CITATION ANALYSIS IN RESEARCH EVALUATION

Information Science and Knowledge Management

Volume 9

The titles published in this series are listed at the end of this volume.

CITATION ANALYSIS IN RESEARCH EVALUATION

HENK F. MOED

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-3713-9 (HB)
ISBN-13 978-1-4020-3713-9 (HB)
ISBN-10 1-4020-3714-7 (e-book)
ISBN-13 978-1-4020-37146 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springeronline.com

Printed on acid-free paper

Contents

Preface

This book is written for members of the scholarly research community, and for persons involved in research evaluation and research policy. More specifically, it is directed towards the following four main groups of readers:

- All scientists and scholars who have been or will be subjected to a quantitative assessment of research performance using citation analysis.
- Research policy makers and managers who wish to become conversant with the basic features of citation analysis, and about its potentialities and limitations.
- Members of peer review committees and other evaluators, who consider the use of citation analysis as a tool in their assessments.
- Practitioners and students in the field of quantitative science and technology studies, informetrics, and library and information science.

Citation analysis involves the construction and application of a series of indicators of the 'impact', 'influence' or 'quality' of scholarly work, derived from citation data, i.e. data on references cited in footnotes or bibliographies of scholarly research publications. Such indicators are applied both in the study of scholarly communication and in the assessment of research performance. The term 'scholarly' comprises all domains of science and scholarship, including not only those fields that are normally denoted as science – the natural and life sciences, mathematical and technical sciences – but also social sciences and humanities.

The term 'research policy' in this book is used in a broad sense, and comprises policies at various levels: science policy of a national government by ministers responsible for scholarly research; research policy at the level of research organisations or institutions dealing with quality control and the

allocation of research funds; and research management, carried out by directors of research groups or departments, including hiring, promoting and retaining individual scholars.

This book deals with a crucial aspect of research performance: the contribution of scholarly work to the advancement of scholarly knowledge. Its principal question is: how can citation analysis be used properly as a tool in the assessment of such a contribution? Although the major part of the analysis relates to the basic science – a domain in which citation analysis is used most frequently – this book also addresses its uses and limits in the applied and technical sciences, social sciences and humanities.

It provides a wide range of important 'facts', and corrects a number of common misunderstandings about citation analysis. It introduces basic notions and distinctions, and deals both with theoretical and technical aspects, and with its applicability in various contexts, at the level of individual scholars, research groups, departments, institutions, national scholarly systems, disciplines or subfields, and scholarly journals. It reveals the enormous potential of quantitative, bibliometric analyses of the scholarly literature for a deeper understanding of scholarly activity and performance, and highlights their policy relevance. But this book is also critical, underlines the limits of citation analysis in research evaluation, and issues warnings for potential misuse. It proposes criteria for proper use of citation analysis as a research evaluation tool.

It describes primarily the use of data extracted from *the Science Citation Index* and related Citation Indexes published by the *Institute for Scientific Information* (ISI). Although this institute's name was recently changed to *Thomson Scientific*, its original name is still used throughout this book. It focuses on the use of the ISI Citation Indexes in the study of the scholarly communication system, and particularly in research evaluation, rather than on their outstanding usefulness in scholarly literature retrieval. But many aspects to which this book dedicates attention relate to citation analysis in general, regardless from which databases the analysed bibliographical data is extracted.

Quantitative studies of science and technology is a rapidly developing field, and its development is closely linked to a number of general tendencies in the global scholarly system. National governments and research organisations and institutions need systematic evaluations for optimising their research allocations, re-orienting their research support, rationalising research organisations, restructuring research in particular fields, or augmenting research productivity. *Evaluative bibliometrics* is a subfield of quantitative science and technology studies, aimed to construct indicators of research performance from a quantitative analysis of scholarly documents. Citation analysis is one of its key methodologies.

Whereas in the USA companies such as *ISI/Thomson Scientific* and *CHI Research* were the primary contributors to evaluative bibliometric methodologies, in other countries, particularly in Europe and in Australia, these were mainly further developed in academic research departments. The aim of this book is to further contribute to an *academic basis* for evaluative bibliometrics, by presenting it as a multi-disciplinary scientific–scholarly activity, with its own methodologies and theoretical debates.

Applying citation analysis in research evaluation in a proper way requires a high level of competence. It is not something that anyone with access to the ISI Citation Indexes can do easily. On the contrary, this book illustrates that one needs detailed technical knowledge and theoretical understanding to carry out citation analysis properly. The book does not, however, provide a detailed technical manual of how to carry out citation analysis. Instead, it focuses on its main lines, basic principles and assumptions, and its uses and limits in the various domains of scholarship.

It is up to the members of the scholarly community and the policy arena, and *not* to the author, to decide whether or not citation analysis is to be used for evaluative purposes. In order to make such a decision properly, it is essential that all participants have insight into the nature of citation analysis, how its indicators are constructed and calculated, what the various theoretical positions state about what they measure, and what are their potentialities and limitations in scholarly research evaluation. This book aims at providing such insight.

Structure of the book

This book presents a number of studies undertaken by the author, some in collaboration with his colleagues at the Centre for Science and Technology Studies (CWTS) at Leiden University (the Netherlands), and building upon the earlier works of many other scholars in the field. The following table provides an overview of the general structure, and indicates the primary target audience for each part.

Part	*Primary target audience*
Executive Summary	Research policy makers, managers and evaluators
Part 1: General Introduction and Main Conclusions	Interested scholars from all domains of scholarship
Part 2: Empirical and Theoretical Chapters	Practitioners and students in quantitative science studies

Part 1 provides an introduction, and presents the basic notions and main conclusions as regards the use of citation analysis in research evaluation. Moreover, it includes a synopsis providing summaries of all later chapters, each briefly introducing the main topics and conclusions in each chapter, but excluding most of the technical details. Part 1 aims at presenting the reader, particularly those who are not active in the field of quantitative science studies, what this book is all about and its conclusions.

Part 2 presents more detailed conclusions of 24 empirical and theoretical chapters, arranged into 8 sub-parts (Parts 2.1–2.8):

Part Number	*Part Title*
2.1	Assessing basic science research departments and scientific journals
2.2	The ISI Citation Indexes
2.3	Assessing social sciences and humanities
2.4	Accuracy aspects
2.5	Theoretical aspects
2.6	Citation analysis and peer review
2.7	Macro studies
2.8	New developments

Part 2.1 provides a further introduction to later parts, discussing a large number of issues as regards the types of citation analysis that are applied most frequently: the assessment of the past performance of research departments in basic science, and the measurement of journal impact using journal impact factors and related citation measures.

Acknowledgements

I wish to thank the following colleagues for their most valuable comments on draft versions of particular chapters: Maarten Coolen, University of Amsterdam, the Netherlands (Chapter 2); Henry Small, Thomson Scientific, USA (Chapters 6, 7, 12, 13, 14); Paul Wouters, Royal Netherlands Academy of Arts and Sciences (Chapters 15 and 16); Liv Langfeldt, Norwegian Institute for Studies in Research and Higher Education (Chapter 18); Marc Luwel, Ministry of the Flemish Community, Belgium (Chapters 19 and 20); Jan Reedijk, Leiden University, the Netherlands (Chapters 19 and 21); Rolf Lehming, National Science Foundation, USA (Chapter 22); and Abraham Bookstein, University of Chicago, USA (Chapter 24).

In addition I am grateful to the following co-authors of journal articles that played an important role in this book: Anthony van Raan, CWTS (Chapter 4); Thed van Leeuwen, CWTS (Chapters 4 and 5); Renger de Bruin, former CWTS, Centraal Museum, Utrecht, the Netherlands (Chapter

4); Jan Reedijk, Leiden University, the Netherlands (Chapter 5); Wolfgang Glänzel, Steunpunt O&O Statistieken, Leuven, Belgium (Chapter 5); Martijn Visser, CWTS (Chapters 8 and 10); Marc Luwel, Ministry of the Flemish Community, Belgium (Chapters 4 and 11); and Anton Nederhof, CWTS (Chapter 11).

I am also grateful to my employing institution (Leiden University) and to Anthony van Raan (Director of CWTS) for giving me the opportunity to write this book and to use CWTS resources, to Suze van der Luijt (CWTS) for clerical assistance in its preparation, and to Peter Negenborn and Erik van Wijk (CWTS) for their efforts in creating and maintaining at CWTS the bibliometric database version of the ISI Citation Indexes applied in numerous chapters.

A substantial part of this book is based upon a series of lectures I delivered in 2003 and 2004 at the CWTS Graduate Course on Measuring Science. I acknowledge the Ministry of the Flemish Community, The Netherlands Organisation for Scientific Research (NWO) and the scientific publishing company Elsevier for funding parts of the work described in this book.

Finally I wish to express my gratitude to Eugene Garfield (Chairman-Emeritus, Institute for Scientific Information, USA), for his strong support during the preparation of this book, for writing a note on accuracy issues included in Chapter 12, and for his most valuable comments on the book's structure and on all chapters.

EXECUTIVE SUMMARY

This book is about scientific, or more generally, scholarly research. It focuses on a type of research that is characterised as 'basic', 'fundamental' or 'strategic'. It recognises its crucial importance for global economic progress and social welfare, but at the same time it acknowledges that a firm political or societal basis for this type of research can be maintained only by further developing a system of internal quality control and performance enhancement. This book aims at showing that citation analysis is a useful tool in such a system.

It primarily concerns the assessment of the contributions scholars make in their research publications to the advancement of valid scholarly knowledge. It deals with the assessment of research performance of individual scholars, research groups, departments and institutions, scholarly journals and national scholarly systems, and with the analysis of general characteristics of global science and scholarship.

It explores the uses and limits of citation analysis, involving the construction and application of a wide range of 'bibliometric' indicators of the 'impact', 'influence' or 'quality' of scholarly work, and derived from citation data, i.e. data on references cited in footnotes or bibliographies of scholarly research papers. It focuses on the Citation Indexes produced by the Institute for Scientific Information (ISI, currently Thomson Scientific), but many findings are also relevant in the use of other citation indexes.

This book aims to provide useful information for members of the scholarly community and research policy officials about basic technical aspects of citation analysis, what it measures, and how it can be properly applied in research evaluation and policy processes, by systematically discussing numerous statements about its value made by scholars and policy makers, correcting misunderstandings and illustrating its strengths and limits, particularly in relation to peer review.

It is argued that the use of citation analysis in the evaluation of individuals, groups and institutions is more appropriate the more it is:

- Formal – i.e., previously known to evaluators or decision makers and to scholars or institutions subjected to evaluation that indicators are used as one of the sources of information.
- Open – those subjected to the bibliometric analysis have the opportunity to examine the accuracy of underlying data, and to provide background information that in their view is relevant for a proper interpretation of the quantitative outcomes.
- Scholarly founded – that bibliometric investigators present their outcomes within a scholarly framework, discuss issues of validity, explicitly state theoretical assumptions, and underline their potentialities and limits.
- Supplemented with expert and background knowledge about the substantive contents of the work under evaluation, the conditions under which evaluated scholars operated, and their research objectives.
- Carried out in a clear policy context – i.e., applied in the framework of an evaluation procedure of which both the evaluative perspective and the objectives are clear to all participants.
- Stimulating users to explicitly state basic notions of scholarly quality, its dimensions and how they were operationalised and weighted.
- Enlightening rather than formulaic – the indicators are used to obtain insight in a particular aspect addressed in the process, rather than as inputs in formulas designed to algorithmically generate the process' outcomes.

Application of citation analysis in the assessment of past research performance in basic science and of scientific journals has reached a *high level of sophistication*. This book discusses numerous issues raised by scientists subjected to citation analysis, by journal editors and policy makers, and shows how such issues can in principle be accounted for or solved technically.

The extent to which citation analysis based on the ISI Citation Indexes can be validly applied in all domains of scholarship, including the applied and technical sciences, social sciences and humanities, is often debated. This book thoroughly examines differences in the *structure of the written communication systems* among the various domains of scholarship, and the extent to which these systems are *covered by the ISI Citation Indexes*.

The ISI Indexes do not claim to have complete journal coverage, but rather to include the most important. Their founder, Eugene Garfield, developed a powerful and unique criterion for expanding the database beyond the core of journals whose importance to a given field is obvious: the frequency at which journals are cited in those sources that are already included in the index.

Applying a 'database internal' criterion, this book shows that ISI coverage tends to be *excellent* in physics, chemistry, molecular biology and biochemistry, biological sciences related to humans and clinical medicine; *good*, yet not excellent, in applied and engineering sciences, biological sciences related to animals and plants, geosciences, mathematics, psychology and other social sciences related to medicine and health; and *moderate* in other social sciences including sociology, political science, anthropology and educational sciences, and particularly in humanities.

A principal cause of non-excellent coverage is the importance of sources other than international journals, such as books and conference proceedings. In fields with a moderate ISI coverage, language or national barriers play a much greater role than they do in other domains of science and scholarship. In addition, research activities may be fragmented into distinct schools of thought, each with their own 'paradigms'.

This book distinguishes and illustrates *four types of bibliometric studies* in which the ISI database plays different roles. The decision as to which type of study is appropriate in a discipline depends upon the extent to which it is covered by the ISI Indexes. Compared to a 'standard' analysis in fields with excellent coverage, this database may be expanded in several ways in fields with good but not excellent coverage, or it may play a limited role or no role at all when field coverage is moderate.

If the extent to which research findings reach beyond a purely national or local viewpoint and are exposed to criticisms from a wide international scholarly audience is considered as a relevant criterion of research quality in social sciences and humanities, a major task would be to develop for the various subfields valid indicators of this aspect of research performance. This book argues that it cannot be taken for granted that the ISI Citation Indexes provide such indicators in all subfields of these domains of scholarship. A challenge would be to systematically explore alternative data sources and methodologies. The expertise and perceptions of scholars active in the various subfields should play an important role in such an exploration.

As regards *journal impact factors*, this book provides a technical and historical explanation of how ISI impact factors are calculated, and highlights a number of problems affecting their accuracy and applicability. It illustrates how alternative journal impact measures solve many of these problems, but at the same time underlines that there is no single 'perfect' indicator of journal performance. Although the status of the journals in which a research group publishes is an aspect of research performance in its own right, journal impact factors should not be used as surrogates of citation impact of a group's publications.

Data accuracy is a next crucial issue. It is illustrated how uninformed data collection and analysis may substantially distort the outcomes of

citation analysis. Use of inaccurate data may not only distort results for particular groups, but also affect the credibility and hence the usefulness of a bibliometric study as a whole. But accuracy problems can be overcome in advanced data handling and in data verification procedures involving evaluated scholars and their institutions.

The next key issue concerns *what citations measure*. Outcomes of citation analysis of basic science research groups tend to statistically correlate in a positive way with peer ratings of the groups' past performance. This book presents more empirical case studies revealing such a positive correlation. Findings provide a further theoretical justification for applying citation analysis in research evaluation, but correlations are not perfect.

It is argued that citation counts can be conceived as manifestations of intellectual influence, but the concepts of citation impact and intellectual influence do not coincide. Distinct notions of the concept of intellectual influence may exist, and evaluators assessing scholarly work may have different views upon which are the most crucial aspects to be taken into account. Outcomes of citation analysis must be *valued* in terms of a qualitative, evaluative framework that takes into account the substantive contents of the works under evaluation.

The interpretation of citation impact involves a quest for possible biases. It is therefore crucial at which level of aggregation citation analysis is carried out. Evaluating aggregates of entities can be carried out in such a way that the effects of special characteristics and circumstances of individual entities to some extent cancel out. It must be underlined that systematic biases as regards the aggregate as a whole may still occur and should be taken into account.

The conditions for proper use of bibliometric indicators at the level of individual scholars, research groups or departments tend to be more readily satisfied in a *peer review* context than in a policy context. It can therefore be argued that bibliometric analyses at such lower aggregation levels normally best find their way to the policy arena through peer assessments. But it does *not* follow that citation analysis is *merely* a tool to be used by peers.

This book illustrates the use of citation analysis as a tool to *assess peer review procedures* and to keep the peer review process honest. From the latter perspective, it is a tool for policy decision makers as well. It shows that citation analysis has its strengths and limits, and that the same is true for peer review. The challenge is to combine the two methodologies in a proper, productive way.

A study of *research assessment exercises*, in which a small peer committee evaluated research departments in an entire national discipline, raised the question whether such exercises are capable of identifying truly excellent or 'top' research departments. This finding underlines the need for

research policy makers to thoroughly reflect upon the objectives of such exercises, taking into account their cost effectiveness.

This study also provided evidence that a peer rating system (e.g., in terms of 'excellent', 'good', 'less good', 'poor') tends to generate a distribution of ratings among departments that depends upon the rating system itself, and that is to some extent independent of the overall performance level of the departments under evaluation.

A study of funding procedures of a *national research council* provided evidence that proximity relationships between applicants and expert committees responsible for the evaluation of grant proposals made their outcomes inequitable. It illustrates how quantitative, bibliometric methods can fruitfully contribute to an *internal* debate within a funding agency about funding procedures and evaluation criteria, and to a *public* debate between a funding agency and the national science policy sphere.

Citation analysis is a most valuable tool in policy studies addressing general issues regarding the academic system, with a complexity that reaches beyond the capabilities of expert panels. Studies of the global academic system and 'macro' studies of national academic systems are excellent examples. This book presents four studies that deal with 'classical' issues in the field of quantitative science studies and that have a high policy relevance:

- Did scientists' global publication productivity increase during the 1980s and 1990s?
- How to measure trends in national publication output?
- Does international scientific collaboration pay?
- Do US scientists overcite papers from their own country?

A first macro study presented in this book examined trends during the 1980s and 1990s in *global publication productivity*, defined as the total number of articles published in a year per scientist active in that year. It was found that, although an 'average' individual scientist can justly claim to have published in recent years more research articles than in the past, from a global perspective scientific publication productivity did not increase during the past two decades. One interpretation is that raising both the internal productivity of the science system, its economic relevance and collaboration, are to some extent conflicting policy objectives for basic science.

Nowadays many countries publish National Science Indicators Reports and analyse what bibliometric macro indicators express about the state of *a nation's research system*, and about the level of its research performance. Not infrequently, the various indicators and methodologies seem to lead to different conclusions. This makes bibliometric indicators vulnerable to selective use and manipulation. A second macro study presented in this book

provides technical information as regards the construction and interpretation of publication based macro indicators.

Assessing the trend in a single country's publication output, it explores a categorisation of publishing authors into *domestic* (i.e., working in institutions located in the country itself) and *foreign* (active in other countries). Indicators are considered that give an answer to the following questions: did the country's scientific workforce expand or shrink, and did the number of papers in which it participated per domestic author increase or decline? It concludes that it is essential to calculate a *series* of indicators and to provide them with a consistent interpretation. Isolating a single measure from the others may distort the results and lead to biased conclusions.

A third macro study addressed the 'classical' issue '*Does international scientific collaboration pay*?' It concludes that when scientifically advanced countries collaborate with one another, they profit in around 7 out of 10 cases from such bi-lateral collaboration, in the sense that both raise their citation impact compared to that of their 'purely domestic' papers. But when advanced countries contribute in bi-lateral international collaboration to the development of scientifically less advanced countries – and thus to the advancement of science in the longer term than the perspective normally adopted in research evaluation – this activity tends to negatively affect their short-term citation impact, particularly when their role is secondary.

It has been claimed that US authors excessively cite other US colleagues. This would lead to a *US bias* in the selection of journals for the ISI Citation Indexes and would distort the outcomes of citation analysis. This book argues that the crucial issue at stake is the adequacy of the norm against which referencing practices of US scientists is evaluated. A fourth macro study found *no* conclusive evidence that US scientists in science fields excessively cite papers originating from their own country.

Finally, this book discusses *recent trends* in the development of indicators and in scholarly publication. The need is emphasised to carry out systematic studies of the *conditions* under which citation analysis is actually applied in research evaluation, and of the *effects* of its use upon the scholarly community, its evaluators and the policy arena. Such insights may contribute to the further development of the 'critical' potential of citation analysis as a research evaluation tool.

Analyses of changes in publication and citation practices are illuminating, but the principal question is not whether or not scholars' practices change under the influence of the use of bibliometric indicators, but rather whether or not the application of such measures as a research evaluation tool enhances research performance and scholarly progress in general.

As more and more scholarly documents become available in electronic form through the World Wide Web, their use as sources in citation analysis is expected to increase in the near future. From the perspective of research evaluation, including more sources does not necessarily lead to more valid assessments of the contributions scholars make to the advancement of scholarly knowledge. The extent to which the sources' documents contain new knowledge and meet professional quality standards is a critical issue.

Outcomes of citation analysis are often presented to the 'outside world' in the form of *rankings* of entities such as individual scholars, research departments or institutions. This also occurs with outcomes of peer reviews. It is argued that the need for policy makers and the wider public to obtain insight into the scholarly quality of the various groups is legitimate, but that scholarly quality is not as straightforwardly measured and ranked as performance is in many other societal domains. Moreover, rankings disregard how the performance of one entity depends upon that of others. Bibliometric investigators should look for means to express these notions in the outcomes they produce.

PART 1

GENERAL INTRODUCTION AND MAIN CONCLUSIONS

Chapter 1

GENERAL INTRODUCTION

1.1 The ISI Citation Indexes as search and research tools

Eugene Garfield's book *Citation Indexing* starts with the following paragraph that describes the basic concept underlying a journal citation index:

> The concept of citation indexing is simple. Almost all the papers, notes, reviews, corrections and correspondence published in scientific journals contain citations. They cite – generally by title, author and where and when published – documents that support, provide evidence for, illustrate, or elaborate on what the author has to say. Citations are the formal, explicit linkages between papers that have particular points in common. A citation index is built around these linkages. It lists publications that have been cited and identifies the sources of the citations. Anyone conducting a literature search can find from one to dozens of additional papers on a subject just by knowing one that has been cited. And every paper that is found provides a list of new citations with which to continue the search (Garfield, 1979, p. 1.).

Eugene Garfield Associates was founded in 1954 and launched numerous editions of *Current Contents* by 1960. In that year, the company name was changed to the Institute for Scientific Information (ISI). In 1964, ISI launched the *Science Citation Index* (*SCI*), as a quarterly multidisciplinary index covering at that time some 600 scientific journals (Garfield, 1964). From 1964 onwards, the *SCI* expanded rapidly. Subsequently, ISI began publishing indexes covering the social sciences (the *Social Sciences Citation Index*, *SSCI*) and the Arts and Humanities (*Arts and Humanities Citation Index*, *A&HCI*). In 1988, the printed indexes were supplemented by CD-ROM editions, then in 1997, ISI began the *Web of Science*, a comprehensive

citation index made available to subscribers through the Internet, and covering some 7,500 scholarly journals from all areas of scholarship.

Throughout this book the term 'ISI Citation Indexes' is used to denote the various information products based on citation indexing of scholarly literature, produced during the past 50 years by the Institute for Scientific Information. ISI recently changed its name to Thomson Scientific, Inc. The ISI Citation Indexes were designed primarily for the purpose of retrieval and dissemination of scholarly literature. Citation indexing is used to augment traditional natural language (titles) by utilising cited references (citations) as indexing terms. Such use of the author, journal, title, and citation indexing elements can be characterised as *bibliographic*.

Garfield viewed the cited work as symbolic of specific content, such as a method, a concept, a fruitful hypothesis, or specific data. The citing papers one retrieves from a citation index search are assumed to have a subject relevance to the idea symbolised by the cited item targeted for the search. Citations can be viewed as indicators of document content, as document descriptors, or indexing terms.

Once citation indexing became available for bibliographic research, it was apparent that it could be used to answer inquiries into the nature of scholarly activity: how it is structured, how it develops and how its actors perform. Garfield expressed this as follows:

> If the literature of science reflects the activities of science, a comprehensive, multidisciplinary citation index can provide an interesting view of these activities. This view can shed some useful light on both the structure of science and the process of scientific development (Garfield, 1979, p. 62).

It was Derek de Solla Price (1970) who underlined that science on the one hand and humanities on the other are two distinct domains of scholarship with essentially different substantive contents. According to Price, the different substantive contents in science and humanities have "erected different social apparatuses of information pooling and exchange".

> Scholarship is a conspiracy to pool the capabilities of many men, and science is an even more radical conspiracy that structures this pooling so that the totality of this sort of knowledge can grow more rapidly than any individual can move by himself. The humanities, by resting with the capability of the individual, eschew this growth rate and certainty (Price, 1970, p. 6).

He conceived a scholarly publication as not merely a piece of information but also as an expression of "the state of a scholar or group of scholars at a particular time" and hypothesised:

> If the paper is an expression of a person or several persons working at the research front, we can tell something about the relations among the people from the papers themselves (Price, 1970, pp. 6–7).

For a historical account of the creation and application of the ISI Citation Indexes, the reader is referred to Wouters (1999). During the past four decades, hundreds if not thousands of studies have used data from the citation indexes to provide some type of quantitative, statistical analysis. These applications can be denoted as *bibliometric*, as they extract, aggregate and analyse quantitative aspects of bibliographic information. As statistics related to scholarship are applied mainly in the sciences, the term scientometric is also often used.

The use of the *SCI* as a search tool is well documented in numerous publications, for which the book *Citation Indexing* (1979) provides an excellent introduction. While it did discuss various non-bibliometric uses, when the book was published the bibliometric use of *SCI* for evaluation and study of scholarly activities had not yet fully matured. The current book takes into account what has been learned about the bibliometric uses of *SCI*, and in particular, the study of scholarly communication and research performance.

A good example of the use of the ISI citation indexes for studying *the structure of the scholarly communication system* is the following statement by Garfield in which counts of cited references from the *SCI* are related to a concept of 'quality' of a scientific journal.

> Since authors refer to previous material to support, illustrate, or elaborate on a particular point, the act of citing is an expression of the importance of the material. The total number of such expressions is about the most objective measure there is of the material's importance to current research. The number of times all the material in a given journal has been cited is an equally objective and enlightening measure of the quality of a journal as a medium for communicating research results (Garfield, 1979, pp. 23–24).

This statement should be viewed in the context of his original and illuminative studies of the scientific communication system, in which relationships among journals were analysed in terms of citations from one journal to another, and core journals and more peripheral ones were identified. Equally important, his analysis provided the basis for a unique and highly useful 'internal' monitor of the adequacy of coverage of the *SCI* itself.

The journal statistics he derived were soon isolated from the study context and published by ISI in rankings of journals by impact factor, probably the bibliometric construct most widely used in the scholarly and publishing community. Journal impact factors found their way into the arena

of research policy, research management and library collection management. Nowadays they are used to evaluate scholars, to develop publication strategies and to monitor library collections.

Citation counts are used to study not only communication artefacts such as scholarly journals, but also individual scholars, research groups, departments and institutions, scholarly disciplines and entire nations. In the 1960s, sociologists recognised the usefulness of bibliometric statistics in sociological research. In their important paper 'Measuring the quality of sociological research: Problems in the use of the Science Citation Index', Jonathan and Stephen Cole stated:

> The problem of assessing the "quality" of scientific publications has long been a major impediment to progress in the sociology of science … The invention of the Science Citation Index (*SCI*) a few years ago provides a new and reliable tool to measure the significance of individual scientists' contributions … The number of citations an individual receives may be tabulated and used as an indicator of the relative scientific significance or "quality" of that individual's publication … This should lead to major advances in the sociology of sociology (Cole and Cole, 1971, p. 23).

The use of bibliometric data in research performance assessments went far beyond the boundaries of sociological research, but soon entered the policy arena in many countries. The application of performance indicators, however, was – and still is – controversial.

A distinction can be made between two contexts of use of bibliometric data or indicators in the study of scholarly activity: a *scholarly research context*, and a *policy context*. This distinction clearly emerges from the following statements by Stephen Cole:

> A crucial distinction must be made between using citations as a rough indicator of quality among a relatively large sample of scientists and in using citations to measure the quality of a particular individual's work (Cole, 1989, pp. 9, 11). In sociological studies our goal is not to examine individuals but to examine the relationships among variables (ibid., p. 11). Citations are a very good measure of the quality of scientific work for use in sociological studies of science; but because the measure is far from perfect it would be an error to reify it and use it to make individual decisions (ibid., p. 12).

Citation indicators in *a scholarly research context* are used as tools in testing hypotheses or examining universal relationships among variables within a theoretical framework. It is the validity of a particular hypothesis that is at stake. In a *policy context*, citation indicators may be used in reaching some type of policy decision. This decision may relate to an individual, but also to aggregates of individuals such as research groups, institutes or disciplines. Outcomes of citation analysis may have *practical*

consequences for the position of individual scholars and the institutions in which they carry out their research.

1.2 Quantitative science and technology studies

Quantitative studies of science and technology is a rapidly developing field. Its development is closely linked to a number of general tendencies in the global scholarly system. During the past few decades, research institutions have been subjected to new influences and pressures emerging from the increasing need for accountability in scholarly research and training of students.

In most OECD countries, there is an increasing emphasis on the effectiveness and efficiency of government-supported research. Governments need systematic evaluations for optimising their research allocations, re-orienting their research support, rationalising research organisations, restructuring research in particular fields, or augmenting research productivity. In view of this, they have stimulated or imposed evaluation activities.

Universities have become more diverse in structure and are more oriented towards economic and industrial needs. In most member states of the Organisation for Economic Cooperation and Development (OECD) the following trends in the university system were identified (OECD Group on the Science System, 1998).

- Declining government R&D finance: Government research and development (R&D) budgets have been reduced in a number of OECD countries.
- Changing nature of government finance: Government funding for academic research is more and more mission-oriented and contract-based, and more dependent upon performance criteria.
- Increasing industry R&D finance: Private industries are funding an increasing proportion of university research.
- Growing demand for economic relevance: Universities are expected to contribute more and more to their national innovation systems.
- Increasing systemic linkages: Universities are encouraged to enter into joint ventures and co-operative research with industry and other research institutions, in order to improve the effectiveness of networks in national innovation systems.
- Growing research personnel concerns: Both ageing of the workforce and the declining interest in some fields of science by young people in a number of countries raise concerns about the availability of sufficient numbers of well-trained researchers in the future.

– Internationalisation of university research: Globalisation, stemming partly from advances in information and communication technologies (e.g., the Internet), influences research activities and networks.

On the one hand, *quantitative studies of science and technology* explore and apply methodologies enabling policy makers to carry out their research and innovation policies; on the other, they provide tools to critically assess the effectiveness of such policies. As a result, science and technology indicators are becoming increasingly important in research policy. This trend is clearly illustrated in the recently published *Handbook of Quantitative Science and Technology Research* (Moed et al., 2004).

All chapters in the *Handbook* deal with the study of conditions that positively or negatively influence scientific and technological performance, defined in terms of the needs and criteria expressed by the societies in which these systems are embedded. The basic assumption underlying these studies is that one must have a proper insight into how an S&T system works in order to design effective policies aimed at improving its performance. Five broad, partly overlapping themes can be identified that apply bibliometric methodologies and focus on *scholarly research*.

1. *The assessment of the contribution made by various bodies in the scholarly system to the advancement of scholarly knowledge*. Typical examples of such bodies are individual scholars, research groups or departments, research institutions such as universities, and national systems. The contribution to scholarly knowledge does not merely relate to the progress achieved in a particular research specialty, but also to the extent to which it contributes to surrounding research areas. This theme comprises comparative assessments of research performance, and particularly the citation impact of their publications on the international research front (van Raan, 2004a). Studies of research groups and departments play a role in national research assessment exercises of scholarly disciplines (van Leeuwen, 2004b).
2. *Analyses of the global scholarly system*. These analyses comprise studies of various characteristics of the scholarly system and their relationships to research performance, including its internationalisation and globalisation, (Zitt and Bassecoulard, 2004), scholarly collaboration networks (e.g., Glänzel and Schubert, 2004), multi- and inter-disciplinarity (Bordons et al., 2004), the dissemination of scholarly information (Arunachalam, 2004), and the participation of women (Naldi et al., 2004). Several authors applied approaches from statistical physics describing the behaviour of complex physical systems to the science and technology system and use bibliometric data to characterise it (e.g., van Raan, 1990; Katz, 1999; van Raan, 2000; Amaral et al., 2001).

3. *Analyses of scholarly fields*. This theme involves mapping of the structure of scholarly or technological fields or disciplines and their development over time on the basis of quantitative, bibliometric analysis of their literatures. Typical examples are co-citation (e.g., Small, 1973; Small, 1977), co-word (Callon et al., 1983; Noyons, 2004) and author co-citation analysis (White and McCain, 1998). Analysing co-occurrence matrices, such studies aim at identifying and analysing emerging research specialties or 'hot' topics of great strategic or technological importance, their principal actors, and their relationships to other areas of research.
4. *Analyses of the science–technology interface and the economic contributions of science*. This theme focuses on the role of science in innovation processes and on assessments of the economic outputs of basic research. There are many ways to analyse bibliometrically the science–technology interface (Bassecoulard and Zitt, 2004). Combined analysis of scientific publications and patents reveals knowledge networks among academic scientists and industrial researchers. The study of references in patents to the scientific literature sheds light upon the science base of modern technology (e.g., Carpenter and Narin, 1983). Studies of inventors of patents reveal the extent to which basic scientists employed in academic institutions contributed to technological developments (Noyons et al., 2003; Schmoch, 2004; Tijssen, 2004).
5. *Assessment of educational, social and cultural contributions of basic research*. This theme comprises a variety of topics that may be closely related to the economic function of scientific research. To the extent that analyses of documents play a role, these topics include assessments of the citation impact of basic medical research upon medical practice and the wider public, by analysing citations to basic science papers that are given in clinical guidelines, textbooks, government policy documents, international or national regulations and newspaper articles (Lewison, 2004). Other approaches examine the extent to which globalisation of research leads to more general welfare, particularly in developing countries (Arunachalam, 2004; da Motta e Albuquerque, 2004). As regards the contribution of social sciences and humanities, their enlightenment function towards the general public constitutes an important topic (e.g., Nederhof and Zwaan, 1991).

Table 1.1 presents an overview of how *citation analysis* can be used to study the science and technology system and the relationships between science and technology. In the former, citations from the scientific literature are analysed, particularly from journals processed for the ISI Citation Indexes, and in the latter those from the patent literature obtained from

major patent offices such as the *US Patent and Trademark Office* and the *European Patent Office*.

Pioneering work on the analysis of patent citations was carried out by Francis Narin and co-workers (Carpenter and Narin, 1983; Albert et al., 1991; Narin, 1994; Narin et al., 1997). One basic hypothesis underlying their work is that the number of times a patent is cited from other patents provides an indication of its technological importance. Citations from one patent to another are even used to assess the economic value of patents (Sampat and Ziedonis, 2004), knowledge networks in innovation (Breschi and Lissoni, 2004), and a patent holder's stock market performance (Narin et al., 2004).

Table 1.1. The role of citation analysis in the study of the relationships between science and technology

Influencing / cited	*Influenced / citing*	
	Science	*Technology*
Science	Contribution of science groups to scientific progress *Citations in science papers to other science papers (this book)*	The science base of technology *Citations in patents to scientific literature*
Technology	The influence of technology upon scientific development *Citation gap*	Contribution of technologies to technological progress *Citations in patents to other patents*

A second hypothesis holds that citations in patents in a field to the scientific literature reflect that field's science base. This hypothesis was further developed in many subsequent publications (e.g., van Vianen et al., 1990; Schmoch, 1993; Meyer, 2000; Tijssen et al., 2000).

It was Cees Le Pair (1988) who underlined that the influence of technology and instrumentation upon scientific development is not properly reflected in cited references in the scientific literature. Valuable technical products such as the electron microscope are heavily used in the research described in numerous scientific publications (Bakker, 1977). The term *citation gap* indicates that, although publishing authors may mention the product in the full texts of their papers, they do not *cite* it in the papers' reference lists. Citations to patents in the scientific literature are relatively rare (Glänzel and Meyer, 2003) and their significance and usefulness are not yet fully explored.

1.3 Scope and structure of the book

This book concerns primarily the assessment of the contributions made by scholars in their research publications to the advancement of valid scholarly knowledge. In terms of the distinction in main themes presented in the previous section, it focuses on the first and second; in terms of citation relationships among documents, it explores references made in scholarly documents to other scholarly documents. The core of this book deals with the assessment of research performance of individual scholars, research groups, departments, institutions and countries, and with the analysis of general characteristics of the global scholarly system.

Basic research can be defined as the type of research that is primarily carried out to increase scholarly knowledge. Following Salter and Martin (2001), it includes both 'curiosity-driven' – sometimes also denoted as 'pure' – as well as 'strategic' or 'application oriented' research. The latter is undertaken in a quest for a particular application, even though its precise details are not yet known. A large part of this book is dedicated to the use of citation analysis in the assessment of basic research. However, it also addresses its usefulness and limits in the applied and technical sciences. In addition, it focuses on 'science', but also dedicates attention to the social sciences and humanities.

Table 1.2. Classification of scientific–scholarly activities into three broad domains

Aggregate term	*Disciplines (non-exhaustive list)*
Science	Natural sciences, including chemistry, physics, astronomy, geosciences Life sciences, including biological sciences, clinical medicine Mathematics Applied and technical sciences, including engineering
Social sciences	Psychology, psychiatry Economics Sociology, political sciences, education, pedagogical sciences, anthropology
Humanities	Law, literature, language and linguistics, historical sciences, philosophy
Scholarship	All domains of science, social sciences and humanities

The concepts '*science*', '*social science*', '*humanities*' and '*scholarship*' may cover different aggregates of substantive contents in different countries or cultures. In this book the term science is used to indicate research activities in the natural sciences, biological and life sciences, mathematics, and the applied and technical sciences. Social sciences include amongst others psychology, economics and sociology, and humanities comprise

amongst others law, literature, language, history and philosophy. Table 1.2 presents an overview. This classification does *not* fully coincide with the arrangement of disciplines into the three Citation Indexes: *Science Citation Index*, *Social Science Citation Index* and *Arts and Humanities Citation Index*. One of the main differences is that in this book the field of law is categorised as a part of the humanities.

In any assessment of research quality, two fundamental dimensions must be clarified: its time horizon, and its scope. The first relates to the time period taken into account in an assessment of the quality of a piece of work under evaluation. Although historical studies of scholarly development may cover a time period of several decades, in many current research assessment exercises it is often much shorter than that. This book focuses on research assessments adopting a time horizon of 5 to 10 years, but also underlines the relevance of analyses covering longer time periods.

The scope of an assessment can be further specified on the basis of distinctions made by Alvin Weinberg in his classic paper 'Criteria for scientific choice' (Weinberg, 1962). He distinguished between internal and external evaluation criteria. The first relate to the quality of research (and the researchers undertaking it) compared to that of other research activities in the same research (sub)field or specialty. External criteria are generated outside the (sub)field and relate to the question "why pursue this particular science?". Weinberg distinguished technological, social and scientific merit. He sharpened the latter criterion as follows: "that field has the most scientific merit which contributes most heavily to and illuminates most brightly its neighboring scientific disciplines." The core of this book deals with the application of citation analysis in research evaluations based on internal criteria, but it also dedicates attention to its use as a tool to assess the external, scientific merit, as proposed by Weinberg.

It is assumed that in science the *research group* is the 'natural' *unit of scientific activity*. Its scientific staff normally includes a group leader, one or more senior researchers, postdoctoral researchers and several PhD students. Members of research groups tend to interact intensively one with another, and jointly carry out the group's research programme. In many areas of social sciences and humanities the organisational structure of research activities tends to be different from that in science. Scholarly research tends to be more an individual activity. The term *research department* is used to indicate an institutionalised aggregate of research groups or individual scholars covering the same subfield, normally reflected in the departments' name. In science, it may include a single or several research groups.

Citation analysis comprises a variety of ways to analyse references cited in scholarly publications. This book focuses on simple and sophisticated 'counting' of citations to particular sets of scholarly publications, but also on