

The Evaluation of Research and the Evolution of Science Indicators^{*}

Current Science, 2005

Loet Leydesdorff

Université de Lausanne, School of Economics (HEC) &
University of Amsterdam, Amsterdam School of Communications Research (ASCoR)
Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands
loet@leydesdorff.net; <http://www.leydesdorff.net>

Abstract

Research evaluation is based on a representation of the research. Improving the quality of the representations cannot prevent the indicators from being provided with meaning by a receiving discourse different from the research system(s) under study. Since policy decisions affect the systems under study, this configuration generates a tension that has been driving the further development of science indicators since World War II. The article discusses historically the emergence of science indicators and some of the methodological problems involved. More recent developments have been induced by the emergence of the European Union as a supra-national level of policy coordination and by the Internet as a global medium of communication. As science, technology, and innovation policies develop increasingly at various levels and with different objectives, the evaluative discourses can be expected to differentiate with reference to the discourses in which they are enrolled.

Introduction

The use of scientometric indicators in research evaluation emerged in the 1960s and 1970s, first in the United States and then also in various European countries. Before that time, research evaluation had not been formalized other than through the peer review system, on the one hand, and through economic indicators which could only be used at the macro-level of a national system, on the other. The economic indicators (e.g., percentage of GDP spent on R&D) have internationally been developed by the *Organization of Economic Co-operation and Development* (OECD) in Paris. For example, the *Frascati Manual for the Measurement of Scientific and Technical Activities* (1963) can be considered a response to the increased economic importance of science and technology which had become visible in economic statistics during the 1950s.

The idea that scientific knowledge can be organized deliberately and controlled from a mission perspective (for example, for military purposes) was a result of World War II. Before that time the intellectual organization of knowledge had largely been left to the internal mechanisms of discipline formation and specialist communications. The military impact of science and technology through knowledge-based development and mission-oriented research during World War II (e.g., the Manhattan project) made it

^{*} A previous version of this paper was published in Chinese under the title 科研评价和科学计量学的研究纲领——二者关系的历史演变与重新定义 (The evaluation of research and the scientometric research program: historical evolution and redefinitions of the relationship). *科学学研究* (*Studies in Science of Science*), 22(3), 2004, 225-232.

necessary in 1945 to formulate a new science and technology policy under peacetime conditions.

Vannevar Bush's (1945) report to the U.S. President entitled *The Endless Frontier* contained a plea for a return to a liberal organization of science. Quality control should be left to the internal mechanisms of the scientific elite, for example, through the peer review system. The model of the U.S. National Science Foundation (1947)¹ was followed by other Western countries. For example, the Netherlands created its foundation for Fundamental Scientific Research (ZWO) in 1950. With hindsight, one can consider this period as the institutional phase of science policies: the main policy instrument was the support of science with institutions to control its funding.

The Sputnik shock

The launching of Sputnik by the Soviet Union in 1957 turned the tables for science policy. The Soviets—who had used a non-liberal model—seemed to have become more successful than the West in mission-oriented research. President Eisenhower felt the pressure of the alliance between the scientific elite and the military for enlarging the funding of the science system after the Sputnik shock. In addition to his warning (in his farewell speech) against the pressure of a 'military-industrial complex,' he formulated as a less known 'second warning:'

Yet in holding scientific research and discovery in respect, as we should, we must also be alert to the equal and opposite danger that public policy could itself become the captive of a scientific-technological elite. (York, 1970: 9).

A far-reaching reorganization of the American research system was one of the consequences. The National Aeronautics and Space Administration (NASA) and the Advanced Research Projects Agency (ARPA), in particular, were created in response to the launch of Sputnik and the perception of military threat (Edwards, 1996).

During this same period, it became increasingly clear that the continuing growth rates of Western economies could no longer be explained in terms of traditional economic factors such as land, labour, and capital. The 'residue' (Abramowitz, 1956; OECD, 1964) had to be explained in terms of the emerging knowledge-base of the economy (Rosenberg 1976). Alongside the military coordination by NATO, the *Organization for Economic Co-operation and Development (OECD)* was created in 1961 in order to organize and to coordinate science and technology policies among its member states, that is, the advanced industrial nations.² As noted, this led in 1963 to the *Frascati Manual* in which parameters were defined for the statistical monitoring of science and technology on a comparative basis. Comparisons among nation states, however, make it possible to raise questions with respect to strengths and weaknesses in the

¹ Actually, the first N.S.F. act of 1947 was vetoed by President Truman so that the creation of the N.S.F. was postponed until 1950. The later and more detailed "Steelman Report," *Science and Public Policy*, published in 1947, set in motion the eventual role the federal government would play in supporting fundamental research at universities (NSF, 2000).

² The OECD was based on the OEEC, the Organization for European Economic Cooperation, that is, the organization which had served for the distribution of the U.S. and Canadian aid under the Marshall Plan during the postwar period.

underlying portfolios. During the latter half of the 1960s national S&T policies thus began to emerge in the OECD member states.

For example, the statistics made visible that physics had been very successful after World War II in organizing its interests both within the various nation states and at the level of international collaborations (e.g., CERN). Other scientific communities (e.g., molecular biology) claimed more budgetary room given new developments and increases in overall science budgets. During the period 1965-1975, the preferred instrument for dealing with these issues at the national level was a differentiation in the increase rates of budgets at the disciplinary level. In summary, the focus remained on (financial) input-indicators, while the system relied on peer review for more fine-grained decision-making at lower levels (Mulkay, 1976).

Output indicators

The attention for the measurement of scientific communication originated from an interest other than research evaluation. During the 1950s and 1960s, the scientific community itself had become increasingly aware of the seemingly uncontrolled expansion of scientific information and literature during the postwar period. In addition to its use in information retrieval, the *Science Citation Index* produced by Eugene Garfield's Institute of Scientific Information came soon to be recognized as a means to objectify standards (Price, 1963; Elkana *et al.* 1978). The gradual introduction of output indicators (e.g., numbers of publications and citations) could be legitimated both at the level of society—because it enables policy makers and science administrators to use arguments of economic efficiency—and internally, because quality control across disciplinary frameworks becomes difficult to legitimate unless objectified standards can be made available in addition to the peer review process.

In 1976 Francis Narin's pioneering study *Evaluative Bibliometrics* was published under the auspices (not incidentally) of the U.S. National Science Foundation. Henry Small (1973) had proposed a method for mapping the sciences based on the co-citations of scientific articles. While Small's approach tried to agglomerate specialties into disciplinary structures, Narin focused on hierarchical structures that operate top-down (Carpenter & Narin, 1973; Pinski & Narin, 1976). This program appealed to funding agencies like the N.S.F. and N.I.H. that faced difficult decisions in allocating budgets across disciplinary frameworks.

Was it possible to classify the sciences in terms of journal clusters, both substantively and in terms of rank-orders? Might weights be attributed to publications in terms of standards, for example, such as expected versus observed citation rates (Braun *et al.* 1985; Schubert and Braun, 1993)? A scientometric research program at the macro-level could thus increasingly be formulated. While the American studies of the 1970s had focused on the organization of scientific literature, the application of these indicators to research evaluation on an institutional basis was developed in the European context during the 1980s.

Following the publication of Martin & Irvine's (1983) study of the relative research performance of various (expensive) installations for radio-astronomy, the idea of the assessment of institutional units took hold among policy makers. Leiden University in

the Netherlands pioneered during the early 1980s with a fine-grained model for introducing output in terms of publications and recognition through citations as feedback parameters into the finance scheme of departments during the early 1980s (Moed *et al.*, 1985). This idea was generalized in the UK-model of the Research Assessment Exercises for the funding of university research (since 1992).

The other European countries did not follow the UK in this rationalization of a budget model for research at the national level (Hicks & Katz, 1996), but pressures prevailed during the 1990s to make publication and citation rates visible in evaluation exercises. For example, after the German unification in 1990, extensive evaluation of the research portfolio of Eastern Germany was immediately placed on the science policy agenda (Weingart, 1991). But can the scientometric indicators carry the political burden of research evaluation?

Methodological limitations

Publication and citation analyses have become standard tools for research evaluation. However, some methodological problems remain unresolved. The consequent uncertainties have sometimes been reflected in hesitations to apply these tools as standards in policy making processes and research management decisions. How shaky is the ground on which the research evaluations stand?

First, one can raise the question of the unit of analysis in scientific knowledge production and control (Collins, 1985). The intellectual organization of the sciences does not coincide with their institutional organization. Furthermore, the relations between these layers of organization can be expected to vary among disciplines (Whitley, 1984). The assumption that one can compare 'like with like' in terms of institutional parameters (Martin & Irvine, 1983) is problematic from the perspective of communication studies. New scientific developments (e.g., artificial intelligence) emerge in very different institutional settings, and in order to make a fair comparison one should perhaps first define a *cognitive* unit of analysis. However, the intellectual organization of the sciences cannot easily be observed or measured (Leydesdorff, 1995).

An alternative way to define a unit of analysis would be to base the operationalization on the reflection of scientific developments in the scientific journal literature. The scientific literature is organized in relatively discrete clusters of journals. For example, an article in a biochemistry journal will not often cite an article in condensed matter physics, or vice versa. The relations between these textual units of analysis and the institutional units under evaluation generate further research questions since publication and citation rates differ among disciplines.

The relative decomposability of the literature was central to the above noted attempt of Narin to cluster the database of aggregated journal-journal citations. However, the clustering algorithms provide a snapshot. The structure at any given moment in time does not take into account the dynamic development of the sciences over time. One expects scientific specialties to develop in parallel and not in a hierarchical order. Furthermore, Narin (1976) had proposed to *fix* a journal set *ex ante* in order to make comparisons over time possible. However, advanced industrial nations tend to publish

in newly emerging areas (and accordingly new journals) relatively more than research units in more conservative systems. The so-called ‘decline of British science’ (Irvine *et al.*, 1985)—a subject of intense political debate during the 1980s—could with hindsight be deconstructed as partially an artifact of this type of methodological decisions. Within a dynamic database the U.K. is more stable than in a fixed set, since losses on one side tend to be compensated at the other (Leydesdorff, 1988; Braun *et al.*, 1989; Martin, 1991).

Multivariate and dynamic analysis

Journal literature can be considered as a huge network that is knit together in terms of aggregated citations among articles and co-occurrences of words among texts and titles. Factor analysis enables us to hypothesize a latent structure of this network at each moment in time in terms of groupings. However, factor analysis is often computationally too intensive to address the entire database in a single run, let alone to address these questions both comprehensively and dynamically.

Graph analytical techniques are based on recursive (‘bottom-up’) procedures that operate on relations and can therefore be applied more easily to large datasets. By definition the results of relational analysis, however, can exhibit only relations and hierarchies. The positional analysis of the groupings in a network—that is spanned in terms of relations—is different in kind (Burt, 1982). But, as noted, the positional analysis remained confined to relatively restricted datasets (Leydesdorff, 2005).

A group of researchers at the *École Nationale Supérieure des Mines* in Paris proposed to focus on words and relations among words (‘co-words’) as an alternative to citation and co-citation analysis (Callon *et al.*, 1983). One advantage would be that the words and co-words occur not only in the scientific journal literature, but also in policy reports and patent applications. Can the strengths of the relations among words be used as an indicator of the survival value of an indicated concept during these ‘translations’ across domains? These authors envisaged that the evaluation of research in terms of performance would become possible by using words and their co-occurrences as indicators of ‘translation’ (Callon *et al.*, 1986; Latour, 1987).

The analysis of the co-word patterns proceeded technically in a manner analogous to the co-citation analysis being further developed at the Institute of Scientific Information (ISI) by Small’s group (Small, 1973). In the meantime, the ISI group had produced an *Atlas of Science* that was based on agglomerative clustering techniques using a graph analytical algorithm (Small *et al.*, 1985). These mappings, however, were flawed by the initial decision to focus on hierarchical *relations* for the study of structure and strategic *positions*. Structure can be analyzed only in terms of differentiations into latent dimensions of the system. These dimensions can be revealed using factor analytical techniques (Leydesdorff, 1987 and 1992; cf. Lazarsfeld & Henry, 1968).

While the sciences are discursively constructed as networks of communication in terms of relations among words and sentences, the aggregated constructs can be expected to differentiate over longer periods of time according to rules which are functional to the further advancement of the intellectual organization of specialized,

and therefore relatively autonomous, structures of scientific communication (Luhmann, 1984 and 1990). The various discourses continuously update and rewrite reflexively their understandings of the relevant history. This ‘self-organization’ of the (paradigmatic) discourses takes place from the hindsight perspective of current understanding, both at the level of the active scientists involved and—perhaps with a time lag—at the level of the policy agencies. The latter can direct and influence the development of the sciences only by budgetary and institutional means. The research fronts, however, develop with a communication dynamics different from the institutional organizations.

For example, what one nowadays considers as ‘biotechnology’ or ‘artificial intelligence’ is very different from what both policy makers and the scientists involved considered as relevant to these categories in the 1980s when the priority programs in these areas were first developed (Nederhof, 1988). The double task of reconstructing the history and of rewriting it as it develops generates the need for a ‘double hermeneutics’ in the reconstruction (Giddens, 1976 and 1984). Both the observable variation and the selection mechanisms (latent eigenvectors at the network level) can be expected to change over time (Leydesdorff, 1997).

Whereas the variation is visible in the data, the selection mechanisms remain latent and can therefore only be hypothesized. On the one hand, these constructs are needed as dimensions for the mapping of the data. On the other hand, constructs remain ‘soft,’ that is, open for debate and reconstruction. De Solla Price’s (1978) dream of making scientometric mapping a relatively hard social science can with hindsight be considered as fundamentally flawed (Wouters and Leydesdorff, 1994; cf. Price, 1970). When both the data and the perspectives are potentially changing, the position of the analyst can no longer be considered as neutral.

During the 1990s, scientometric research evaluation would suffer the kind of fragmentation that is well known to the social scientist. Research evaluation became increasingly contingent on the question of the evaluating agency. The crisis became manifest in the journal *Scientometrics* when this leading journal of the field devoted a special issue to the question of ‘Little Scientometrics, Big Scientometrics... and Beyond?’ in 1994 (Glänzel & Schöpfli, 1994). Not only the reflections became more uncertain than before, but also the subjects under reflection had begun to change because of the increasing focus on the techno-sciences and science-based innovations.

The European Union and the field of research evaluation

Whereas research evaluation was shaped as an agenda at the level of national agencies, the *Single Act* of the European Community in 1986 and the *Maastricht Treaty* of the European Union in 1991 have marked a gradual transition within Europe to a supra-national technology and innovation policy. The EU policies continuously referred to science and technology, because these are considered as the strongholds of the common heritage of the member states. However, the ‘subsidiarity’ principle prescribes that the European Commission should not intervene in matters that can be left to the nation states. Therefore, a ‘federal’ research program of the EU could not be developed without taking the detour of a focus on *innovation* as a science-based practice (Narin & Noma, 1985).

The *Research, Technology, and Development (RTD) Networks* of the European Union promote transnational and transsectoral collaboration by rewarding the participation of research groups that capitalize on *complementarities* among national origins and institutional spheres. Thus, the operation of ‘a triple helix of university-industry-government relations’ has been reinforced by the European level of policy making. A system of negotiations and translations among expectations tends to add a dynamic overlay to the nationally institutionalized systems (Etzkowitz and Leydesdorff, 2000).

Interactions can be optimized *ex post* with objectives other than the institutional rationales *ex ante*, and when repeated over time, the network systems increasingly provide their own dynamics. From this perspective, the institutional layer can be considered as the retention mechanism for a network that tends to develop further in terms of its functions. The European Union has provided a feedback on the functions by stabilizing the next-order system’s level. The overlay system can be conceptualized as a network mode—named ‘Mode 2’ by Gibbons *et al.* (1994)—or as ‘international’ when compared with the ‘national systems of innovation’ previously studied by evolutionary economists (e.g., Lundvall, 1992; Nelson, 1993).

The new models can be considered neo-evolutionary insofar as they provide a heuristics for measurement other than the institutional ones that have prevailed at the level of the nation states. For example, national governments have been limitedly successful in developing transdisciplinary programs (Van den Daele *et al.*, 1979; Studer and Chubin, 1982). The bureaucratic focus of Europe, however, legitimated a shift from scientific publications in the traditional format towards achievements and so-called ‘deliverables,’ relatively unhindered by the evaluation schemes of national research councils and scientific communities. These ‘deliverables’ have become a carrier for next rounds of policy formation in a new mode of research evaluation. From this perspective, the scientific literature can be expected to lag behind the research agenda (Lewison and Cunningham, 1991).

Research groups can be sorted by this (trans-national) evaluation scheme in terms of their reliability in providing deliverables to the bureaucracy and therefore in terms of their competencies to serve their audience. In ‘Mode 2’ research not only the social, but also the intellectual organization of projects and programs is increasingly functionalized in terms of serving relevant audiences (Kobayashi, 2000). It should be noted that this shift does not imply necessarily a commercialization of science, since the mechanisms remain mainly institutional and therefore non-market (Nowotny *et al.*, 2001).

Leydesdorff and Etzkowitz (1998) have argued for ‘innovation’ as the analytical unit of operation of ‘innovation systems’ that incorporate knowledge-based developments. However, the delineation of a knowledge-based innovation system itself begs the question. Evolutionary economists have emphasized the national character of innovation systems (Lundvall, 1992; Nelson, 1993; Skolnikoff, 1993), while a focus on technological developments suggests the sectoral level as the most relevant system of reference (Pavitt, 1984; Freeman, 2002). Others (e.g., Carlsson, 2002) have argued in favour of new technosciences like biotechnology as the frameworks of knowledge integration. The various subsystems of innovation crisscross the organization of

society and can be expected to drive (or to inhibit) one another. Thus, the specification of the system of reference for the research evaluation becomes itself increasingly a research question.

The Internet and the development of Cybermetrics and Webometrics

The emergence of the Internet during the 1990s has turned the tables again. Globalization takes place at a supra-institutional level and it reinforces direct relations between science, technology, and the market economy as different mechanisms potentially functional for the coordination. The institutional organizations (e.g., at the national levels) can then be considered as providing niches of communication that develop themselves (or stagnate) by drawing on resources from and by earning credit in the global environments.

Although the carrying organizations provide the original material for their representation at the level of the communication networks, the representations can circulate as ‘actants’ in the networked relations (Callon and Latour, 1981). The actors behind the actants are increasingly black-boxed when the interactions among the representations begin to resonate at the network level. Thus, the represented systems may become dependent on their representations at the network level, and the carrying systems become reflexively aware of their being reflected (Wouters, 1999).

Under these conditions, research evaluation can only position itself reflexively with reference to the representations which are being evaluated because it is analytically unclear (and sometimes not easily accessible) what is precisely reflected and represented (Rip, 1997). Texts are embedded in *contexts*, but the latter provide the former with meaning given specific codifications of the communication. The different audiences can be served by different discourses using different databases or from different perspectives on the same data. The representations are evaluated in terms fulfilling their functions at interfaces.

For example, an innovation like the introduction of a new drug on the market has a meaning for the corporation carrying the market introduction which is different from its meaning for the patients suffering from a disease or from that of the scientists and pharmacists who developed the drug. The latter may use *Chemical Abstracts* as their system of reference, or perhaps the *Medline* for searching and reference. The generic name of the drug used in scientific communications will not be familiar to most of its users, who know the same substance only by its tradename. The evaluation schemes of these different audiences can be expected to vary, for example, between the molecular biologists and the medical scientists involved in clinical testing (Leydesdorff, 2001a).

The agencies carrying these different discourses have only a limited capacity to interface with relevant discourses in their environments (‘contexts’). This structuration of the discourses enables each of them to focus on the quality of its own communications. Overarching ‘Mode 2’ research runs the risk of relating representations to representations without access to the relevant substances because of the formalization. Although the modeling of these complex interactions is a legitimate enterprise, the abstract results of the simulations are very different from the

substantive insights of the research evaluations. While the latter can intuitively be understood and translated into policies, the former require first a theoretical interpretation.

Reflexive scientometrics

Are there still options for evaluating research and other knowledge-based communication given these complex dynamics? When ‘All that is solid, melts into air’ (Marx, 1848), the melting still leaves behind traces of communication. Communication systems communicate with other communication systems across interfaces and if this mutual information is sufficiently repeated the systems may ‘lock-in’ and temporarily stabilize into a co-evolution of mutual shaping at an interface. How is one able to operationalize the specificity and therewith the quality of the interfacing communications?

Communication systems (e.g., reflexive discourses) cannot be considered as givens with clear delineations. However, the (sub-)systems of communication can be reconstructed because they have been historically constructed, and insofar as their trajectory has been stabilized, this unit of analysis can then also be made amenable to the measurement and eventually management (Leydesdorff, 2001b). One is able to observe the events, but they can be provided with meaning only in relations to discourses. However, one can hypothesize these systems of communication on the basis of theoretical information about the specificity of their communication. By raising first the question of ‘What is communicated when the system under study communicates?’ an analyst is able to reconstruct a code of the communication system under study. Only after addressing this substantive question can one meaningfully proceed to the question of ‘how is this codification expected to be communicated?’ and then to indicators for the measurement given the specification of the system of reference.

Communications are amenable to measurement because a specific (codified) substance is communicated, and therefore redistributed. A change in the distribution generates a probabilistic entropy (that can be expressed in terms of bits of information). Whereas this mathematical definition of ‘information’ is still content-free, the specification of a system of reference provides the uncertainty with meaning. As noted, the specification of a system of reference itself is a difficult and analytical task, since one needs information both about the relations of the communications under study with the relevant systems in its environment (at each moment in time) and about the internal development of meaning within the system(s) under study.

Historical research becomes relevant to scientometric analysis from this perspective. The analytical clarification in the reconstructing discourse reproduces the discourses under study as best as it can. Still, the representations may serve purposes other than those of the systems represented. The quality of the representation—both in terms of the analytical specificity and the quantitative precision—becomes crucial to the costs involved in making the management of these representations socially acceptable to the systems represented and therewith to their further development.

Where does this leave us in a seemingly exploding universe of communications and representations? From the perspective of research evaluation, the development of scientometric indicators can only improve the quality of the communication reflexively after asking what a given discourse has been communicating. Qualitative discourses can be expected to be functionally specific. How well has a discourse under study served the internal missions of the field and/or the missions of the agencies who committed the research?

It seems to me that different objectives can be distinguished. First, there is the need at the system's level to provide high-quality information when making decisions about S&T in the public sphere or R&D at institutional levels. This information can only be considered as partial indicators that are locally constructed with reference to specific research questions. However, it is well-known that windows of appreciation and evaluation do terrible things to the tangents of the systems under study (Casti, 1989). One should proceed very carefully in this direction (Van Raan, 1988).

Second, scientometrics has made us aware that science is amenable to measurement, however imperfect the representation may be. The history of science, the sociology of science, and the philosophy of science can be recognized as qualitative representations of the sciences under study. The relations between qualitative and quantitative approaches can be reformulated: qualitative descriptions and insights inform the measurements as hypotheses and heuristics, while the measurements can be updated and refined by taking interaction terms into account.

Qualitative approaches inform the model from different perspectives, and the results of the quantitative analysis can again be provided with meaning from the various perspectives. The model can then be considered as a machine that enables us further to develop our theories of science. However, as against the modernist concept, the reflexive model can be expected also to fail to carry out this function. All representations remain necessarily incomplete in comparison with the represented system. The sum of the partial representations is not necessarily more informative than their *differences*. Because of this focus on the potential differences in status of the various contributions, the research agenda of quantitative science and technology studies or scientometrics cannot avoid to take a methodological turn: the measurement can no longer avoid the question of what is being measured and why.

Conclusion

Whereas the research program of the measurement of scientific communications emerged in a context where the delineations between academia, government, and industry were institutionalized, the systemic development of these relations during the second half of the 20th century has changed the system of reference for the evaluation of research. In a knowledge-based economy science fulfills functions that change the definitions of what is considered research and the globalization has changed the relevance of a national system of reference. In Europe notably the transnational level has taken the lead in developing innovation policies in an attempt to address the internationalization of industry and the globalization of innovations. Science, of course, has been internationally oriented from its very beginnings, but the entrainment

of the research process in these global developments is reflected in the research evaluation and the scientometric measurement.

In other words, the systems under study have become more complex. A complex dynamics can analytically be decomposed in several subdynamics. For example, one can raise the question of whether international collaboration in science and coauthorship across national boundaries has emerged during the 1990s as a new subdynamics with characteristics other than domestic collaboration and coauthorship (Persson *et al.*, 2003; Wagner & Leydesdorff, 2003). Can a different (e.g., global) subdynamic be hypothesized and then also be measured? Can this new dimension of scientific output also be accounted for in schemes that were developed for institutional management within national systems? The questions generate puzzles at the interfaces between the sciences and the economic and political contexts. The evolving systems and subsystems communicate in different dimensions and the evaluation has become part of the codification of these communications.

References

- Abramowitz, M. 1956. Resource and Output Trends in the United States since 1870, *American Economic Review*, 46: 5-23.
- Braun, T. (ed.). 1998. Topical Discussion Issue on Theories of Citation, *Scientometrics* 43: 3-148.
- Braun, T., W. Glänzel, and A. Schubert . 1985. *Scientometric Indicators. A 32-Country Comparative Evaluation of Publishing Performance and Citation Impact*. Singapore/Philadelphia: World Scientific Publications.
- Braun, T., W. Glänzel, and A. Schubert. 1989. Assessing Assessments of British Science. Some Facts and Figures to Accept or Decline, *Scientometrics*, 15: 165-70.
- Burt, R. S. 1982. *Toward a Structural Theory of Action*. New York, etc.: Academic Press.
- Bush, V. 1945. *The Endless Frontier: A Report to the President*. Reprinted New York: Arno Press, 1980.
- Callon, M., and B. Latour. 1981. Unscrewing the big Leviathan: how actors macro-structure reality and how sociologists help them to do so. In *Advances in Social Theory and Methodology. Toward an Integration of Micro- and Macro-Sociologies*, edited by K. D. Knorr-Cetina and A. V. Cicourel. London: Routledge & Kegan Paul, 277-303.
- Callon, M., J.-P. Courtial, W. A. Turner, and S. Bauin. 1983. From Translation to Problematic Networks: An Introduction to Co-Word Analysis, *Social Science Information*, 22: 191-235.
- Callon, M., J. Law, and A. Rip. 1986. *Mapping the Dynamics of Science and Technology*. London: Macmillan.
- Carpenter, M. P., and F. Narin. 1973. Clustering of Scientific Journals, *Journal of the American Society for Information Science*, 24(6): 425-436.
- Carlsson, B. (Ed.). 2002. *New Technological Systems in the Bio Industries -- an International Study*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Casti, J. 1989. *Alternate Realities*. New York, etc.: Wiley.
- Collins, H. M. 1985. The Possibilities of Science Policy, *Social Studies of Science* 15: 554-558.

- Edwards, P. 1996. *The Closed World, Computers and the Politics of Discourses in Cold War America*, Cambridge, MA: MIT Press.
- Elkana, Y., J. Lederberg, R. K. Merton, A. Thackray, and H. Zuckerman. 1978. *Toward a Metric of Science: The advent of science indicators*. New York, etc.: Wiley.
- Etzkowitz, H., and L. Leydesdorff . 2000. The Dynamics of Innovation: From National Systems and 'Mode 2' to a Triple Helix of University-Industry-Government Relations, *Research Policy*, 29(2): 109-123.
- Freeman, C. 2002. Continental, national and sub-national innovation systems—complementarity and economic growth, *Research Policy* 31: 191-211.
- Garfield, E. 1979. *Citation Indexing*. Philadelphia, PA: ISI.
- Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, and M. Trow. 1994. *The new production of knowledge: the dynamics of science and research in contemporary societies*. London: Sage.
- Giddens, A. 1976. *New Rules of Sociological Method*. London: Hutchinson.
- Giddens, A. 1984. *The Constitution of Society*. Cambridge: Polity Press.
- Glänzel, W., and U. Schöpfli. 1994. Little Scientometrics, Big Scientometrics.... and Beyond, *Scientometrics*, 30(2-3): 375-384.
- Hicks, D., & J. S. Katz. 1996. Science Policy for a Highly Collaborative Science System. *Science and Public Policy*, 23, 39-44.
- Irvine, J., B. Martin, T. Peacock, R. Turner. 1985. Charting the decline of British science, *Nature* 316: 587-90.
- Kobayashi, S.-i. 2000. Applying Audition Systems from the Performing Arts to R&D Funding Mechanisms: Quality Control in Collaboration among the Academic, Public, and Private Sectors in Japan. *Research Policy*, 29(2), 181-192.
- Latour, B.. 1987. *Science in Action*. Milton Keynes: Open University Press.
- Lazarsfeld, P. F., and N. W. Henry. 1968. *Latent structure analysis*. New York: Houghton Mifflin.
- Lewis, G., and P. Cunningham. 1991. Bibliometric Studies for the Evaluation of Trans-National Research, *Scientometrics*, 21: 223-244.
- Leydesdorff, L. 1987. Various methods for the Mapping of Science, *Scientometrics*, 11: 291-320.
- Leydesdorff, L. 1988. Problems with the measurement of national scientific performance, *Science and Public Policy*, 15: 149-52.
- Leydesdorff, L. 1992. A Validation Study of 'LEXIMAPPE', *Scientometrics*, 25: 295-312.
- Leydesdorff, L. 1995. *The Challenge of Scientometrics: The development, measurement, and self-organization of scientific communications*. Leiden: DSWO Press, Leiden University; at <http://www.upublish.com/books/leydesdorff-sci.htm> .
- Leydesdorff, L. 1997. Why Words and Co-Words Cannot Map the Development of the Sciences, *Journal of the American Society for Information Science* 48(5): 418-27.
- Leydesdorff, L. 2001a. Indicators of Innovation in a Knowledge-based Economy, *Cybermetrics*, 5, Issue 1, Paper 2, at <http://www.cindoc.csic.es/cybermetrics/articles/v5i1p2.html>
- Leydesdorff, L. 2001b. *A Sociological Theory of Communication: The Self-Organization of the Knowledge-Based Society*. Parkland, FL: Universal Publishers; at <http://www.upublish.com/books/leydesdorff.htm> .

- Leydesdorff, L. 2005. Can Scientific Journals be Classified in terms of Aggregated Journal-Journal Citation Relations using the *Journal Citation Reports*? *Journal of the American Society for Information Science and Technology* (forthcoming).
- Leydesdorff, L., and P. van den Besselaar. eds.. 1994. *Evolutionary Economics and Chaos Theory: New directions in technology studies*. London: Pinter.
- Leydesdorff, L., and P. van den Besselaar. 1997. Scientometrics and Communication Theory: Towards Theoretically Informed Indicators, *Scientometrics* 38: 155-74.
- Leydesdorff, L., and H. Etzkowitz. 1998. The Triple Helix as a model for innovation studies, *Science and Public Policy*, 25(3): 195-203.
- Leydesdorff, L., and P. Wouters. 1999. Between Texts and Contexts: Advances in Theories of Citation? *Scientometrics*, 44: 169-182.
- Luhmann, N. 1984. *Soziale Systeme. Grundriß Einer Allgemeinen Theorie*. Frankfurt a. M.: Suhrkamp.
- Luhmann, N. 1990. *Die Wissenschaft Der Gesellschaft*. Frankfurt a.M.: Suhrkamp.
- Lundvall, B.-Å. (ed.). 1992. *National Systems of Innovation*. London: Pinter.
- Martin, B. R.. 1991. The Bibliometric Assessment of UK Scientific Performance. A Reply to Braun, Glänzel and Schubert, *Scientometrics* 20: 333-357.
- Martin, B., and J. Irvine. 1983. Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy, *Research Policy* 12: 61-90.
- Marx, K. 1848. *The Communist Manifesto*. Paris.. Translated by Samuel Moore in 1888. Harmondsworth: Penguin, 1967.
- Moed, H. F., W. J. M. Burger, J. G. Frankfort, and A. F. J. Van Raan. 1985. The Use of Bibliometric Data for the Measurement of University Research Performance, *Research Policy*, 14: 131-49.
- Mulkay, M. J. 1976. The mediating role of the scientific elite, *Social Studies of Science*, 6: 445-470.
- Narin, F. 1976. *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*, Washington, DC: National Science Foundation.
- Narin, F., and E. Noma. 1985. Is Technology Becoming Science? *Scientometrics*, 7: 369-381.
- Nederhof, A. J.. 1988. Changes in publication patterns of biotechnologists: An evaluation of the impact of government stimulation programs in six industrial nations, *Scientometrics*, 14: 475-485.
- Nelson, Richard R. (ed.). 1993. *National Innovation Systems: A comparative study*. Oxford and New York: Oxford University Press.
- NSF (2000). *Science and Technology Then and Now*. NSF Fact Sheet at <http://www.nsf.gov/od/lpa/news/media/2000/bckgrdrst.htm> (last visited Sep. 8, 2003).
- Nowotny, H., P. Scott, and M. Gibbons. 2001. *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge: Polity Press.
- OECD. 1963, ³1976. *The Measurement of Scientific and Technical Activities: 'Frascati Manual'*. Paris: OECD.
- OECD. 1964. *The Residual Factor and Economic Growth*. Paris: OECD.
- Pavitt, K. 1984. Sectoral patterns of technical change: towards a theory and a taxonomy, *Research Policy*, 13: 343-73.
- Persson, O., W. Glänzel, & R. Danell. (2003). *Inflationary Bibliometric Values: The Role of Scientific Collaboration and the Need for Relative Indicators in Evaluative Studies*. Paper presented at the 9th International Conference on

- Scientometrics and Informetrics, Beijing.
- Pinski, G., and F. Narin. 1976. Citation Influence for Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics, *Information Processing and Management*, 12(5): 297-312.
- Price, D. de Solla. 1963. *Little Science, Big Science*. New York: Columbia University Press.
- Price, D. de Solla. 1970. Citation measures of hard science, soft science, technology and non-science. In: *Communication among Scientists and Engineers*, edited by C. E. Nelson & D. K. Pollock. Lexington, MA: D. C. Heath & Co, pp. 3-22.
- Price, D. de Solla. 1978. Editorial statement, *Scientometrics*, 1: 7-8.
- Rip, A. 1997. Qualitative Conditions for Scientometrics: The New Challenges, *Scientometrics*, 38(1): 7-26.
- Rosenberg, N. 1976. *Perspectives on Technology*. Cambridge: Cambridge University Press.
- Rousseau, R. 1997. Sitations: An exploratory study, *Cybermetrics* 1:1, at <http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html>
- Schubert, A., and T. Braun. 1993. Standards for Citation Based Assessments, *Scientometrics* 26: 21-35.
- Skolnikoff, E. B. 1993. *The Elusive Transformation*. Princeton, NJ: Princeton University Press.
- Small, H. 1973. Co-citation in the scientific literature: a new measure of the relationship between two documents, *Journal of the American Society for Information Science*, 24: 265-269.
- Small, H., E. Sweeney, and E. Greenlee. 1985. Clustering the Science Citation Index Using Co-Citations II. Mapping Science, *Scientometrics*, 8: 321-340.
- Studer, K. E., and D. E. Chubin. 1982. *The Cancer Mission. Social Contexts of Biomedical Research*. Beverly Hills, etc.: Sage.
- Van den Daele, W., W. Krohn, and P. Weingart (eds.). 1979. *Geplante Forschung*. Frankfurt a.M.: Suhrkamp.
- Wagner, C. S., & L. Leydesdorff. (2003). *Mapping Global Science Using International Co-Authorships: A Comparison of 1990 and 2000*. Paper presented at the 9th International Conference on Scientometrics and Informetrics, Beijing.
- Weingart, P. (ed.). 1991. *Die Wissenschaft in osteuropäischen Ländern im internationalen Vergleich—eine quantitative Analyse auf der Grundlage wissenschaftsmetrischer Indikatoren*. Bielefeld: Kleine Verlag.
- Whitley, R. R. 1984. *The Intellectual and Social Organization of the Sciences*. Oxford: Oxford University Press.
- Wouters, P. 1999. *The Citation Culture*. Unpublished Ph.D. Thesis, University of Amsterdam.
- Wouters, P., and L. Leydesdorff. 1994. Has Prices Dream Come True: Is Scientometrics a Hard Science? *Scientometrics*, 31: 193-222.
- York, H. F.. 1970. *Race to Oblivion: A Participants View of the Arms Race*, New York: Simon and Schuster; at <http://www.learnworld.com/ZNW/LWText.York.Race.Access.html>