dynamics of each specialty are then identified and compared with those of other institutions or countries active in the same fields. Van Raan (2003) uses co-citation analysis in a similar way.

Yet co-citation analysis, co-word analysis and bibliographic coupling cannot be used to evaluate performance as such. For this purpose, maps have to be combined with publication counts and citation analyses. These indicators are then juxtaposed with the map, but they could just as well be used on their own. Mapping enables you to visualize the cognitive structure of research fields, but the information it provides has little effect on the evaluation of national performance. It is more useful, however, for the identification of emerging fields, which is discussed in the next section of this report.

4.4 Parameters

The parameters affecting bibliometric methods are outlined below.

4.4.1 Levels of analysis

Most of the methods of bibliometric evaluation presented above are applicable at a number of levels of analysis: the individual scholar, the university department (or a research group in an institution), the institution, the region, the country and the international level. While this report focuses primarily on measurement of national performance, the evaluation of institutions and departments often proves useful, especially for ascertaining a country's strengths and weaknesses. The level of an analysis can also be the project or the program.

The main consideration with respect to level of analysis is the amount of data available, which must be large enough for evaluation results to be reliable. According to Van Raan (1998), bibliometric evaluation of scholars is problematic, particularly in the SSH, where numbers of publications and citations are too low (yet Norris and Oppenheimer (2003) use a bibliometric evaluation method). Evaluation of research groups or departments, which publish an average of 10 to 20 articles annually, would be more appropriate (van Raan 2003). Where the amount of data at a particular level is too small, it is preferable to move up to the next level of aggregation, as long as it meets the purpose of the project.

When benchmarking institutional or departmental performances, an actor's productivity can be compared with that of regional, national and international competitors. In this context, Nederhof and Noyons (1992b) say that it is important to draw comparisons at several levels. In a country where performances in a given discipline are uniformly strong, an institute performing very strongly at the international level may rank very low nationally. An evaluation based on national comparisons alone would not capture this difference. Accordingly, Nederhof and Noyons suggest comparing the performance of a given department with those of the departments with the best results in the discipline internationally. However, the language question would have to taken into consideration if such an approach were taken.

4.4.2 Citation windows

The time period, or window, covered by a bibliometric analysis is a parameter that impacts significantly on the validity and usefulness of the results. It has even more influence on citation analysis. An article is published just once, but the number of citations it receives builds up over the years. The key question is this: When do you stop counting them? In general, what period of time should bibliometric evaluations cover in order to be reliable? As has already been said, too short a window may undermine the validity of the performance measurement, but too long a window may reduce its usefulness for science policy purposes.

While a two-year window seems adequate for the NSE, a window of four to five years would be more appropriate for the SSH. Measurement windows referred to in the literature covered in this report ranged from two to five years. Thomson ISI uses the shortest period, calculating its impact factor on the basis of citations accumulated by journal articles over the two years following their publication (CNER 2002; Uzun 1998). Such a window is considered too short for SSH evaluation (CNER 2002).

Villagra Rubio (1992) prefers to compare the performances of universities according to average annual productivity over mutually exclusive three-year periods. Glänzel (1996) also uses a three-year (1990-92) window for benchmarking national performances in six social science disciplines. Unlike Villagra Rubio, Glänzel calculates the total number of articles published and citations received over the three-year period instead of calculating an average. Thus citations to documents published in 1990 are calculated from 1990 to 1992, citations to those published in 1991 from 1991 to 1993, and citations to those published in 1992 from 1992 to 1994.

Nederhof and Noyons (1992b) propose calculating the citation window on the basis of the specific characteristics of the research field being analysed. They chose three years for experimental psychology but considered four years more suitable for general linguistics. In another article (Nederhof and Noyons 1992a) they propose counting the number of citations in a specific year following publication of an article, selecting the third year for their study.

In an attempt to solve some problems regarding the citation window and the slow pace at which citations are accumulated, van Raan (1998 and 2003) proposes evaluating performance on the basis of overlapping five-year blocks. For example, for an analysis of performance over the period 1987-96, blocks are formed from 1987 to 1991, from 1988 to 1992, and so on up to 1992-96. The researcher calculates total numbers of articles published in each block and also calculates the number of citations received by the articles during each specific block. According to van Raan, this method helps solve the problems of low numbers and slow accumulation. It also facilitates the analysis of very recent periods, since citations are not counted beyond the measurement blocks. However, the amount of accumulated data hides the fact that citations to articles published in the last two years of the last block are not counted. Note also that the analysis cannot cover a period shorter than eight years, because of the window covered by the blocks. On the one hand, the method may require more effort to collect data than the ones described earlier; on the other, it is as fine-grained as they are, if not more so.

According to van Raan (1998 and 2003), the problems and dilemmas faced when deciding on a citation window reflect an important fact. Scholarly research is a process that needs time to produce results and disseminate them. The performance of a research group or country cannot be evaluated over a period of one or two years, whether by bibliometric methods or by peer review. Van Raan feels that a five-year window is always preferable, since it ensures that the analysis will be relevant for science policy purposes and yield reliable results.

4.4.3 Self-citation

Self-citation practices can sometimes appear to be an impediment to bibliometric analysis. In general, number of citations received and impact are calculated with and without self-citations. Specific measures are taken when the self-citation rate is too high. Van Leeuwen, van der Wurff and van Raan (2001) consider self-citation rates of over 50% and under 20% to be problematic. Actors with a self-citation rate of over 50% are over-evaluated in bibliometric analysis, while those with a rate under 20% are under-evaluated. Self-citation can be excluded from counts in fields where such problems occur too frequently.

Overall, self-citation does not seem to be a major obstacle to bibliometric evaluation. Recent research has shown that self-citation rates in the SSH are generally lower than in the natural sciences. Furthermore, there is no correlation between international visibility and frequency of self-citation (Glänzel and Thijs 2004).

4.5 Recommendations

In selecting a method for benchmarking national performances, the researcher must take into account the limitations identified in Section 3. The most important one for national performance is the variety of knowledge dissemination media. For a number of SSH disciplines, any analysis based solely on academic articles will be incomplete. That being said, factoring in other publication media can be problematic too.

A manual analysis of several types of publication will require harnessing significant resources, particularly if national performances are being compared. The use of several specialized databases could solve the problem, since a number of such databases contain data on books, theses and other types of publication. However, few studies of specialized databases have been conducted, so the validity of their data and the reliability of their coverage have yet to be proven.

Regarding methods of analysis, the only viable one for evaluating national performances on the basis of heterogeneous sets of documents is the publication count. An exhaustive, complete citation analysis covering books, articles, conference reports and other document types, even in a single field, would be extremely labour-intensive, time-consuming and expensive.

For fields where the analysis of articles alone is sufficient, SSCI is still the best database to use. It is the only one that can be used for citation analysis, which can be a particularly productive method. Since there is no equivalent of the *Journal Citation Reports* for the humanities, it may be wise to calculate specific impact factors for the disciplines concerned, using a much wider citation window than the one Thomson ISI uses for the JCRs covering the NSE and the social sciences — for example, five years instead of two. For analysis as such, the battery of indicators used by van Raan (2003), combined with the specialization index, will serve to identify national strengths and weaknesses.

5 Bibliometric methods for identifying emerging SSH research fields

The identification of emerging research themes and fields is vital for science policy making. The knowledge generated by the process helps to set priorities for an organization like a granting council and may enable a country or region to stay at the cutting edge of scientific research. This explains why emerging research fields have recently been the focus of attention from some granting councils (Bertrand, Gingras et Archambault 2003). This section outlines the various bibliometric methods for identifying emerging research themes and fields.

5.1 General considerations

There are few studies focusing specifically on the emergence of new research fields and the use of bibliometric methods to identify them, and most of the articles on the subject deal with the NSE.

The various methods for identifying emerging fields do not identify them at the same stages of development. Some identify them only at the point where they become institutionalized, while others will do so when the field is at a fairly low level of development but has reached a higher "noise" level. It is important to define the stage of development at which it is possible to identify a specialty so that the appropriate method can be selected. Bibliometric methods can confirm experts' intuition or help to discover specialties that the academic community was generally unaware of. This is a matter that needs to be considered when selecting an identification method.

5.2 Characteristics of an emerging field

Before outlining these methods, it will be helpful to give the characteristics of an emerging field.

5.2.1 Development

Drawing in particular on the work of Thomas Kuhn, Morris (2004) has provided a precise and accurate description of how a new field emerges. The start of the process usually follows the release of publications announcing new discoveries or new analysis schemes. These publications form the core around which a new specialty develops. Subsequently other publications present results of experiments designed to test the validity of the theory in question. References are made to other disciplines and represent instances of "knowledge or method loans." The theory is then developed and expanded, and further publications are produced, with the result that it is disseminated and popularized, applied, and eventually consolidated and standardized. Publications of discoveries serve as exemplars, i.e., according to Kuhn's theory, as models for solving scientific problems specific to the speciality concerned. Exemplars are of necessity heavily cited because they serve as the cognitive and scientific foundations of the discipline. It follows that, by identifying publications that serve as exemplars, it is then possible to identify emerging fields. Small and Griffith (1974) agree with Morris on the role he attributes to these seminal publications. In their view, a specialty is defined by a few key articles published early in its history.

An emerging field cannot be accepted or recognized as valid by experts in a discipline. Research in that field must be published in peripheral academic journals representing a variety of research fields (Debackere and Rappa 1994). This impacts on the identification of emerging fields, because it may well be the case that SSCI and AHCI do not cover those particular journals. This explains why launching a journal devoted to research work in the new field is an important stage, especially when the speciality is multidisciplinary (McCain 1998).

5.2.2 Multidisciplinarity

A number of the most recently established specialties are multidisciplinary, combining concepts, methods and analytical frameworks from a variety of research fields (Etemad and Lee 2003; Fujigaki 2002; Ponzi 2002). Emerging multidisciplinary fields are characterized by the lack of a well-defined core of published works relating specifically to the subject (as stated above, articles tend to be published in peripheral journals). It is therefore harder to define a body of literature for analysis. Articles relevant to a given field may be published in a number of different journals. Coverage of the content by databases may be inadequate, as in the case of single-discipline databases, which provide only partial coverage (McCain and Whitney 1994).

Hollman, Murrey and Homaifar (1991) have drawn a bibliometric profile of insurance, an emerging, multidisciplinary discipline, highlighting its openness and lack of a strong identity or the lack of close ties between senior researchers. In bibliometrics, this situation is reflected in citations being attributed to a number of other disciplines and in research often being published in media other than journal articles. All these characteristics of increasingly multidisciplinary emerging fields must be factored into the selection of an identification method.

5.3 Thomson ISI services

One starting point for identifying emerging fields is Thomson ISI's services, including Research Fronts, in-cites and SCI-BYTES. Their purpose is to identify the scholars and articles with the best performance. ISI uses the number of citations or co-citations received to identify "hot" research topics and fields. However, its services pay little attention to the SSH. In SCI-BYTES, for example, the social sciences as a whole make up only two of the 22 categories, and the two that are used are necessarily catch-all categories. There is no data at all on the humanities. Apart from Research Fronts, on which more SSH-related research should be conducted, the Thomson ISI services cannot form a solid basis for the science policy-relevant identification of emerging fields, particularly with respect to the SSH.

5.4 Direct article and citation counts

Counts may seem the simplest and most direct method of identifying emerging fields. Intuitively, it is reasonable to assume that the annual number of publications in an emerging field is bound to be low and, unless the field made a "false start," the number will tend to increase each year. The same could be said of the total number of citations attributed to the field, which should also grow each year. However, before a number of publications and citations can be assigned to a field, it has to be identified and delineated, and this step lies at the heart of the identification process. While very useful and even necessary, counts merely characterize the rate and type of growth of the speciality under consideration. The methods of co-citation and co-word analysis outlined in Section 4.5 offer a decided advantage in that serve both to delineate and to characterize the field.

Co-citation and co-word analysis can be a preliminary step before counts (Glänzel and Schubert 2003; Glänzel, Schubert, Schoepflin and Czerwon 1999; Laudel 2003; Leydesdorff and Cozzens 1993). There are other possible methods. For example, the judgment of an expert who identifies a few recent and influential articles on a new topic may be the starting point for identifying and delineating an emerging field. Another method is a survey of a number of experts in a discipline. It is also possible to draw up a list of recent, heavily cited articles in a field and then conduct bibliometric monitoring and analysis for all the publications on the list to determine whether a new research theme is emerging around them.

Only a few articles bear on the use of article counts for identifying emerging fields. In a study on the emergence of bioelectronics, Hinze (1994) shows an annual increase in the number of articles in the discipline. Yitzhaki and Shahar (2000) counted articles in alternative medicine, defined as a fast-growing field, combining the count with the application of growth models. Other authors have developed or adapted mathematical models for analysing growth of research literature (Goffman 1971; Egghe and Rao 1992). Egghe and Rao in particular focused on SSH literature growth models. Such models could be used to generate rough estimates of the future growth of specialties, which could be helpful for such processes as priority setting.

Other studies use counts as a means of analysing structural dynamics. Abrahamson and Fairchild (1999) examined fads in management science and analysed quality circles as a speciality. Annual output of articles on this subject increased from a handful to more than one hundred in the space of six years, but 10 years later it had dropped back to a very low figure. Over a 15-year period, quality circles as a specialty emerged, developed exponentially, and then faded away. According to Ponzi and Koenig (2002), fads in management science generally reach their publication peak after five years. It therefore seems possible to make a distinction between an emerging speciality that is here to stay and a fad. The above examples illustrate how counts can be used for identifying and characterizing emerging fields.

5.5 Co-citation analysis, co-word analysis and bibliographic coupling

A number of bibliometricians have developed methods to illustrate the structure of research fields. The goal of the methods is to highlight relationships between articles, journals or authors graphically. Co-citation and coupling could potentially be used to identify emerging fields: "Specialties can be expected to use specific (jargonistic) words and to cite in specialist domains. Citations are known to be even more highly codified than title words" (Leydesdorff 2002).

By analysing specific words and citations, it is possible to bring out the structure and development of scientific fields and thus identify "hot," emerging research topics. An article by Tijssen and van Raan (1994) gives an overview of co-citation analysis and co-word analysis, and a report by Noyons et al. (2000) provides a clear, detailed description of them.

Two distinct stages are involved in co-citation and co-word analysis. The first identifies relationships within the field of concern on the basis of one of the following methods: co-citation analysis, co-word analysis, bibliographic coupling, journal-to-journal citation analysis and co-occurrences of classification codes. Making a choice among these methods is not a random one, because the choice also reveals a specific theoretical perspective on science (Tijssen and van Raan 1994). The second is mapping, which generates a graphic representation of the relationships. The various co-citation and coupling methods and factors to consider when applying them are outlined below.

Co-citation analysis leads to the identification of research fronts illustrated graphically by clusters. A group of heavily cited articles may represent the cognitive core of exemplars around which a specialty is articulated. Broadly speaking, these are the premises established by Thomson ISI in building the Research Fronts database.

Co-citation analysis is the most common coupling method. Several studies include co-citation analyses applied to the SSH (Bayer, Smart and McLaughlin 1990; Eom 1996; McCain 1986; Small and Crane 1979). However, the results are of limited use for science policy purposes. The studies clearly illustrate the structure of SSH disciplines, but the results are not recent or informative enough to be of any use for evaluation. They have only limited relevance to sociological issues and cannot be used to generate projections of developments in the near or more distant future.

Our review of the literature identified only one example of the use of co-citations that may be really useful for detecting emerging fields. Schwechheimer and Winterhager (1999) applied co-citation analysis in a particularly interesting and productive way. After establishing clusters of documents representing the various specialties in a field, they identify the ones with dynamic research fronts. The authors define a dynamic specialty as a research front with a high proportion of recent and cited articles. Once the dynamic specialties have been identified, the articles in them can be ranked by country, institution, etc. for performance evaluation purposes.

Co-word analysis too can be used to identify research fronts, which are represented by clusters of articles sharing a key term in the title or abstract. The method is based on the assumption that articles using the same terms are related on the level of cognitive content (Noyons et al. 2000). Co-word analysis enjoys a significant advantage over co-citation analysis. Because it is completely independent of citations, it can be applied to fields like those in the humanities, where citation data builds up slowly (Tijssen and van Raan 1994). Furthermore, co-citation analysis covers only heavily cited publications, whereas co-word analysis can cover all publications available in a database. On the other hand, co-word analysis cannot be used by itself to identify the seminal articles in a field.

It would be possible to use the method proposed by Schwechheimer and Winterhager (1999) in the context of co-word analysis. Clusters made up of articles with the same keywords could be analysed to determine whether they have a high percentage of recent articles. This would be very efficient because the researcher would not have to wait for citations to accumulate and could therefore apply the method to the most recent publications.

To offset some of the problems identified by Hicks (1987) and outlined in Section 3.2.3, Glänzel and Czerwon (1996) advocate *bibliographic coupling* (Kessler 1963) to identify "hot" research topics. This method can be used to identify documents forming the core of a specialty without having to wait until there is a critical mass of citations. Once the core documents have been identified, they can be analysed to determine recent developments in research topics, major actors, etc.

Journal-to-journal analysis is based on an aggregate of the citations linking articles at the journal level. Journals become institutionalized representations of scientific fields or specialties, and changes in citation data illustrate changes in the structure of science. Journal-to-journal relationships can then be visualized for different time periods and other parameters. An emerging field lacks a stable core of journals in which articles in that field could be published. Inclusion of a new journal in the Thomson ISI databases and the way that inclusion influences journal-to-journal citations may indicate the emergence of a specialty (Leydesdorff, Cozzens and van den Besselaar 1994). Leydesdorff is the leading proponent of this method. He has used it on numerous occasions (Leydesdorff 2002; Leydesdorff, Cozzens and van den Besselaar 1994; van den Besselaar and Leydesdorff 1996; Wagner and Leydesdorff 2003a) and has also applied it to the social sciences (Leydesdorff 2003).

The last method for identifying specialties is the *analysis of co-occurrences of classification codes* assigned to publications by indexes and thesauri. The analytical process is similar to that of co-word analysis. It is a little used method (Hinze 1994; Noyons and van Raan 1998).

Limitations of co-citation/co-word analysis and other coupling methods

The use of co-citation and other coupling methods for identifying emerging SSH research fields poses some problems. Most of them necessarily rely on data from the Thomson ISI databases, and the results are therefore affected by the shortcomings of those databases. As well as suffering from the limitations discussed earlier, SSCI and AHCI report only in first author when attributing citations, with the result that some authors are under-represented (Gmür 2003). In addition, co-citation analysis is not sufficiently refined — it cannot differentiate between research subfields, the level at which many specialties emerge (Gmür 2003). This is particularly true of interdisciplinary fields in the SSH.

Another limitation of the co-citation method is that it was developed initially to characterize the structure and historical development of a field or research group (Eom 1996; Gmür 2003). This limitation is compounded by the fact that, in many instances, it is technically impossible to create series of maps covering time periods that are close to one another and can be compared, especially for the purpose of visualizing the development of fields and the emergence of specialties over time.

While co-citation and other coupling methods seem to have some science policy potential, their value is in many instances limited to sociological and historical issues (Perry and Rice 1998; Ponzi 2002; Rowlands 1999; Small 1977; Zitt and Bassecoulard 1994). In short, their usefulness for making predictions and detecting emerging fields has yet to be proven.

Furthermore, there may not be any citations to an emerging research field, particularly in the SSH. As mentioned above, some authors conclude that emerging fields have been rejected or looked down on by the rest of the academic community (Debackere and Rappa 1994). In these circumstances, the co-citation method may overlook those fields because they do not receive the required number of citations for analysis.

In light of the variety of methods available, the lack of consensus, the warnings issued by Hicks (1987), problems of application to a science policy context and the overall lack of empirical experience in applying the methods to the SSH, co-citation and other coupling methods must be used with due caution for identifying emerging fields. It is nonetheless worth keeping track of developments in their use, since their level of refinement and usefulness may improve over the medium term. In particular, the work of the Centre for Science and Technology Studies (CWTS) in Leiden, the Netherlands, shows considerable promise (see Noyons 2001 for a description of the CWTS approach). Chen's work is promising too (Borner, Chen and Boyack 2001; Chen, Cribbin, Macredie and Morar 2002; Chen and Kuljis 2003; Chen, Paul and O'Keefe 2001). In both cases, the goal of the work is to create user-friendly interfaces making it easier to produce and interpret maps. In the U.S., there is even a patent on a process using co-word analysis to identify "hot" and "cold" research areas (Cardona 2002). There do not yet seem to be any commercial applications for the process, but it would be worthwhile examining the patent in greater depth.

Two approaches can be taken to enhancing the validity of co-citation and other coupling methods. First they can be combined to switch from one level of analysis to another and to compare the results of different methods (Braam, Moed and van Raan 1991a and 1991b; Glänzel and Czerwon 1996; Hinze 1994). Maps can also be validated by experts in the field under study (McCain and Whitney 1994; Peters and van Raan 1993a). Noyons (2001) proposes involving the scholars being evaluated and end users at all levels of analysis.

5.6 Mapping and other visualization methods

Once the data is collected, relationships between the articles, journals and scholars under study can be represented graphically by means of a variety of computer programs and methods. One such method is multidimensional scaling (Tijssen and van Raan 1994). Based on geometric measurements in which the position of each element reflects the strength of its relationships to the other elements, the method generates real maps. The elements are assigned spatial coordinates, and then the configuration of the elements is mapped on the basis of those coordinates. Other potential methods are cluster analysis (Noyons et al. 2000; Small and Griffith 1974) and factorial analysis (Leydesdorff

2003). A number of computer programs can be used, including SPSS and SAS (Noyons et al. 2000). Some researchers have developed their own computer applications to perform the required operations (Noyons 2001).

Example of mapping applied to the humanities

Kreuzman (2001) used the co-citation method to make a map of 62 authors active in philosophy. The target period was 1980–93. The results are presented in Figure 4. The location of an author in one quadrant rather than another is not significant. The map gives a spatial representation of author-to-author relationships.

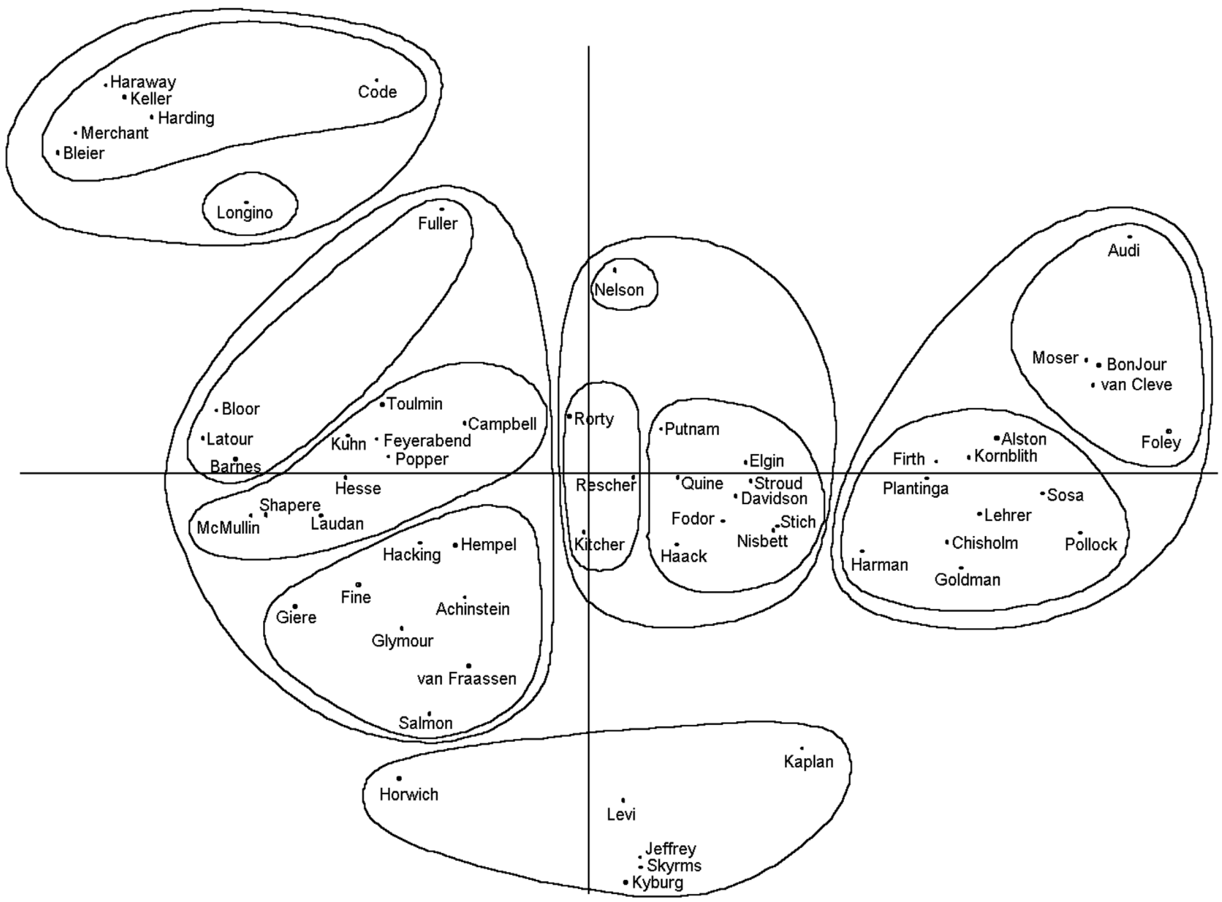

Figure 4 Multidimensional scaling map of 62 philosophers

Source: Kreuzman 2001

5.7 Proposed methods

Our review of the literature shows that co-citation, co-word analysis and bibliographic coupling are the only purely bibliometric methods that can help to identify emerging research themes and fields. As yet, however, they are not particularly suited for science policy purposes and present some shortcomings. That being said, they should not be rejected. Even though the available methods are imperfect, they are still useful for a number of SSH disciplines.

There are at least two promising methods for the identification of emerging fields. Both involve a two-step process of (1) identifying and delineating a speciality and then (2) characterizing its development:

- Identify seminal, core articles and then research fronts by means of bibliographic coupling, co-word analysis or co-citation analysis or a combination of these methods. The advantage of co-word analysis is that it is independent of citations and therefore is not affected by the drawbacks of the Thomson ISI databases and can be applied to the latest articles. The same goes for bibliographic coupling. On the other hand, co-citation analysis seems to be considered more valid by bibliometrics practitioners. Bibliographic coupling may also be inappropriate for multidisciplinary fields because seminal articles may cite references from different fields. If a field is delineated by means of mapping, articles belonging to the resulting sets are classified by year of publication, institution, author, etc. If only the seminal articles are identified, the literature must be monitored first in order to delineate all the articles belonging to the specialty concerned. Based on growth in the number of publications, it is possible to generate an estimate to determine whether the field is emerging, has been stabilized, or is a passing fad. It would then be worthwhile having the results validated by experts.
- One or more experts can identify a set of recent articles on a new subject. A monitoring process is then launched, using the articles as the starting point. The monitoring must detect all exemplars in the field and determine whether documents cited in them belong to the research front. The enlarged set is then analysed. This method can be applied to emerging multidisciplinary fields because it uses citations and monitoring to build its source set of publications.

In the final analysis, there is still a need to determine whether bibliometric methods are more reliable or less reliable and more expensive or less expensive than peer review.

6 Conclusions and recommendations

The preceding sections contained an examination of the application of bibliometrics to the SSH and methods available for mapping and benchmarking SSH research output and identifying emerging fields. The analysis shows that the application of bibliometrics to the SSH is valid and useful as long as its limitations are taken into account. Bibliometric measurements of the SSH serve to identify research dynamics that would otherwise be overlooked by decision makers. SSH bibliometric analysis has some shortcomings, but as van Raan (2003) has pointed out, they are not serious enough to discredit the entire approach. Bibliometrics adds an invaluable empirical content and objectivity to the process of understanding and evaluating the strengths and weaknesses of the knowledge production system in the SSH. In addition, it is virtually indispensable for mapping collaborative research output, and it is very useful for identifying loci of concentration of expertise.

This section presents a summary of the report's findings and recommendations based on those findings.

6.1 Limitations of application of bibliometrics to SSH

In Section 3 the drawbacks of applying bibliometrics to the SSH were identified . The study showed that specific structural factors make SSH research literature different from that of the NSE and make it difficult to apply bibliometric methods to the SSH. They include the following:

- use of dissemination media other than journals
- local orientation of some research
- slower, more limited accumulation of citations

These factors have consequences for the application of bibliometric tools to the SSH, including the following:

- The intrinsic limitations of the databases that have traditionally been used for bibliometrics, i.e., the Thomson ISI databases. Generally speaking, they only cover journals, thereby excluding books, conference proceedings and other publication media, and often neglect literature with a more local orientation.
- The inadequacy of other databases. They could offset some of the shortcomings of the Thomson ISI databases, but they are less exhaustive and are inadequately documented.
- The amount of resources required to conduct a reliable bibliometric analysis that would cover literature of local interest and all research communication media.
- A smaller and therefore less reliable citation pool, in that the law of large numbers applies in fewer cases when data has to be disaggregated.

Consequently, the degree of validity of traditional bibliometric analysis of the SSH will vary by discipline. Therefore an effort must be made to reduce the impact of these limitations as much as possible or to determine to what extent they affect results. This issue is discussed below.

6.2 Extent of limitations

In order to use bibliographic tools for evaluation and mapping in the SSH, it is important to know the extent to which the limitations identified above apply to each SSH discipline. The literature highlights the limitations and proposes at most some partial solutions. No research has been done on quantifying the limitations. The article by Larivière, Archambault, Gingras and Vignola-Gagné (2004) is a first step in this direction. Their data provides a basis for determining the most suitable bibliometric practices for each discipline. Table XI gives the percentage share of citations to research articles by articles covered in the Thomson ISI databases from 1996 to 2000. Given the tendency of books to cite books and articles to cite articles (Cronin, Snyder and Atkins 1997; Line 1981), it can be assumed that the results in the Table are slightly higher than they would be if all types of research publications were taken into account.

Table XI Percentage of citations to journal articles, by discipline, 1996-2000

Discipline	Citations to articles
Psychology and Psychiatry	68%
Law	59%
Economics and management	55%
Other Social Sciences	45%
Education	44%
History	33%
Other Humanities	28%
Literature	22%
SSCI – AHCI Average	**48%**

Source: Larivière, Archambault, Gingras and Vignola-Gagné 2004

In psychology and psychiatry, the percentage of citations to articles is comparable to that of several NSE disciplines, and the same could be said of law and economics and management. For these disciplines, bibliometric evaluation can be based on SSCI and research articles alone and still generate a sound evaluation of research output. Ideally, the results would nonetheless be validated through peer review so as to identify potential biases caused by the failure to consider other publication media and local research.

In "other social sciences" and education, the number of citations to articles is lower than the number of citations to other types of publication, so an SSCI-based bibliometric evaluation would be incomplete. Ideally, the evaluation would have to incorporate data on the other types of publication.

For history, literature and "other humanities," an AHCI-based bibliometric evaluation would identify the dynamics of only a small part of each field. The analysis would still be very relevant and interesting, but it would have to be pointed out, when interpreting the data, that the observations are based on a small portion of the research output and that that portion might not reflect the behaviour of the rest of the output. An evaluation that included data on books would be much more useful in this case.

A number of other questions remain unanswered. For example, it would be very useful for SSH bibliometric evaluation to know the precise proportions of articles, books, conference reports, government reports and other types de publications in each discipline. Ideally, the proportion of locally oriented research should be calculated for each discipline too, with a view to ascertaining whether it is adequately covered by international research databases, particularly those that disseminate their results in English. To produce these results, a large-scale study could be conducted in a number of Canadian and foreign university departments, initially focusing on a few disciplines and then broadening its coverage as needs dictated.

Another component could be included in the project — the characterization of more locally oriented research to determine how much of this type of research was being carried out and how much it influenced bibliometric evaluation. The study would have to calculate the proportion of research of local interest and especially the proportion published in journals with limited distribution.

6.3 Working around existing limitations

In spite of its limitations for SSH analysis, bibliometrics is still a very useful tool. Decision makers can and must use it wisely, following specific guidelines. When working with the existing pool of information services, the first requirement is to determine how bibliometrics should be applied to each SSH discipline. In some cases it may be better to avoid using bibliometric methods for research evaluation if the available resources cannot offset the limitations of the existing databases through the addition of monographs, book chapters, articles, conference proceedings and other relevant bibliographic items. In other cases, bibliometrics could be used, but only publication counts would be valid. For another group of disciplines, articles form an adequate information base for solid bibliometric evaluations. The analysis must also determine to what extent the results generated by bibliometrics can and must be combined with peer review. For example, bibliometric data can be provided to peers to help them carry out their review.

Data sources must be selected in light of the discipline being studied and the methods being used. On the one hand, the Thomson ISI databases make it possible to carry out a number of large-scale bibliometric analyses, but the representativeness will vary by discipline because the databases do not cover types of publication other than the article and do not adequately cover journals from countries other than the U.K., the U.S., the Russian Federation, the Netherlands and Switzerland. On the other hand, databases with better coverage of types of publication other than the article and of literature of local interest cannot often be used for any bibliometric analyses other than publication counts.

These recommendations apply mainly to national performance benchmarking. For the identification of emerging SSH fields, the only methods that have been extensively developed so far are applicable only to journal articles. Peer review and monitoring can be used to identify emerging fields in disciplines in which articles play a lesser role, but they are not bibliometric methods. Even when based on the analysis of journal articles alone, bibliometric methods for identifying emerging fields, i.e., co-citation/co-word analysis and bibliographic coupling, do not seem to have proven their worth, regardless of the issues of local orientation and the small citation pool. The co-citation, co-word and bibliographic coupling methods are worth further examination and development, but Science-Metrix does not recommend their use for collective decision making until they have been properly adjusted for SSH applications.

Because coverage of the various disciplines in bibliometric databases varies widely, comparisons between disciplines for norm-referenced evaluation should be avoided. Of course, it would be interesting to use bibliometrics to gain a greater understanding of structural differences between fields, but it would not be appropriate for comparing the performances of scholars in different fields. Similarly, researchers must exercise due caution when processing data in aggregates encompassing, for example, all the social sciences or all the humanities or, worse still, all the SSH.

6.4 Future applications of bibliometrics in SSH

The findings and recommendations above concern the application of bibliometrics to the SSH in their present state. It is very possible that characteristics and specific properties of the SSH will change and that the changes will influence the conditions for applying bibliometrics. It is also very likely that the existing databases will evolve and that new data sources will become available.

Two attitudes can be adopted toward these changes: a passive one and a proactive one. Each of them is explored below.

6.4.1 The laissez faire or passive attitude

One can take a laissez faire attitude to potential changes in the SSH. It would be perfectly legitimate because a number of indicators suggest that the SSH and the databases are gradually evolving into forms that would make it more effective to conduct bibliometric evaluation. Science-Metrix has identified three possible changes.

Changes in fields

Section 3.2.1 showed that the general trend is for the results of SSH research to be disseminated increasingly by means of articles at the expense of other types of publication. The importance of locally oriented SSH research could decline too. Data from the Academy of Finland, for example, indicates that the proportion of research published in international journals is rising at the expense of research published in journals with a more local distribution. Thus the two factors with the greatest influence on bibliometric analysis in the SSH may be gradually fading in importance. There is therefore reason to believe that SSH bibliometric analysis will become more representative over time. Over the medium or long term, all the social sciences could potentially be evaluated on the basis of SSCI data. However, the situation is likely to be very different in the humanities, where the role of journals is not growing and may even be declining.

Changes to existing databases

It may be that, in light of criticism and in an effort to enhance coverage, the existing databases will expand their coverage. For example, Thomson ISI could modify its evaluation criteria and include more journals written in languages other than English. Complementary databases could be merged or connected and data input could be coordinated. One significant fact militating in favour of expanding Thomson ISI coverage is that data input costs have dropped dramatically in recent years with the advent of electronic data interchange (EDI). It will cost Thomson ISI very little to expand coverage. In addition, the possibility that a new player will arrive on the information services market and thereby change the amount of available data should not be excluded.

In this regard, a project was recently launched to build a European database on the humanities to complement AHCI. The European Science Foundation has been mandated to build the database, called the European Citation Index in the Humanities (ECIH) (Kiefer et al. 2004). If completed, the database would be very useful for bibliometrics (see the May 2004 issue of the *Lettre d'information SHS* published by the CNRS for information that complements this report and covers specific issues of SSH knowledge dissemination).

Open access

If open access continues to grow, it may well bring radical changes to processes that are currently part and parcel of any bibliometric analysis. Self-archiving within institutional archives is the most promising avenue in this context. Under the model targeted by the Open Archive Initiative, each institution would have its own archive in which its scholars would deposit their articles. All the archives would follow the same standards to facilitate coordination and liaison among archives and thereby create a meta-archive. Tools can be used (and some already are being used) to conduct research across all these archives and, with a few refinements, to collect the data required for bibliometric analysis. A system of this type would solve the problem of covering locally oriented research because the only limitation on the meta-archive's coverage would be the level of participation of the institutions concerned. Alternatively, the proliferation of "gold" journals may help to increase citation numbers, although it will not change the existing model to any significant extent.

Open access should not have much effect on the different types of publication media. It primarily affects academic journals, which in some disciplines constitute a minor publication medium. Banks (2004) claims nonetheless that the current interest in open access could potentially be exploited to promote indexing, processing and archiving of grey literature, particularly official government documents. Just as open access increases dissemination of and access to journals, so indexing of the literature makes it easier to locate and thus easier to access. It could also be claimed that design and processing work, the setting of standards and the development of technological innovations for establishing open access archives may be a starting point for eventually indexing books and monographs and the associated citations. For example, the experience gained in coordinating a number of archives (e.g., through the OAI) could eventually be used to link up and coordinate many databases covering individual SSH disciplines, or even SSCI and AHCI, with a database on books and grey literature.

Open access is not a magic solution for problems in applying bibliometrics to the SSH. The model will probably not work for books and conference proceedings, because the need for a financial return limits opportunities for free distribution. Furthermore, a number of technical, economic and political obstacles may limit development of this option over the short term. That being said, open access may have a very positive influence on bibliometrics and the science system in general. Therefore Science-Metrix recommends that SSHRC keep close track of developments in open access and even fund research on it in order to gain a better understanding of the issues revolving around open access and the potential consequences of adopting it.

6.4.2 The proactive attitude

To offset the limitations of bibliometric methods, it is possible to take direct action instead of waiting for the developments outlined above to take concrete shape. Two possible actions are discussed below.

Building databases

The most effective way of eliminating the problems of applying bibliometrics to the SSH is to take action at the source. A database may contain all the information needed for a complete bibliometric analysis. It may include data on books, conference proceedings and academic articles from all countries, along with data on citations and institutional addresses. Its scope is, of course, limited by the resources available for data input. It can be built ad hoc for a specific evaluation project or be built for the longer term.

Experience shows that building new databases is a long, laborious and expensive process (Cronin, Snyder and Atkins 1997; Godin, Archambault and Vallières 2000). Yet the option should not be cast aside too hastily. Recent developments in the new information and communications technologies have probably made it easier and less costly to build and maintain a database. In general, the traditional manual indexing of bibliographic information can now be replaced with electronic processing of information from publishers and editors. An alternative to building new databases is to standardize, merge and coordinate existing ones.

An exhaustive process of reflection should precede the creation of a database and should take account of a variety of factors: the current scope and reliability of bibliometric analyses, the amount of potential increase in scope and reliability, and the cost of the human and financial resources and time required to achieve the increase.

Reforming the Common CV System

As mentioned in Section 3.4, Canada has a tool that offers great potential but is under-utilized because of some design flaws.It is the Common CV System, a repository of Canadian scholars' curricula vitae. If this database were standardized and cleaned up, it could provide some very useful tombstone data on the various types of research publications and the local orientation of SSH research, in addition to providing data for publication counts. It could be used to determine, for each discipline, the exact percentage share of research publications taken up by articles, books, conference proceedings and other contributions. Unfortunately, it could be used only for evaluations of Canadian research output, unless other countries were to develop similar systems. In fact, Canada could become a promoter of such a tool.

SSHRC is strategically positioned to ensure that the System is completed, standardized and updated so as to be able to play a more significant role in Canadian research evaluation and mapping. Therefore Science-Metrix recommends that SSHRC take action to ensure that the System can be exploited to its full potential.

6.5 Recommendations

Science-Metrix advises SSHRC to implement the following recommendations on the use and development of bibliometric tools for SSH research evaluation.

Recommendation 1. Assign bibliometrics-based mapping and evaluation work only to qualified entities. Organizations specializing in the use of bibliometrics are very familiar with the limitations of their tools and know how to interpret results with due care and caution. In particular, projects involving bibliometrics must entail explaining how the following variables affect study results:

- What types of publication (articles, books, etc.) are used in the discipline under consideration and what is the rate of coverage of these media in the information sources used?
- How are the indicators used in the study (count, citation/co-citation/co-word analysis, bibliographic coupling) affected by the internal variables of each discipline and the specific characteristics of the data sources used?

Recommendation 2. Promote research on determining the specific characteristics of SSH publication practices, and particularly on the following aspects:

- The proportion of the literature in each discipline represented by journal articles, monographs, book chapters, conference proceedings and other document types;
- The proportion of the literature in each discipline devoted to topics of more local interest and the proportion of research published in publications with limited distribution;
- The size and scope of the pool of citations from both books and articles.

Recommendation 3. Promote research on devising quantitative methods for identifying emerging fields and on methods for tracking their development.

Recommendation 4. (A) Promote research on understanding the influence of open access publications and self-archiving on trends and developments in SSH knowledge dissemination media;

and (B) promote research on using data available through open access for research mapping and evaluation.

Recommendation 5. Play a proactive role in formatting data in the Common CV System so that it can be used to map and evaluate Canadian research.

Appendix 1: Databases of potential use for SSH bibliometric analysis

1. ABELL Online: Annual Bibliography of English Language and Literature
2. ABI/INFORM Global
3. American Humanities Index Online
4. L'Année philologique
5. Anthropological Index Online
6. Anthropological Literature
7. Archivia
8. Art Abstracts
9. ARTbibliographies Modern
10. Arts and Humanities Citation Index
11. ArticleFirst
12. ATLA Religion Database
13. Avery Index to Architectural Periodicals
14. BHInet - British Humanities Index
15. BL Online: the bibliographical database of linguistics
16. CBCA Reference
17. CIAO: Columbia International Affairs Online
18. CISTI Source
19. Communication Abstracts
20. CSA Worldwide Political Science Abstracts
21. Design and Applied Arts Index (DAAI)
22. Doctrinal
23. Econlit
24. Education Full Text
25. eHRAF Collection of Archaeology
26. eHRAF Collection of Ethnography
27. Emerald Intelligence
28. ERIC
29. Film Index International
30. Francis
31. Global Books in Print
32. Historical Abstracts
33. History of Science, Technology and Medicine
34. Index to Canadian Legal Literature
35. Index to Foreign Legal Periodicals
36. Index to Law School Theses and Dissertations
37. Index to Legal Periodicals
38. Ingenta
39. International Bibliography of the Social Sciences
40. International Index to Music Periodicals, Full Text Edition
41. International Index to the Performing Arts
42. International Political Science Abstracts
43. LegalTrac
44. Library Literature
45. Library & Information Science Abstracts (LISA)
46. Linguistics and Language Behavior Abstracts (LLBA)
47. MLA Directory of Periodicals
48. Music Index Online
49. PAIS International
50. Philosopher's Index
51. Population Index
52. Proquest Digital Dissertations (PQDD)
53. PsycINFO
54. Risk Abstracts
55. Social Sciences Citation Index
56. Social Work Abstracts
57. Sociological Abstracts (SocioFile)
58. Wilson Humanities Abstracts
59. Wilson Social Sciences Abstracts
60. World History FullTEXT

Appendix 2: Methods

Analyses of the coverage provided by the Thomson ISI databases are by and large based on a comparison of the journals included in them with those in the Ulrich directory. The lists of journals in the Thomson ISI databases — the Sciences Citation Index (SCI), the Sciences Citation Index Expanded (SCI_E), the Social Sciences Citation Index (SSCI) and the Arts and Humanities Citation Index (AHCI) — were taken from the Thomson ISI website (http://www.isinet.com/journals/) as of June 3, 2004. Data on the journals in the Ulrich directory was taken from the 2004 second edition of the CD-ROM of *Ulrich's International Periodicals Directory* (ISSN 0000-0507).

Journals in the two sources were matched by means of their ISSN number. Over 95% of the Thomson ISI journals were matched with the Ulrich database. Since 93.5% of the matched ISI journals are peer-reviewed (SCI = 98.6%; SSCI = 97.4%; AHCI = 73.4%), the analysis in this report is based only on peer-reviewed journals in the Thomson ISI databases and in Ulrich.

Language of editor and language of journal are directly coded in Ulrich. More than one language is provided for some journals. In such cases, equal fractions have been assigned to all the main languages of the journal concerned. Thus for a journal that mainly contains texts in English and French and only occasionally includes Spanish texts, 0.5 goes into the “English” column and 0.5 into the “French” column.

For the analysis of the languages of editors’ countries, the first language attributed to countries in the *World Factbook* (http://www.cia.gov/cia/publications/factbook/index.html) was used.

Bibliography

Abrahamson E. and Fairchild G. 1999. Management fashion: lifecycles, triggers, and collective learning processes. *Administrative Science Quarterly*. 44: 708-740.

Academy of Finland. 2003. *Scientific Research in Finland: A Review of Its Quality and Impact in the Early 2000s.* Publications of the Academy of Finland.

Almind T. and Ingwersen P. 1997. Informetric analyses on the World Wide Web: Methodological approaches to "Webometrics". *Journal of Documentation*. 53: 404-426.

Andersen H.1996. Acta Sociologica pa den internationale arena. *Dansk Sociologi*. 7(2): 72-79.

Andersen H. 2000. Influence and reputation in the social sciences - how much do researchers agree? *Journal of Documentation.* 56(6): 674-692.

Baltagi B.H. 1999. Applied Econometrics Rankings: 1989-1995. *Journal of Applied Econometrics*. 14: 423-441.

Banks M. 2004. Connections between open access publishing and access to gray literature. *Journal of the Medical Library Association*. 92(2): 164–166.

Bayer A.E., Smart J.C. and McLaughlin G.W. 1990. Mapping Intellectual Structure of a Scientific Subfield through Author Cocitations *Journal of the American Society for Information Science*. 41(6): 444-452.

Bertrand F., Gingras Y. and Archambault É. 2003. *Examen des programmes et des pratiques d'évaluation des projets de recherche novateurs et interdisciplinaires*. Research report presented to FQRNT by Science-Metrix and CIRST.

Björneborn L. and Ingwersen P. 2001. Perspective of webometrics. *Scientometrics*. 50(1): 65-82.

Börner K., Chen C. and Boyack K.W. 2002. Visualizing knowledge domains. *Annual Review of Information Science and Technology* 37: 179-255.

Braam R.Ft., Moed H.F. and van Raan A.F.J. 1991a. Mapping of Science by Combined Co-Citation and Word Analysis. I. Structural Aspects. *Journal of the American Society for Information Science*. 42(4): 233-251.

Braam R.Ft., Moed H.F. and van Raan A.F.J. 1991b. Mapping of Science by Combined Co-Citation and Word Analysis. II. Dynamical Aspects. *Journal of the American Society for Information Science*. 42(4): 252-266.

Brody T., Stamerjohanns H., Vallières F., Harnad S., Gingras Y. and Oppenheim C. 2004. The effect of open access on citation impact. Online. http://www.ecs.soton.ac.uk/%7Eharnad/Temp/OA-TAadvantage.pdf.

Cardona C. *System and method for Database Retrieval, Indexing and Statistical Analysis*. Patent issued May 7, 2002. Number: US 6,385,611 B1. USPTO: Alexandria, Virginia.

Chen C., Cribbin T., Macredie R. and Morar S. 2002. Visualizing and tracking the growth of competing paradigms: Two case studies. *Journal of the American Society for Information Science and Technology*. 53(8): 678-689.

Chen C.M. and Kuljis J. 2003. The rising landscape: A visual exploration of superstring revolutions in physics. *Journal of the American Society for Information Science and Technology*. 54(5): 435-446.

Chen C., Paul R.J. and O'Keefe B. 2001. Fitting the jigsaw of citation: Information visualization in domain analysis. *Journal of the American Society for Information Science*. 52(4): 315–330.

Clemens E.S., Powell W.W., McIlwaine K. and Okamoto D. 1995. Careers in Print: Books, Journals and Scholarly Reputations. *American Journal of Sociology*. 101(2): 433-494.

Social Science and Humanities Research Council of Canada (SSHRC). 2004. From Granting Council to Knowledge Council: Renewing the Social Sciences and Humanities in Canada. Ottawa. 27 pages.

Comité national d'évaluation de la recherche (CNER). 2002. *Évaluation de la recherche publique dans les établissements publics français.* Paris.

Community Research and Development Information Service (CORDIS). 2004. Commission launches review of Europe's scientific publications system. *CORDIS focus*. 248: 5.

Coupé T. 2003. Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1990-2000. *Journal of the European Economic Association*. 1(6): 1309-1345.

Cronin B., Snyder H. and Atkins H. 1997. Comparative Citation Rankings of Authors in Monographic and Journal Literature: A Study of Sociology. *Journal of Documentation.* 53(3): 263-273.

Debackere K. and Glänzel W. 2004. Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key. *Scientometrics*. 59(2): 253-276.

Debackere K. and Rappa M.A. 1994. Institutional variations in problem choice and persistence among scientists in an emerging field. *Research Policy*. 23: 425–441.

Dusansky R. and Vernon C.J. 1998. Ranking of U.S. Economics Departments. *Journal of Economic Perspectives*. 12(1): 157-170.

Egghe L. 2000. New informetric aspects of the Internet: Some reflections, many problems. *Journal of Information Science*. 26(5): 329–335.

Egghe. L. and Ravichandra Rao I.K. 1992. Classification of Growth models based on growth rates and its applications. *Scientometrics.* 25(1): 5-46.

Eom S.B. 1996. Mapping the intellectual structure of research in decision support systems through author cocitation analysis (1971-1993). *Decision Support Systems*. 16(4): 315-338.

Etemad H. and Lee Y. 2003. The knowledge network of International Entrepreneurship: Theory and evidence. *Small Business Economics*. 20(1): 5-23.

Filiatreau G. et al. 2003. Analyse bibliométrique de la recherche en génomique au cours de la décennie 90. Paris: Observatoire des sciences et des techniques and Laboratoire de traitement du signal et de l'image, Inserm, Université Rennes I.

Finkenstaedt T. 1990. Measuring research performance in the humanities. *Scientometrics.* 19: 409-417.

Fujigaki Y. 2002. Analysis on dynamics of research sub-domains in interdisciplinary fields: Analysis using personal distribution versus papers. *Scientometrics.* 54 (1): 63-74.

Garfield E. 1979. Most-cited authors in the Arts and Humanities, 1977-1978. Current Contents. 32 In Garfield E. *Essays of an information scientist*. Volume 4: 238-244. Philadelphia: Institute for Scientific Information.

Garfield E. 1996. A bibliometric analysis of science journals in the ISI database. *Science Citation Index. Journal Citation Reports*. Printed guide to the microfiche edition of the SCI JCR.

Gingras Y. 1984. La valeur d'une langue dans un champ scientifique. *Recherches Sociographiques*. 25(2): 286-296.

Gingras Y. 1995. Performance Indicators: Keeping the Black Box Open. *Proceedings of the Second International Symposium on Research Funding*. Ottawa, Canada, 13-15 September 1995.

Gingras Y. 2002. Les formes spécifiques de l'internationalité du champ scientifique. *Actes de la recherche en sciences sociales*. 141-142: 31-45.

Glänzel W. 1996. A bibliometric approach to social sciences. National research performance in 6 selected social science areas 1990-1992. *Scientometrics.* 35(3): 291-307.

Glänzel W. and Czerwon H.J. 1996. A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*. 37(2): 195-221.

Glänzel W. and Schoepflin U. 1999 A Bibliometric Study of Reference Literature in the Sciences and Social Sciences. *Information Processing and Management*. 35: 31-44.

Glänzel W. and Schubert A. 2003. A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics.* 56(3): 357-367.

Glänzel W., Schubert A., Schoepflin U. and Czerwon. 1999. An item-by-item subject classification of papers published in journals covered by the SSCI database using reference analysis. *Scientometrics*. 46(3): 431–441.

Glänzel W. and Thijs B. 2004. World flash on basic research - the influence of author self-citations on bibliometric macro indicators. *Scientometrics.* 59(3): 281-310.

Gmür M. 2003. Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics*. 57(1): 27-57.

Godin B. 2002. *The Social Sciences in Canada: What Can We Learn From Bibliometrics? Project on the Measurement of the Social Sciences*. Working Paper No. 1.

Godin B., Archambault E. and Vallières F. 2000. La production scientifique québécoise : mesure basée sur la BRSQ. *Argus.* 29(1): 15-24.

Goffman W. 1971. Mathematical method for analyzing the growth of a scientific discipline. *Journal of the Association of Computing Machinery.* 18(2): 173-185.

Grayson L. and Gomersall A. 2003. *A difficult business: finding the evidence for social science reviews*. ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 19.

Harnad S., Brody T., Vallières F., Carr L., Hitchcock S., Gingras Y., Oppenheim C., Stamerjohanns H. and Hilf E.R. 2004a. The green and the gold roads to open access. *Nature Focus*. Online. http://www.nature.com/nature/focus/accessdebate/.

Harnad S., Brody T., Vallières F., Carr L., Hitchcock S., Gingras Y., Oppenheim C., Stamerjohanns H. and Hilf E.R. 2004b. The Access/Impact Problem and the Green and Gold Roads to Open Access. Online. http://www.ecs.soton.ac.uk/%7Eharnad/Temp/impact.html

Hicks D. 1987. Limitations of co-citation analysis as a tool for science policy. *Social Studies of Science*. 17: 295-316.

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

Hicks D. 2004. The Four Literatures of Social Science. Forthcoming in *Handbook of Quantitative Science and Technology Research*. Ed.: Henk Moed. Dordrecht: Kluwer Academic.

Hinze S. 1994. Bibliographical cartography of an emerging interdisciplinary scientific field: The case of bioelectronics. *Scientometrics*. 29: 353–376.

Ho K.K. 1998. Research output among the three faculties of business, education, humanities & social sciences in six Hong Kong universities. *Higher Education*. 36(2): 195-208.

Hodgson G.M. and Rothman H. 1999. The Editors and Authors of Economics Journals: A Case of Institutional Oligopoly? *The Economic Journal*. 109: F165-F186.

Hollman K.W., Murrey J.H. and Homaifar G. 1991. The Structure and Disciplinary Boundaries of Insurance – a Citational Analysis of JRI Articles. *Journal of Risk and Insurance.* 58(4): 714-721.

Ingwersen P. 1997. *The Central International Visibility of Danish and Scandinavian Research 1988-1996: A General Overview of Science & Technology, the Humanities and Social Sciences by Online Publication Analysis*.

CIS Report 5.3. Centre for Informetric Studies, Royal School of Library and Information Science, Denmark.

Ingwersen P. 1998. The calculation of Web Impact Factors. *Journal of Documentation*. 54: 236-243.

Ingwersen P. 2000. The international visibility and citation impact of Scandinavian research articles in selected Social Science fields: The decay of a myth. *Scientometrics.* 49(1): 39-61.

Ingwersen P., Larsen B. and Noyons E. 2001. Mapping national research profiles in social science disciplines. *Journal of Documentation*. 57(6): 715-740.

Ingwersen P. and Wormell I. 1999. Publication behavior and international impact: Scandinavian clinical and social medicine, 1988-1996. *Scientometrics.* 46: 487-499.

Jeannin P. 2002. *Pour une « revuemétrie » de la recherche en Sciences sociales : Un point de vue français*. Communication to CNRS scientific council, Paris, 1-2 July 2002.

Jeannin P. 2003. *Revuemétrie de la recherche en sciences humaines et sociales*. Report presented to Minister with Responsibility for Research and New Technologies, France.

Kalaitzidakis P., Mamuneas T.P. and Stengos T. 2002. *Rankings of Academic Journals and Institutions in Economics*. Discussion Papers in Economics 01/8, Department of Economics, University of Leicester.

Katz J.S. and Hicks D. 1998. *Indicators for Systems of Innovation - a Bibliometrics-based Approach*. IDEA Paper.

Katz J.S. 1999. *Bibliometric Indicators and the Social Sciences*. Prepared for the Economic and Social Research Council (ESRC), U.K.

Kessler M.M. 1963. Bibliographic coupling between scientific papers. *American Documentation*. 14:10-25.

Kiefer F., Mustajoki A., Peyraube A. and Vestergaard E. 2004. Building a European Citation Index in Humanities (ESF). *Lettre d'information SHS.* 69: 8-10.

King J. 1987. A review of bibliometric and other science indicators and their role in research evaluation. *Journal of Information Science*. 13: 261-276.

Kreuzman H. 2001. A Co-Citation Analysis of Representative Authors in Philosophy: Examining the Relationship Between Epistemologists and Philosophers of Science. *Scientometrics*. 51(3): 525-539.

Kurtz M.J. 2004. *Restrictive Access Policies Cut Readership of Electronic Research Journal Articles by a Factor of Two*. Working Paper - http://opcit.eprints.org/feb19oa/kurtz.pdf.

Kyvik S. 1988. Internationality of the social sciences: The Norwegian case. *International Social Sciences Journal*. 115: 163-172.

Laband D.N. and Piette M.J. 1994. The Relative Impacts of Economics Journals: 1970-1990. *Journal of Economics Literature*. 32: 640-666.

Larivière V., Archambault É., Gingras Y. and Vignola-Gagné E. 2004. The Place of Serials in Referencing Practices: Comparing Natural Sciences and Engineering with Social Sciences and Humanities. Submitted to *Journal of the American Society for Information Science and Technology*.

Larivière V. and Godin B. 2001. *Profil de la production scientifique du Service canadien des forêts*. Report prepared for Canadian Forest Service.

Larson R. 1996. Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace. In Hardin S. (Ed.) *Proceedings of the 59th Annual Meeting of the American Society for Information Science*. 33: 71-78.

Laudel G. 2003. Study the brain drain: Can bibliometric methods help? *Scientometrics.* 57(2): 215-237.

Lawrence S. 2001. Online or Invisible? Online. http://www.neci.nec.com/~lawrence/papers/online-nature01/. Edited version published in *Nature*. 411(6837): 521.

Lewison G. 2001. Evaluation of books as research outputs in history of medicine. *Research Evaluation*. 10(2): 89-95.

Lewison G. 2004. James Bond and citations to his books. *Scientometrics*. 59(3): 311-320.

Leydesdorff L. 2002. Indicators of Structural Change in the Dynamics of Science: Entropy Statistics of the SCI Journal Citation Reports. *Scientometrics.* 53(1): 131-159.

Leydesdorff L. 2003. Can Networks of Journal-Journal Citation be used as Indicators of Change in the Social Sciences? *Journal of Documentation*. 59(1): 84-104.

Leydesdorff L. and Cozzens S.E. 1993. The delineation of specialties in terms of journals using the dynamic journal set of the Science Citation Index. *Scientometric*s. 26: 133-154.

Leydesdorff L., Cozzens S. and van den Besselaar P. 1994. Tracking areas of strategic importance using scientometric journal mappings. *Research Policy*. 23: 217-229.

Li X. 2003. A review of the development and application of the Web impact factor. *Online Information Review*. 27(6): 407-417.

Li X., Thelwall M., Musgrove P. and Wilkinson D. 2003. The relationship between the WIFs or inlinks of Computer Science Departments in UK and their RAE ratings or research productivities in 2001. *Scientometrics*. 57(2): 239–255.

Line M.B. 1981. The structure of social science literature as shown by a large-scale citation analysis. *Social Science Information Studies.* 1: 67-87.

Line M.B. 1999. *Social Science information - the poor relation*. 65th IFLA Council and General Conference.

Lucas R.F. 1995. Contributions to economics journals by the Canadian economics profession, 1981-90. *Canadian Journal of Economics*. 28(4): 945-960.

Luwel M., Moed H.F., Nederhof A.J., De Samblanx V., Verbrugghen K. and Van Derwurff L.J. 1999. Towards indicators of research performance in the social sciences and humanities. An exploratory study in the fields of Law and Linguistics at Flemish Universities. Brussels: Vlaamse Interuniversitaire Raad.

McCain K.W. 1986. Cocited author mapping as a valid representation of intellectual structure. *Journal of the American Society for Information Science.* 37(3): 111-122.

McCain K.W. 1998. Neural networks research in context: A longitudinal journal cocitation analysis of an emerging interdisciplinary field. *Scientometrics*. 41(3): 389–410.

McCain K.W. and Whitney P.J. 1994. Contrasting assessments of interdisciplinarity in emerging specialties, the case of neural network research. *Knowledge: Creation, Diffusion, Utilization*. 15: 285-306.

Moed H.F., Nederhof A.J. and Luwel M. 2002. Towards performance in the humanities. *Library Trends.* 50: 498-520.

Morris S.A. 2004. *Manifestation of emerging specialties in journal literature: a growth model of papers, references, exemplars, bibliographic coupling, co-citation and clustering coefficient distribution*. Working Paper.

Nederhof A.J., Luwel M. and Moed H.F. 2001. Assessing the quality of scholarly journals in linguistics: An alternative to citation-based journal impact factors. *Scientometrics*. 51(1): 241-265.

Nederhof A.J. and Noyons E.C.M. 1990. *Trends in publication activity and in transnational co-publications in the social and behavioral sciences and the humanities*. Leiden: Center for Studies of Science and Technology (CWTS).

Nederhof A.J. and Noyons E.C.M. 1992a Assessment of the International Standing of University Departments Research – A Comparison of Bibliometric Methods. *Scientometrics*. 24(3): 393-404.

Nederhof A.J. and Noyons E.C.M. 1992b. International comparison of departments' research performance in the humanities. *Journal of the American Society for Information Science*. 43(3): 249–256.

Nederhof A.J. and van Raan A.F.J. 1993. A bibliometric analysis of six economics research groups: A comparison with peer review. *Research Policy*. 22: 353-368.

Nederhof A.J. and Zwaan R.A. 1991. Quality Judgments of Journals as Indicators of Research Performance in the Humanities and the Social and Behavioral Sciences. *Journal of the American Society for Information Science*. 42(5): 332-340.

Nederhof A.J., Zwaan R.A., Debruin R.E. and Dekker P.J. 1989. Assessing the Usefulness of Bibliometric Indicators for the Humanities and the Social and Behavioral Sciences: A Comparative Study. *Scientometrics* 15(5-6): 423-435.

Nicolaisen J. 2002. The scholarliness of published peer reviews: a bibliometric study of book reviews in selected social science fields. *Research Evaluation*. 11(3): 129-140.

Norris M. and Oppenheim C. 2003. Citation counts and the Research Assessment Exercise V - Archaeology and the 2001 RAE. *Journal of Documentation*. 59: 709-730.

Noyons E.C.M. 2001. Bibliometric mapping of science in a science policy context. *Scientometrics*. 50(1): 83-98.

Noyons E.C.M., Buter R.K., van Raan A.F.J., Schwechheimer H., Winterhager M. and Weingart P. 2000. *The Role of Europe in World-Wide Science and Technology: Monitoring and Evaluation in a Context of Global Competition*. Report for the European Commission. Leiden: Universiteit Leiden.

Noyons E.C.M., Moed H.F. and Luwel M. 1999. Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. *Journal of the American Society for Information Science.* 50(2): 115-131.

Noyons E.C.M. and van Raan A.F.J. 1998. Monitoring scientific developments from a dynamic perspective: Self-organized structuring to map neural network research. *Journal of the American Society for Information Science.* 49(1): 68–81.

Odlyzko A. 2002. The rapid evolution of scholarly communication. *Learned Publishing*. 15: 7-19.

Perry C.A. and Rice R.E. 1998. Scholarly communication in developmental dyslexia: Influence of network structure on change in a hybrid problem area. *Journal of the American Society for Information Science.* 49(2): 151-168.

Peters H.P.F. and van Raan A.F.J. 1993a. Co-word-based science maps of chemical engineering I: Representations by Direct Multidimensional-Scaling. *Research Policy*. 22: 23–45.

Pierce S. 1987. Characteristics of professional knowledge structures: Some theoretical implications of citation studies. *Library & Information Science Research*. 9: 143-171.

Ponzi L.J. 2002. The intellectual structure and interdisciplinary breadth of Knowledge Management: A bibliometric study of its early stage of development. *Scientometrics*. 55(2): 259-272.

Ponzi L.J. and Koenig M. 2002. Knowledge management: another management fad? *Information Research*. 8(1). Working Paper No. 145.

Prime P., Bassecoulard E. and Zitt M. 2002. Co-citations and co-sitations: A cautionary view on an analogy. *Scientometrics*. 54(2): 291 - 308.

Rousseau R. 1997. Sitations: an exploratory study. *Cybermetrics*. 1(1): paper 1.

Rowlands I. 1999. Patterns of author cocitation in information policy: Evidence of social, collaborative and cognitive structure. *Scientometrics*. 44(3): 533-546.

Royle P. and Over R. 1994. The Use of Bibliometric Indicators to Measure the Research Productivity of Australian Academics. *Australian Academic & Research Libraries*. 25(2): 77-88.

Schoepflin U. 1992. Problems of Representativity in the Social Sciences Citation Index. In Weingart P., Sehringer R. and Winterhager M. (eds.). *Representations of Science and Technology: Proceedings of the International Conference on Science and Technology Indicators, Bielefeld, 10-12 June 1990*. Leiden: DSWO-Press. 177-188.

Schwechheimer H. and Winterhager M. 1999. Highly dynamic specialties in climate research. *Scientometrics*. 44 (3): 547-560.

Small H. 1977. A Co-Citation Model of a Scientific Specialty: A Longitudinal Study of Collagen Research. *Social Studies of Science*. 7(2): 139-166.

Small H. and Crane D. 1979. Specialties and Disciplines in Science and Social Science: An Examination of Their Structure Using Citation Indexes. *Scientometrics*. 1(5-6): 445-461.

Small H. and Griffith B.C. 1974. The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies.* 4: 17–40.

Smith A.G. and Thelwall M. 2001. Web impact factors and university research links. *Proceedings of the 8th International Conference on Scientometrics and Informetrics*. 2: 657-664. Sydney, Australia. 16-20 July 2001.

So C.Y.K. 1998. Citation ranking versus expert judgement in evaluating communication scholars: effects of research specialty size and individual prominence. *Scientometrics.* 41(3): 325-333.

Soete L.G. and Wyatt S. M. E. 1983. The use of foreign patenting as an internationally comparable science and technology output indicator. *Scientometrics*. 5(1): 31-54.

Testa J. and McVeigh M.E. 2004. *The Impact of Open access Journals: A Citation Study from Thomson ISI*. Prepared by Thomson ISI.

Thelwall M. 2001. Extracting Macroscopic Information from Web Links. *Journal of the American Society for Information Science and Technology*. 52(13): 1157-1168.

Thelwall M. and Harries G. 2004. Do the Web Sites of Higher Rated Scholars Have Significantly More Online Impact? *Journal of the American Society for Information Science and Technology*. 55(2): 149-159.

Thelwall M. and Wilkinson D. 2004. Finding similar academic Web sites with links, bibliometric couplings and colinks. *Information Processing & Management*. 40: 515–526.

Tijssen R.J.W. and van Raan A.F.J. 1994. Mapping Changes in Science and Technology – Bibliometric Co-citation Analysis of the R and D Literature. *Evaluation Review*. 18(1): 98-115.

Uzun A. 1998. A scientometric profile of social sciences research in Turkey. *The International Information & Library Review*. 30(3): 169-184.

Van den Besselaar P. and Leydesdorff L. 1996. Mapping Change in Scientific Specialties: A Scientometric Reconstruction of the Development of Artificial Intelligence. *Journal of the American Society for Information Science.* 47(6): 415-436.

Van der Meulen B. and Leydesdorff L. 1991. Has the Study of Philosophy at Dutch Universities Changed Under Economic and Political Pressures? *Science, Technology & Human Values.* 16(3): 288-321.

Van Leeuwen T.N., van der Wurff L.J. and van Raan A.F.J. 2001. The use of combined bibliometric methods in research funding policy. *Research Evaluation*. 10(3): 195-201.

Van Raan A.F.J. 1998. Assessment of social sciences: The use of advanced bibliometric methods as a necessary complement of peer review. *Research Evaluation.* 7: 2–6.

Van Raan. A.F.J. 2001. Bibliometrics and internet: Some observations and expectations. *Scientometrics*. 50(1): 59–63.

Van Raan A.F.J. 2003. The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung.* 12(1): 20-29. English translation available: http://www.itas.fzk.de/tatup/031/raan03a.htm

Varki A. 2003. Open access: the JCI has already shown it works. *Nature*. 426: 383.

Villagra Rubio A. 1992. Scientific Production of Spanish Universities in the Fields of Social Sciences and Language. *Scientometrics*. 24(1): 3-19.

Wagner C.S. and Leydesdorff L. 2003. Seismology as a dynamic, distributed area of scientific research. *Scientometrics.* 58(1): 91-114.

Webster B.M. 1998. Polish Sociology Citation Index as an Example of Usage of National Citation Indexes in Scientometric Analysis of Social Science. *Journal of Information Science.* 24(1): 19-32.

Winclawska B.M. 1996. Polish Sociology Citation Index (Principles for Creation and the First Results). *Scientometrics*. 35(3): 387-391.

Yitzhaki M. 1998. The Language Preference in Sociology. *Scientometrics.* 41(1-2): 243-254.

Yitzhaki M. and Shahar T. 2000. *The emergence of alternative medicine as a new field: a bibliometric study of a rapidly-growing field*. 66th IFLA Council and General Conference, Jerusalem.

Zitt M and Bassecoulard E. 1994. Development of a Method for Detection and Trend Analysis of Research Fronts Built by Lexical or Cocitation Analysis. *Scientometrics*. 30(1): 333-351.

Zwaan R.A. and Nederhof A.J. 1990. Some Aspects of Scholarly Communication in Linguistics: An Empirical Study. *Language*. 66: 523-727.