Can Networks of Journal-Journal Citations Be Used as Indicators of Change in the Social Sciences?

Loet Leydesdorff

Science & Technology Dynamics,

Amsterdam School of Communications Research,

Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands

loet@leydesdorff.net; http://www.leydesdorff.net/

Abstract

Aggregated journal-journal citations for 1999 are compared with similar data in the *Journal Citation Reports* 1998 of the *Social Science Citation Index.* In addition to indicating local change, probabilistic entropy measures enable us to analyze changes in distributions at different levels of aggregation. The indicators are compared by elaborating the journal-journal mappings and in relation to similar developments in the *Science Citation Index*. Specialty formation seems a more important mechanism in the development of the social sciences than in the natural and life sciences, but the developments are volatile. The use of aggregate statistics based on the *Science Citation Index* is ill-advised in the case of the social sciences because of differences in the underlying dynamics.

1. Introduction

With hindsight the Scientific Revolution of the 17th century can be recognized as first and foremost a communication revolution. The printing press made it possible to change the dynamics of information storage, archiving, and retrieval (Eisenstein, 1983; Kaufer & Carley, 1993). New editions of scientific texts were no longer transcripts of older manuscripts, but they could thenceforce be expected to contain updated information and new knowledge. This update mechanism was provided, first, by circles of correspondence, but then rapidly scientific journal literature emerged. Price (1961) provides a graphical illustration of the exponential growth of the journal literature since the appearance of the first journals in 1665 (Figure 1).[1]

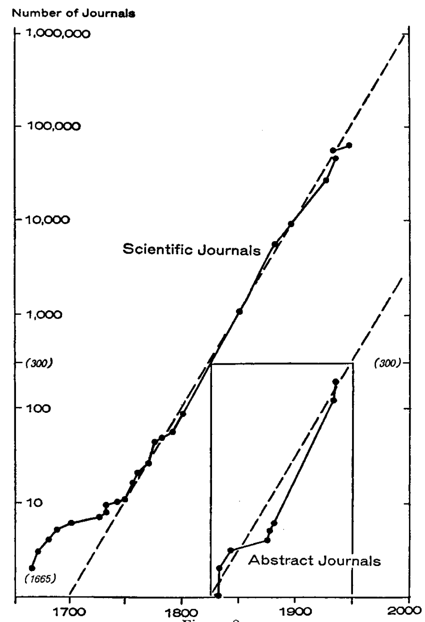


Figure 8.1
Number of journals founded (*not* surviving) as a function of date. The two uppermost points are taken from a slightly differently based list.

The social sciences emerged much later, but particularly since World War II these sciences have adopted formats of communication similar to those of the natural and the life sciences. The international journal literature in the social sciences is still less codified than that in the natural sciences, but citation practices have been established in these younger sciences as well (Price, 1970). The Institute of Scientific Information (ISI) entertains a *Social Science Citation Index* that can be compared to the *Science Citation Index* in many respects.[2]

---

[1] The first scientific journal published were the *Philosophical Transactions of the Royal Society* in 1665, and the French *Journal des Sçavants* shortly thereafter (cf. Leydesdorff, 1998).

[2] The *Science Citation Index* exists since 1961, but the first edition of the *Journal Citation Reports* is from 1975. The *Social Science Citation Index* was first published in 1966, and extended with *Journal Citation Reports* in 1978.

Despite the lower degree of codification in the social sciences and the humanities, the citation network structures are rather similar in a number of respects.

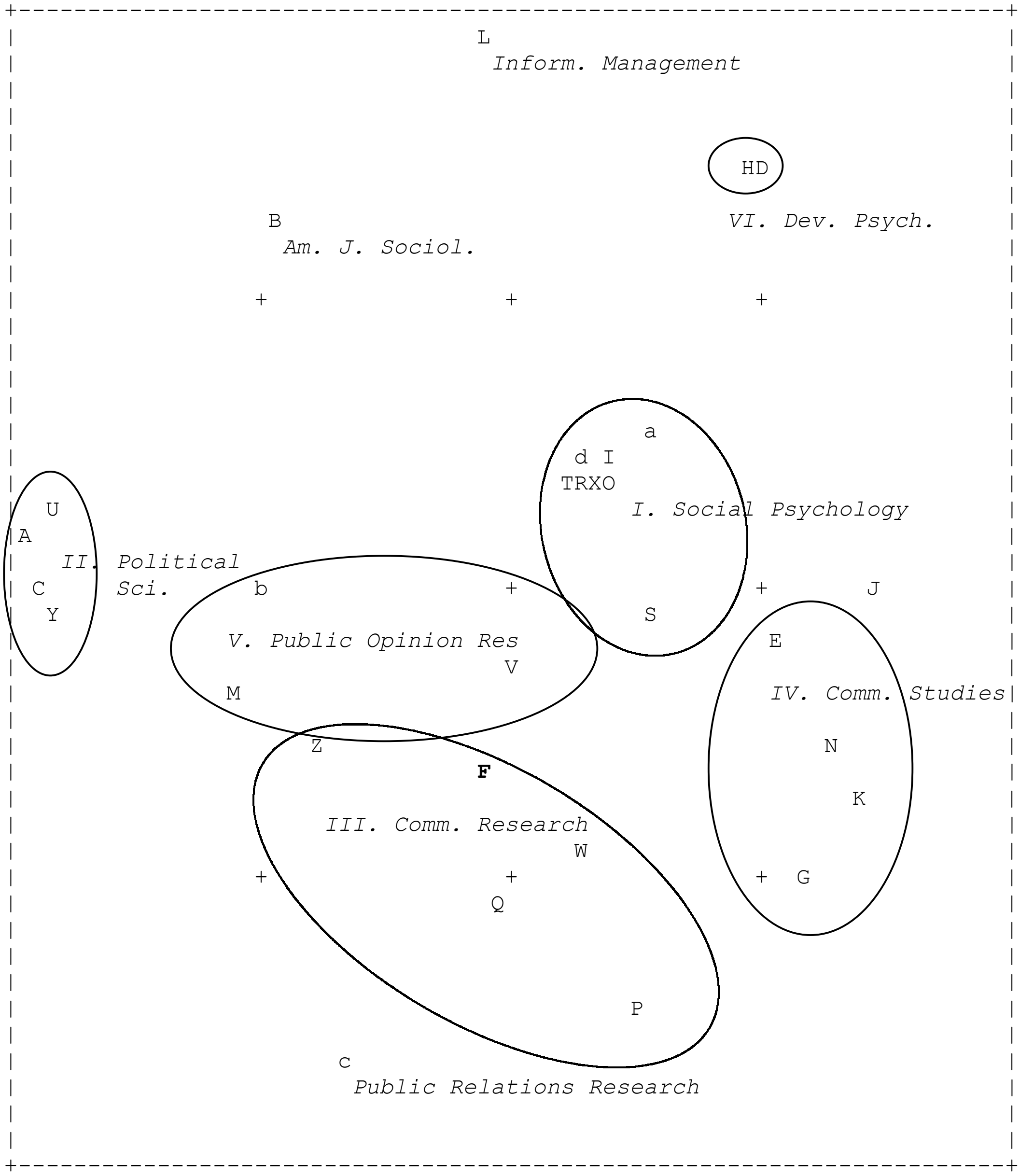
For example, aggregated citations among journals in the natural sciences provide us with a robust network which is reproduced year after year with relatively minor modifications (Carpenter & Narin, 1973; Doreian & Fararo, 1985; Leydesdorff, 1986; Tijssen *et al.*, 1987; Tijssen, 1992). This network can be analyzed using multi-variate (e.g., factor) analysis and then "the structure of science" in terms of citation patterns can be mapped accordingly (e.g., by using multidimensional scaling techniques).

Citations occur in dense clusters indicating specialty structures which operate in parallel. The aggregated citation matrix among journals is nearly decomposable since articles in biochemistry journals, for example, do hardly cite journals in inorganic chemistry, and vice versa. Some journals like *Science* and *Nature* relate journals at a next level, but entertain only relatively weak citation relations with some major journals within the specialty structures (Cozzens & Leydesdorff, 1993; Leydesdorff & Cozzens, 1993).

In (bio-)medicine, this pronounced factor structure of the network is somewhat mitigated by journals relating diseases from a clinical perspective. This relational pattern adds to the recursive production of knowledge within the specialties. However, the subdynamics of within-specialty communication remains more pronounced. In the social sciences and the humanities, an integrating super-structure of scientific communication is provided by scholarly monographs, on the one hand, and by national scientific literatures that are more deeply embedded in the respective (national) cultures, on the other (Nederhof *et al.,* 1989). For example, the French research organization CNRS subsidizes the publication of approximately 225 scientific journals of which 190 in the social sciences and the humanities (De Looze *et al*., 1996; Legentil, *personal communication*).

Monographs are not included in the *Social Science Citation Index* and nationally specific journals are underrepresented. Thus, this database—like in the natural science, but for different reasons—reflects the further differentiation of the sciences into specialties at the global level more than their integration in local instantiations. For example, figure 2 exhibits a mapping of the citation patterns when *Communication Research* is used as a seed journal in 1999. The clustering is based on the factor analysis of the citation patterns on the citing side of the journals which cite or are cited by *Communication Research* to one percent of its total citations. The multi-dimensional scaling provides us with a picture of the fine structure of the neighbouring specialties.

PLOT OF STIMULUS SPACE



X = DIMENSION 1   Y = DIMENSION 2

A. AM J POLIT SCI
B. AM J SOCIOL
C. AM POLIT SCI REV
D. CHILD DEV
E. COMMUN MONOGR
F. COMMUN RES
G. COMMUN THEOR
H. DEV PSYCHOL
I. EUR J SOC PSYCHOL
J. HEALTH COMMUN
K. HUM COMMUN RES
L. INFORM MANAGE
M. INT J PUBLIC OPIN R
N. J APPL COMMUN RES
O. J APPL SOC PSYCHOL
P. J BROADCAST ELECTRON
Q. J COMMUN
R. J EXP SOC PSYCHOL
S. J HEALTH COMMUN
T. J PERS SOC PSYCHOL
U. J POLIT
V. J SEX RES
W. JOURNALISM MASS COMM
X. PERS SOC PSYCHOL B
Y. POLIT BEHAV
Z. POLIT COMMUN
a. PSYCHOL BULL
b. PUBLIC OPIN QUART
c. PUBLIC RELAT REV
d. SOC COGNITION

Figure 2
Multi-dimensional scaling of the citation environment of *Communication Research* in 1999, citing pattern, 1% citation threshold. Clusters are based on the factor analysis of the citation matrix. (See Leydesdorff & Cozzens (1993) for the precise methodology.)

2. The research question

The *Journal Citation Reports* provide a yearly update of these citation networks. In library and information sciences the tables are mainly used for a hierarchical ordering of the journals in terms of their impact factors, etc. Hierarchical ordering is based on transitivity in the relations, while journal mapping is based on structural positions at the level of the latent network (Burt, 1982). These positions, however, are constructs that can be revealed using analytical tools such as factor analysis. From a science policy or a library and information management perspective, one is mainly interested in *change* in this latent structure (ESRC, 1995; Healey, 1997). Can the data be analyzed so that change can be monitored at an early stage?

Structure is an attribute to the network of links, while relations can be attributed to the nodes (that is, the journals). Eigenvector analysis enables us to study the structure in the network of communications among the nodes. As different from graph analysis, the zeros are input to this analysis and the consequential possibility of spatial mapping. A problem, however, is the delineation of data sets that can be processed in a single run. Thus, one has to find a method to cut the database in relevant subdomains.

The factorial decomposition can only be successful if the matrix of journal-journal citations is nearly decomposable (Simon, 1969), for example, in terms of clusters indicating specialties. As noted, this is a reasonable assumption given the prevailing emptiness of the matrix and the ongoing processes of specialization in the sciences. Because of this continuous selection pressure, one can expect the specialization to be reflected in the aggregated journal-journal citation data.

Using a focus on technological breakthroughs in the natural sciences—such as the discovery of superconductivity at relatively high temperatures in 1987 (Leydesdorff *et al.*, 1994)—or the emergence of new fields of science—such as biotechnology, new materials, and artificial intelligence during the 1980s and early 1990s (see Leydesdorff & Gauthier, 1996; Van den Besselaar & Leydesdorff, 1996)—we were able to delineate datasets which could be expected to exhibit structural change. From this previous research we concluded that new developments can be traced in the database mainly in terms of the *being cited* patterns of journals.

New developments attract attention by scholars in neighbouring fields and therefore journals reporting on these new developments are *cited* to a significantly larger extent than the year before. Otherwise stability prevails in "being cited" patterns because this reflection maps the archive of science, whereas citations on the citing side (by intentful authors) provide the variation. Active citation can be considered as the running operator potentially generating change. New combinations are then selected by the codified structures of the specialist communication that prevail.

Since 1994, the ISI produces the *JCR*-data on CD-ROM and thus the comprehensive data is readily available in a computer readable format. In a previous study, I returned to my initial research question of whether one can use this data for indicating change in the natural and life sciences at the aggregate level (Leydesdorff, forthcoming). In this study I extend the approach to the *Social Science Citation Index*. How are processes of change different between the social and the natural sciences? Can indicators derived with reference to the natural sciences legitimately be applied to the social sciences?

3. Static and dynamic aggregation and disaggregation

The sciences do not develop in terms of individual journals, but in terms of journal clusters. As noted, the data does not provide us with direct indicators of change at the level of journal *clusters* and the factor analysis of large data matrices is computationally problematic. Furthermore, factor analysis is essentially a static technique. Substraction of the results for different years, that is, comparative static analysis, does not yet necessarily reveal the dynamics. Using the observable differences, one is no longer able to control whether change is a consequence of structural change or of variation in the data that may have lead to another optimization of the fit in the instances which are compared.

One needs a calculus for studying change in the data in relation to change in structure (Bar-Hillel, 1955). Information calculus or entropy statistics provides us with methods that allow for the study of multi-variate complexity and development over time in a single design (Theil, 1972; Leydesdorff, 1995). As different from the substraction of geometrical representations for various years, the algorithmic account enables us to assess change systematically. Furthermore, information theory is based on the additive quantity of information in bits (Σ), and therefore the decomposition remains fully transparent, and there are no technical limitations of sample size other than disk space.

However, the equivalent of a positional analysis is again too computational intensive on large dataset (Krippendorff, 1986; Leydesdorff, 1995) and therefore, I shall use information theory in this analysis only as a method for the indication of journals that can be considered as indicators of change. Thereafter, I follow up with factor analysis and multidimensional scaling in relevant environments of the so indicated journals.

Following Shannon (1948), Theil (1972) defined the expected information content *I* of a message that an *a priori* distribution Σ p_i has turned into an *a posteriori* distribution Σ q_i , as follows:

$$I = \Sigma_i q_i {}^2\log (q_i / p_i)$$

When the two-base of the logarithm is used, *I* is expressed in bits of information. Furthermore, it can be shown that *I* is necessarily equal or larger than zero (Theil, 1972, at pp. 59 f.). This constraint of a non-negative *aggregated* value for *I* allows for local entropy-changes as negative contributions (ΔI). One expects local structures to contribute to the redundancy by selecting upon the variation.

The expected information is contained in a message that is received by a system *a posteriori*. In other words, the evolutionary analysis changes the time horizon to the operation of the system in the present as building *a posteriori* upon its historical manifestations. While citation analysis has often been used in the social studies of science for the historical reconstruction (e.g., Garfield, 1979; Callon *et al*., 1986), we are here more interested in the relevance of the past for the present.

For example, what was considered 'biotechnology' in 1980 is no longer necessarily defined the same way in later years (Nederhof, 1988). For the prospective policy analysis, however, the current understanding is more relevant than a previous understanding. In other words, the historical axis is inverted when using an evolutionary perspective: the system of reference is *ex post*, whereas the historical analysis tends to fix the framework *ex ante* (Narin, 1976). Data becoming available in each year provide a potential update value for the historically evolving expectations.

I shall focus below on change between the two latest available years at the time of this research, that is, change contained in the data for 1999, given 1998 data. Comparisons with earlier years remain possible, in principle, and often desirable for substantive reasons—that is, for a historical understanding—but these extensions do not add fundamentally to the methodology.

4. Materials

The *Journal Citation Reports* of the *Social Science Citation Index* list the aggregated citation data of 1,690 journals in 1999 versus 1,679 in 1998. This data was reorganized in order to fit legacy software developed for the analysis of this type of data in the 1980s (Leydesdorff & Cozzens, 1993). In general, citation data can be analyzed from the 'cited' and from the 'citing' side. The cell values of the grand matrix can be considered as the mutual information between these two dimensions of the matrix.

The *Citation Indices* of the ISI are generated by processing the publications from the 'citing' side. Literature from the current year is scanned for references to literature in the archives. Then, the matrix is transposed in order to consider also the 'cited' dimension (Wouters, 1999). This operation in itself adds no data to the database. 'Cited', however, are also a number of journals other than those processed by the ISI.

In 1999, 72,468 source items were cited by the citing documents in a total of 2,361,720 citations. The total number of cited items within the domain of the ISI journals was only 1,055,369, that is, 44.7%. The other references are to sources such as monographs. In the *Science Citation* Index, the corresponding percentage was 79.3% indicating the more focused and journal-centered citation behaviour in the natural and life sciences. Thus, the initial selection of journals by the ISI from the total set in the case of the sciences can be more precise than in the case of the social sciences because of these differences in citation behaviour.

I limit the analysis here below to the journals which were processed by the ISI both on the citing and the cited side. This reduces the number of cited references in the distribution considerably, but not in proportion to the above figures. Of the 248,093 unique references contained in the 1999 database, 171,286 point to source materials which were *not* processed by the ISI from the citing side. I work with only the remaining 76,807 citation relations (31.0%) which contain a total of 733,099 citations (that is, 69.5% of the total cited). The other 30.5% are single citation relations which are subsumed by ISI under the category "All others". (This number is only 10.3% for the *Science Citation Index.*)

	1999	*1998*	*1999/1998*
number of source journals processed	1699	1679	1.012
number of items referenced	72,468	72,399	1.001
number of citation relations	248,093	244,339	1.015
total citations 'citing'	2,361,720	2,314,111	1.021
citation relations to source journal	76,807	73,951	1.039
source journal not processed 'citing'	4	2	2
total 'cited'	1,055,369	988,870	1.067
total covered by our analysis[3]	733,099	705,741	1.039

Table 1

Comparison of the data in various relevant dimensions for 1999 and 1998, respectively. (The source journals which were not processed on the citing side will be included into the analysis when the focus is on the 'cited' dimension.)

[3] Single citation relations are compiled by the ISI under the heading "all others" and not included in our analysis.

Table 1 summarizes the data for 1998 and 1999. Additionally, the ISI listed 32 changes of journal names in 1999. Taking this list into account, I was able to match 1660 journals of the 1679 journals listed in 1998 (98.9%). Thirty nine journals (2.3%) were added in 1999.

	SoSCI JCR 1999	*SCI JCR 1999*
number of source journals processed	1699	5550
number of items referenced	72,468 *(42.7)*	194,786 *(35.4)*
number of citation relations	248,093 *(146.0)*	1,371,216 *(249,3)*
total citations 'citing'	2,361,720 *(1390.1)*	20,050,851 *(3645.6)*
citation relations to source journal	76,807 *(45.1)*	771,045 *(140.2)*
source journal not processed 'citing'	4 *(0.0)*	21 *(0.0)*
total 'cited'	1,055,369 *(621.2)*	15,898,944 *(2890.7)*
total covered by our analysis[3]	733,099 *(431.5)*	14,264,510 *(2593.5)*

Table 2

Citation totals compared between the Social SCI JCR and the SCI JCR for 1999.

In italics and between brackets the number per source journal, respectively.

Table 2 provides summary statistics for the comparison between the *Social Science Citation Index* and the *Science Citation Index.* The figures between brackets and in italics are normalized in relation to the number of journals processed in both databases. Thus, one can observe that the number of journals referenced is of the same order of magnitude in the two databases, but that the number of citations to these journals is much larger in the *Science Citation Index*. The network is more densily 'populated' with citations in the natural sciences compared to the social sciences. The citation networks in the latter are relatively thin. As shown in Figure 2 above, however, one can still use this data for indicating network structures. Notably, the factor structure was very pronounced indicating the presence of delineated clusters in the database.

5. Methods

If one conceptualizes the aggregated journal-journal citations as a huge matrix of 1699 journals cited versus these same 1699 journals citing, this matrix contains 1699^2 = 2,886,601 cells. Whereas we have 76,708 unique citation relations (in 1999), only 2.7% of these cells contain a non-missing value. Almost all (97.3 %) of the cells are empty. This emptiness means that the multi-dimensional space corresponding to the matrix representation can be considered as virtually empty. Since citation patterns are highly similar within specialties, one may also consider the system to be nearly decomposable (Simon, 1969).

One implication of the prevailing emptiness of the matrices is that the 1999 matrix is globally very similar to that of 1998 because both sets are mainly empty. Furthermore, we can only make comparisons among journal representations which are present in both years. In other words, the overall pattern can be expected to be rather similar when analyzed at the aggregated level. One expects specific change or, in other words, *change can be considered as the exception*. Can these events also be used as an indicator of newness, obsolescence, etc.?

As noted, the change of a distribution can be measured in bits of information using I as defined above. I is a non-parametric and aggregative measure. The measurement is normalized in terms of the *a posteriori* event, that is, the information is evaluated from a hindsight perspective. The multivariate extension of the dynamic entropy measure to $I_{ijk..} = \Sigma\, q_{ijk..}\, {}^2\log\,(q_{ijk..} / p_{ijk..})$ is straightforward.

Once the information is brought under the control of a database manager, several options for developing indicators using entropy statistics can be distinguished. I shall first compute the contribution of each journal to the overall change of the aggregated journal-journal citations in both the cited and the citing dimension. Thus, we will be able to specify the change in the distribution of total citations in either dimension between the two years, and this change can then be decomposed in terms of the contributions of individual journals to it (ΔI). However, the overall change in the distribution of citation patterns among journals does not yet inform us about the change of the citation patterns of each individual journal as a (one-lower-level) vector of this matrix.

In other words, this first measure provides us with statistics which are normalized in terms of the database. The indicator can be compared with the impact factor, but it is a measure of the dynamics whereas the impact factor is measured for each year separately. However, journals can be compared directly in terms of this indicator, since the values are normalized with reference to the grand total of the citation matrix. Thus, a contribution to the change in the distribution can be expressed as a ΔI for each individual journal in terms of bits of information.

If we wish to use the journal citation pattern as an indicator of specific change, we need to know precisely which journals are cited differently from the year before by each journal separately. In this case, the analysis should be performed at the level of the 76,807 cell values within the matrix in comparison to the 73,951 values available in 1998, and not in terms of the margin totals.

One can then distinguish between the probabilistic entropy generated at the level of each vector and the probabilistic entropy generated at the level of the matrix by specific citation interactions aggregated for each journal. The normalization is different in either case. When normalized with reference to the matrix, contributions to change at each cell can again be compared and aggregated. When comparing vectors, however, the values for different journals cannot be aggregated, since the size of the journal affects the analysis. We will pursue both types of analysis here below, yet with a focus on the 'cited' side for reasons specified above.

In order to keep the computation directly tractable with reference to different possible normalization, let me rewrite the formula for *I* in the following way:

$$I \quad = \quad \Sigma\, q_i \, {}^2\log\,(q_i / p_i)$$

By writing $\Sigma\, q_i$ and $\Sigma\, p_i$ as relative frequency distributions:

$$\Sigma\, q_i = \Sigma\, f_q/n_q \quad \text{and} \quad \Sigma\, p_i = \Sigma\, f_p/n_p \text{ , respectively}$$

$$I \quad = \quad \Sigma\, f_q/n_q \;\; {}^2\log\,\{(f_q/n_q) / (f_p/n_p)\}$$

$$= \quad \Sigma\, f_q/n_q \;\; \{{}^2\log\,(n_p/n_q) + {}^2\log\,(f_q/f_p)\}$$

$$= \quad ({}^2\log n_p - {}^2\log n_q) + (1/n_q)\,\{\Sigma\, f_q\, {}^2\log\,(f_q/f_p)\}$$

The right hand-term enables us to operate directly on the comparable cell values as relative frequencies. Using this formula, the normalization can be performed after the addition is completed. Note that n_q and n_p are different for each vector, but at the level of the complete database or matrix n_q and n_p are constants. In the latter case one can therefore also use $\{\Sigma\, f_q \log\,(f_q/f_p)\}$ directly as an indicator of change.

In addition to the total number of citations in each year (n_p and n_q), the *number of journals* involved in the citation process of each journal under study provides us with a third parameter for the normalization. This journal-specific citation window limits the width of the channel that can be used for producing probabilistic entropy. I shall indicate this number below with *N*.

Let me provide an example in order to explain in greater detail what I will do. Assume that journal A is cited in the year 1998 by Journals B, C, D and E. In 1999, Journal A is cited by Journals C, D, E, and F. The analysis focuses on the number of citations by Journals C, D, and E in this case, since these citations can be compared as relative frequencies in both these years.

On the one hand, the citation of Journal A by Journal F in 1999 can be considered as unpredictable in terms of the 1998 expectation. The inclusion of Journal F would lead to a division by zero (in the *a priori* cell) and therefore to an infinite information value. Only on its second occurrence can a citation be evaluated in terms of its contribution to *structural* change.

On the other hand, the disappearance of Journal B from the citation pattern of Journal A leads to a zero in the denominator and therefore to a term which is equal to zero by definition ($0 \cdot \log 0 \equiv 0$). In other words, the disappearance of events in the past does not add information to our expectation about what will happen in the future and, therefore, it does not add to the value of the dynamic indicator in the present.

Note that one cannot simply focus on the new journals added to the database. New journals are added both because existing fields can expand and because of new developments (cf. Garfield, 1990). The quest is for an indicator which picks up the signal of structural change in the (citation) distribution pattern among journals, but not restricted to the inclusion of new journals in the database.

6. Results

6.1 'Cited' and 'citing' at the level of the aggregated matrix

Since 1660 of the 1699 journals included in 1999 could be matched with journals in 1998, only these 1660 journals can contribute to the change in the overall citation pattern on the cited side.[4] The I_{cited} generated among the cited distributions between these two years was 19 millibits, while the total I_{citing} was 102 millibits, that is, more than five times as much. This result is consistent with the theoretical notion that the 'cited' side represents the archive of the journals, while 'citing' can be considered as the running operator generating the variation.

The corresponding values in the *Science Citation Index* were I_{cited} = 24 and I_{citing} = 88 millibits. Thus, the change of the 'cited' side is of the same order despite a considerably higher number of journals in the *Science Citation Index*. However, the change on the 'citing' side is even higher in the *Social Science Citation Index* than in the *Science Citation Index.* This difference reflects again the less codified nature of citations in the social sciences, on average. There is more change in citation behaviour.

Table 3 exhibits the top twenty journals in terms of their contribution to change in both the cited and citing dimensions. These journals are, in other words, sorted in terms of the ΔI to the change in the distribution of the total citations on either side.

[4] As noted, another two journals were not included as 'citing'.

cited	*citing*
ADV EXP SOC PSYCHOL	URBAN EDUC
STRATEGIC MANAGE J	MONOGR SOC RES CHILD
J EXP SOC PSYCHOL	INT J LANG COMM DIS
MIND LANG	ANN DYSLEXIA
SCHIZOPHR RES	THEOR MED BIOETH
CONSCIOUS COGN	MINN SYM CHILD PSYCH
J INTERPERS VIOLENCE	J BEHAV HEALTH SER R
J CHILD PSYCHOL PSYC	NAT RESOUR J
ACAD MANAGE J	PSYCHOL SCI
PERCEPTION	J ECT
J SPEECH LANG HEAR R	J PEDIATR PSYCHOL
ADDICTION	KOREAN J DEF ANAL
PSYCHIATR SERV	AUST PSYCHOL
NEUROBIOL LEARN MEM	U ILLINOIS LAW REV
COGNITION	J MARKET RES SOC
PERS INDIV DIFFER	VA LAW REV
CLIN NEUROPSYCHOL	NARRAT INQ
HEALTH AFFAIR	AM ANTHROPOL
HEALTH PSYCHOL	J MANAGE
COMP STUD SOC HIST	PSYCHOL HEALTH

Table 3

Twenty journals contributing most to the change of the overall citation pattern between 1998 and 1999, both in terms of 'being cited' and 'citing.'

Let us now consider whether these journals can serve as indicators of structural change. As noted, I performed factor analysis and drew the multidimensional plots for visual inspection so that I could analyze in detail whether the indicated changes were also structural when comparing the solutions for 1998 and 1999.

Using the first journal on the 'cited' side of the list (that is, *Advances in Experimental Social Psychology*) as a seed journal, structural change could be found between 1998 and 1999. On the basis of a one percent threshold, this journal draws 15 journals into its citation environment in 1999, as opposed to 49 in 1998. In both years, the journal loads highest on a first factor of its own citation environment, that is, with 0.952 in 1998 and 0.950 in 1999, but the composition of this factor has changed. While 18 journals obtained their first factor loading on this factor in 1998, only eight had remained so highly correlated in 1999. These latter were exclusively ‘social psychology’ journals, whereas in 1998 this factor was composed of other fields of psychology and more general psychology journals as well.

For example, *Social Cognition* and the *Annual Review of Psychology* entertained high factor loading on this first factor in 1998, while both journals had disappeared from the citation environment of *Advances in Experimental Social Psychology* in 1999. However, the factors in this citation environment, such as ‘applied psychology,’ ‘applied social psychology,’ and clinically oriented journals, were present in both years. Thus, it seems that we are witnessing here the relative closure of the citation environment of ‘social psychology’ which can be associated with specialty formation.

This impression is further enforced when we turn to the *Journal of Experimental Social Psychology*, which is the third journal on the list of cited journals in Table 3. In both years, this journal has the second position on the same first factor in the factor solution of the citation pattern of *Advances in Experimental Social Psychology*. In 1998, the highest correlation with this factor is exhibited by the *Personality and Social Psychology Bulletin* whereas the fourth position is taken by the *Journal of Personality and Social Psychology.* In 1999, these latter two journals have changed positions, but, in sum, the first four journals on this factor remained identical.

When the *Journal of Experimental Social Psychology* itself is used as the seed journal for generating the relevant citation environment (at the one percent level), the solutions for 1998 and 1999 are almost identical. There is a core set of journals in ‘experimental social psychology,’ there are two applied factors (one more specifically on social psychology and one more general), and in 1999 the *British Journal of Social Psychology* and the *European Journal of Social Psychology* show as a separate (third) factor.

In summary, we witness in 1999 a more focused citation pattern for 'experimental social psychology' corresponding to a stronger citation profile. What we observe here is a typical pattern of specialty formation in an experimentally oriented field. These scholars cite each other in 1999 to a considerable larger extent than in 1998. Let me note that we did not find this pattern of specialty formation in the corresponding analysis of the *Science Citation Index.* Perhaps, the explanation is that the developments in the natural and the life sciences among *existing specialties* overshadow the emergence of new specialties.

The second journal on the list in Table 3 is entitled the *Strategic Management Journal*. We find here a similar pattern as in the case above, but to a lesser extent. This journal taken as a seed journal draws into the analysis 24 journals in 1999 as against 33 in 1998. In both years it loads on a first factor with factor loading of 0.775 and 0.813, respectively. In both years, the factor is led by journals like *Academy of Management Journal, Academy of Management Review*, the *Administrative Science Quarterly*, and the *Journal of Management,* in both years. Note that among these, the *Academy of Management Journal* is listed as the ninth journal on the list of Table 3.

If we turn to this *Academy of Management Journal* itself, a pattern similar to the ones discussed above can be observed: 21 journals in 1999 versus 29 in 1998. In 1998, the journal loads with 0.854 on a first factor that is led by the *Administrative Science Quarterly* (0.914), while in 1999 the journal itself is the lead journal on a second factor (with a factor loading of 0.898) behind a first factor that is composed of journals focusing on organizational behaviour. (The lead journal of this group is then the *Journal of Organizational Behavior*). In 1998, this latter group was not yet that pronounced and related to the former group more through a communality with 'applied psychology' journals.

Let me emphasize that these developments are here monitored at a very fine-grained level since the underlying citation matrices are extremely sparse. As noted, the main structural element is provided by the missing values in the citation matrix. Thus, we are witnessing changes that may be feeble and perhaps reversible in subsequent years. But our results suggest that specialty formation is a main driving force in changing the citation patterns in the social sciences at each moment in time.

Specialty formation itself may be rapid and perhaps transient. For example, if we turn to the next journal on the list in Table 3, that is, the journal *Mind and Language*, we see the same cloud of journals around it in 1998 and 1999, but the cloud has shrunk from 30 to 24 journals. In 1999, the journal is positioned together with *Philosophical Psychology* and *Behavioral and Brain Sciences* on a fifth, but separate cluster, while in 1998 it was still an *isolate* in this citation environment. Thus, changes can be rapid; perhaps also because of the relative weakness of the citation relations.

In the *Science Citation Index,* we found at this level of aggregation mainly relative decline of specific clusters and increase of citation patterns of *existing* specialties. New developments could not be observed at this level because the relative changes in positions between existing specialties are more important than the weak signals of emerging fields and new codifications. We had therefore to turn to individual journals as indicators of qualitative development. Let us now make this same turn in analyzing the *Social Science Citation Index.*

6.2 Citation patterns at the level of individual journals

Table 4 provides the listing of journals that exhibited most change between these two years in terms of their being citedness at the level of the row vectors. Thus, the change is now no longer normalized in terms of the overall change at the matrix level, but we focus on change in the citation patterns for each single journal.

cited journals	*sorted on probabilistic entropy production (1999|1998)*	*N* =
ANTHROZOOS	2.088	2
AM BUS LAW J	1.945	4
RACE CLASS	1.768	2
ARMED FORCES SOC	1.590	4
SOC WORK RES	1.379	6
J REHABIL RES DEV	1.370	3
J INST THEOR ECON	1.306	13
COMMUNIST ECON EC TR	1.283	2
J BEHAV THER EXP PSY	1.277	36
NAT RESOUR J	1.245	9
J EDUC PSYCHOL CONS	1.191	4
AUDITING-J PRACT TH	1.181	6
INT J SERV IND MANAG	1.034	3
J CREATIVE BEHAV	1.007	4
COLUMBIA J TRANS LAW	1.001	10
EUR J COMMUN	0.972	3
J CAREER DEV	0.959	7
SPRACHE KOGNIT	0.929	4
CLIN PSYCHOL-SCI PR	0.929	20
J LEGAL EDUC	0.919	21

Table 4

Journals with changing citation patterns in the cited dimension in decreasing order.

The third column in this table lists the number of *citing* journals included in the comparison among the cited journal between the two years. This data will provide us below with a clue for the explanation of the sorting: the low numbers indicate an extremely volatile pattern of change among the years. Only a few citation relations are stable.

Anthrozoos, for example, draws into its environment 15 journals in 1999 against 33 in 1998, but in addition to this pattern known from above only two journals have remained stable in its relevant environment (that is, *Psychological Reports* and *Society & Animals.*). In 1998, *Anthrozoos* loads with *Psychological Reports* on an eight factor. Its citation environment is then related to social and development psychology journals.

In 1999, most psychology journals have become more distanced and sociological issues have come more to the fore, for example, *Sociological Quarterly* and the *Journal of Marriage and the Family*, the journal *AIDS Care* and the *J of Rehabilitation.* Relations with anthropology journals are failing in both years. In summary, the main characteristic of this journal is that it is moving through the database because it is not recognizable by its relevant citation environment as being specifically related to a single specialty cluster.

When we turn to the second journal on this list (the *American Business Law Journal*), we find a somewhat more stable environment in terms of the designation of other factors in its environment, but change prevails in its own being cited pattern. Among the 80 journals drawn into the analysis in 1999 (at the one percent level) only four were present among the 53 journals drawn into this analysis in 1998. Thus, the journal is cited by articles in an increasing number of journals, but also with increasing variety. In other words, these journals are loosely cited in different environments. Thus, they generate 'noise' in the database when comparing between years because the citation environment is almost completely replaced.

Note that this instrument provided us with the best indicator of structural change when we analyzed the *Science Citation Index*. In other words, the addition of new journals to the *Science Citation Index* is a major source of variation in the database, whereas existing journals usually exhibit rather stable patterns of aggregated citations by journals in their relevant environment. If the citation patterns among existing journals change, this indicates underlying changes in the perception of the relative importance of these journals.

In the *Social Science Citation Index*, however, change in citation patterns of individual journals seems to indicate 'looseness' in the orientation of the journal, whereas change at the level of the database indicates *specialty formation*. The relations among journals seem to generate variation endogenously, that is, as a movement in the positions of journals in relation to one another. Let me emphasize that these 'moving targets' perhaps fulfill other important (e.g., dynamic) functions, for example, at interfaces in the database. However, they cannot easily be used as the envisaged yardsticks for indicating change in underlying structures.

6.3 Other relevant dimensions

In addition to the cited dimension the journals can be analyzed in terms of their 'citing' patterns. As noted, the expectation from our experience with the *Science Citation Index* is that journals may vary in aggregated 'citing' behaviour from year to year without indicating structural change when a longer-term perspective is used. Highly codified communication structures wield out the variation in an ongoing process of recursive selections.

A similar pattern could be found in the *Social Science Citation Index.* For example, the journal in the *Social Science Citation Index* with strongest change in 'citing' patterns was *Library Trends*. This journal is part of a cluster of journals which all have the word "library" in their title in 1998, while it exhibits major factor loading on the cluster of information science journals around the *Journal of the American Society for Information Science* in 1999. However, this can be considered as a change in position given an existing interface between these two clusters. Journals can also be functional at interfaces; for example, the edition of a special issue may relocate their position in relation to relevant environments in a given year.

Another indicator which we mentioned above is the value of $\{\Sigma\ f_{1999} \log (f_{1999} / f_{1998})\}$ expressing the difference in terms of cell values between the years, but before normalization.

In the *Science Citation Index* this indicator provided us with an overrepresentation of the leading journals of the database because these journals are able to improve their citation profiles over the years at the expense of journals that are relatively loosing citation profile. Leading journals profit from the ongoing erosion of the fine structure in the database. One can consider this as a manifestation of the so-called "Matthew effect" in science (Merton, 1968) which states that "for he that hath, to him shall be given." (Matt. 4: 25).

cited journal (row vectors of the matrix)	$\Sigma\ f_{1999}\ {}^{2}\log\ (f_{1999}\ /\ f_{1998})$
J PERS SOC PSYCHOL	1257.01
CHILD DEV	1200.19
AM J PSYCHIAT	1104.36
J CONSULT CLIN PSYCH	1086.53
AM PSYCHOL	1085.52
ARCH GEN PSYCHIAT	1067.01
J AM ACAD CHILD PSY	1015.52
PSYCHOL BULL	1009.52
BRIT J PSYCHIAT	981.61
DEV PSYCHOL	974.42
STRATEGIC MANAGE J	938.79
AM SOCIOL REV	928.55
J CHILD PSYCHOL PSYC	901.68
J EXP PSYCHOL LEARN	878.42
AM ECON REV	877.18
J PERS ASSESS	876.23
PSYCHOL REV	870.18
ACAD MANAGE J	852.86
SCHIZOPHRENIA BULL	837.76
J EXP PSYCHOL HUMAN	833.27

Table 5

Top twenty journals sorted in terms of the entropy production at the level of the matrix.

Table 5 lists the journals of the *Social Science Citation Index* organized according to this value and it seems upon visual inspection that this effect is combined in this database with the structural effects of specialty formation that I described above. The *Journal of Personality and Social Psychology* which leads the list, for example, is part of the development described above, but it was not among the indicated journals after normalization. Major journals with the adjective "American" in their title, such as the *American Sociological Review* and the

American Journal of Psychiatry, are part of this list indicating an of size effect similar to the one noted in the case of the *Science Citation Index.*

A third indicator which we used in the *Science Citation Index* is the grouping by the ISI of journals under categories. ISI lists 55 such categories for the *Social Science Citation Index* as against 160 for the *Science Citation Index.* Since journals can be subsumed under more than one category, these "macro-journals" are very unequal in terms of their size (Cozzens & Leydesdorff, 1993).

The aggregated change at this level is highly correlated with the size of the cluster. In the *Science Citation Index* we found a Spearman's ρ of 0.971 between the probabilistic entropy generated and the cluster size (in the later year). In the *Social Science Citation Index* this value was 0.966, indicating a similar effect.

Since the probabilistic entropy of a macro-journal is based on a summation over the journals included, one can also divide by the number of journals in order to obtain a value for the average probabilistic entropy per journal. Table 6 lists these normalized values for the top twenty categories.

ISI-category	*I(cited) / N*	*number of journals N*
ETHNIC STUDIES	0.438	6
LAW	0.364	101
EDUCATION, SPECIAL	0.324	19
PSYCHOLOGY, EDUCATIONAL	0.316	37
HISTORY	0.298	13
ANTHROPOLOGY	0.289	48
PSYCHOLOGY, DEVELOPMENTAL	0.281	48
PSYCHOLOGY, EXPERIMENTAL	0.279	63
LANGUAGE & LINGUISTICS	0.278	39
PSYCHOLOGY, SOCIAL	0.270	40
SOCIAL WORK	0.270	28
SOCIOLOGY	0.269	86
BUSINESS	0.268	49
GERIATRICS & GERONTOLOGY	0.268	24
ENVIRONMENTAL STUDIES	0.267	42

FAMILY STUDIES	0.264	31
REHABILITATION	0.262	43
PSYCHOLOGY, PSYCHOANALYSIS	0.259	13
PSYCHOLOGY, CLINICAL	0.257	84
PSYCHOLOGY, MATHEMATICAL	0.255	10

Table 6

Twenty journal categories of ISI sorted according to the average amount of change per journal in this category (cited dimension).

The second cluster of 101 'law' journals is very large. The first cluster, indicated 'ethnic studies,' consists only of six journals indicating that these journals are gaining importance in terms of "being cited". Because of the noted size differences between macro-journals, however, this indicator seems too confused to be useful. Furthermore, the aggregation rules of the ISI follow automated attribution principles based on *ex ante* criteria that are kept stable over the years under study, whereas our analysis is precise at the level of individual journals and sensitive to changes in the clustering.

7. Conclusions

Using entropy statistics I have explored whether and how changes in citation patterns can be used as indicators of structural change in the *Social Science Citation Index*. The exploration was mainly methodological. For example, this study was restricted to change between only two subsequent years.

The conclusion is that various structures operate as subdynamics which can be distinguished in operational terms. These dynamics are in important respect different from those in the *Science Citation Index*. The generative mechanisms for concepts like "impact" therefore are different and this may have implications for the use of scientometric indicators in evaluation studies.

Let me first note that we did not find effects from hierarchically higher-level journals equivalent to *Science* and *Nature* as are well-known in the *Science Citation Index*. Journals which play a role in a multitude of fields albeit only in the relevant citation environments, are less obvious in the social sciences. Perhaps, this next-order role of reviewing the research process is attributed to books and monographs. The single exception seems to be the *American Sociological Review* which was visible in various environments as a source of citations. However, because of its strong disciplinary affiliation this journal cannot be considered as a 'general science' journal.

Since the citation network is much 'thinner' than in the natural and the life sciences, the dense links within the specialties are very specific. However, there are journals which are more loosely connected. They generate variety and change from year to year. These latter journals could for this reason not be used as indicators of structural change, as was the case in the *Science Citation Index*. Sometimes, one wonders whether these journals should have been included in the database. Does the less codified nature of the citation in the social sciences make it more difficult for the ISI to perform the selection of journals to be included with sufficient precision?[5]

We were able to use change at the level of the overall database as indicators of specialty formation. Given the low values of citations, local densities can be detected in this database more easily than in the *Science Citation Index.* For example, the difference between a value of four versus one in the previous year, is numerically equivalent to a change from 104 to 101 as a difference. However, we use the quotient in the formula and the size effects are then far more pronounced with small numbers. As noted, we were not able to use this indicator in the *Science Citation Index*.

Another effect which was very clear in the *Science Citation Index*, but which we did not note at all in these analyses, is that a major source of change in the natural and life sciences is the disappearance of specialties by further inclusion in a larger set, for example, at the disciplinary level. We did not find this ongoing process of codification in our analysis of the *Social Science Citation Index*, while it seems the major source of change in the *Science Citation Index.* In other words, not the emergence of new specialties, but the inclusion of detailed specialization into larger clusters was the major movement in the *Science Citation Index*. The winners of this erosion of fine-structure are the major journals. We have mentioned the Matthew-effect in this context. Some of that latter effect could also be witnessed in the *Social Science Citation Index.*

In summary, the *Social Science Citation Index* can be considered as considerably less codified, more volatile, internally developing, etc., when compared to the *Science Citation Index.* Specialty formation has an impact at the level of the database, while specialty erosion through codification was the main effect in the *Science Citation Index.* While maintaining citation profile under codification pressure can be associated with impact in the sciences (Small, 1978), the looseness of the specialty structure in the social sciences and the rapid turn-over make impact of individuals and institutions dependent on rapid and transient developments at the level of specialties.

---

[5] In another study we found similarly volatile patterns of citations in engineering sciences and physics in the case of wind and solar energy journals when these fields were under increased policy pressure during the 1980s (Leydesdorff & Van der Schaar, 1987).

To a larger extent than in the *Science Citation Index* the use of impact factors for performance measurement of institutional agencies—let alone individual scholars—seems very doubtful when using this database. The data provide a thin layer of what is going on underneath and because of the volatility relatively small changes in parameter choices (e.g., citation windows; cf. Price, 1970) may lead to dramatically different conclusions. The use of indicators which are based on the citation statistics of the *Science Citation Index* is therefor ill-advised when using data from the *Social Science Citation Index.*

References

Bar-Hillel, Y. (1955). An Examination of Information Theory. *Philosophy of Science, 22,* 86-105.

Burt, Ronald S. (1982). *Toward a Structural Theory of Action.* New York, etc.: Academic Press.

Callon, Michel, John Law, and Arie Rip (Eds.) (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 of Information Science, 24*, 425-436.

Cozzens, Susan E., and Loet Leydesdorff (1993). Journal Systems as Macro-Indicators of Structural Change in the Sciences. In: A. F. J. Van Raan, R. E. de Bruin, H. F. Moed, A. J. Nederhof and R. W. J. Tijssen (Eds.), *Science and Technology in a Policy Context* (pp. 219-233) Leiden: DSWO Press/Leiden University.

De Looze, Marie-Angèle, R.Cronini, M. Legentil, P. Jeannin, & M.-H. Magri (1996). Determining the Core of Journals of a Research Centre: The Example of Researchers from the Department of Rural Economy and Sociology of the Institut National de la Recherche Agronomique, France, *Scientometrics* 36(2), 167-183.

Doreian, Patrick, and T. J. Fararo (1985). Structural Equivalence in a Journal Network. *Journal of the American Society of Information Science, 36,* 28-37.

Eisenstein, Elizabeth L. (1983). *The Printing Revolution in Early Modern Europe*. Cambridge: Cambridge University Press.

ESRC (1995). *Identifying Innovation in Social Science: Some Bibliometric Approaches*. London: Economic and Social Research Council.

Garfield, Eugene (1979). *Citation Indexing*. New York: Wiley.

Garfield, Eugene (1990). How ISI selects Journals for Coverage: Quantitative and Qualitative Considerations, *Current Contents* (28 May 1990) 5-13.

Healey, Peter (1997). *International Social Science Comparison. Report for the ESRC*. London: Science Policy Support Group.

Kaufer, David S., & Kathleen M. Carley (1993). *Communication at a Distance: The Influence of Print on Sociocultural Organization and Change.* Hillsdale, NJ: Erlbaum.

Krippendorff, Klaus (1986). *Information Theory. Structural Models for Qualitative Data.* Beverly Hills, etc.: Sage.

Leydesdorff, Loet (1986). The Development of Frames of Reference, *Scientometrics, 9*, 103-125.

Leydesdorff, Loet (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, Loet (1998). Theories of Citation? *Scientometrics, 43*(1), 5-25.

Leydesdorff, Loet (forthcoming). Indicators of Structural Change in the Dynamics of Science: Entropy Statistics of the SCI Journal Citation Reports, *Scientometrics* (forthcoming).

Leydesdorff, Loet, & Peter Van der Schaar (1987). The Use of Scientometric Indicators for Evaluating National Research Programmes. *Science & Techology Studies 5,* 22-31.

Leydesdorff, Loet, & Susan E. Cozzens (1993). The Delineation of Specialties in Terms of Journals Using the Dynamic Journal Set of the Science Citation Index, *Scientometrics, 26,* 133-154.

Leydesdorff, Loet, Susan Cozzens, & Peter Van den Besselaar (1994). Tracking areas of strategic importance using scientometric journal mappings, *Research Policy, 23,* 217-229.

Leydesdorff, Loet & Élaine Gauthier (1996). The Evaluation of National Performance in Selected Priority Areas using Scientometric Methods, *Research Policy, 25,* 431-50.

Merton, Robert K. (1968). The Matthew Effect in Science, *Science* 159, no. 3810 (5 January): 56-63.

Narin, Francis (1976). *Evaluative Bibliometrics.* Cherry Hill, NJ: Computer Horizons, Inc.

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.

Nederhof, A. J., R. A. Zwaan, R. E. De Bruin, and P. J. Dekker (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural science: A comparative study, *Scientometrics, 15*(5/6), 423-435.

Price, Derek de Solla (1961). *Science since Babylon*. New Haven: Yale University Press.

Price, Derek de Solla (1970). Citation measures of hard science, soft science, technology and non-science. In: C. E. Nelson & D. K. Pollock (Eds.), *Communication among Scientists and Engineers* (pp. 3-22). Lexington, MA: D. C. Heath & Co..

Shannon, Claude E. (1948). A Mathematical Theory of Communication, *Bell System Technical Journal, 27,* 379-423 and 623-56.

Simon, Herbert A. (1969). *The Sciences of the Artificial*. Cambridge, MA/ London: MIT Press.

Small, Henry (1978). Cited documents as concept symbols, *Social Studies of Science*, *7,* 113-122.

Theil, Henry (1972). *Statistical Decomposition Analysis*. Amsterdam/ London: North-Holland.

Tijssen, Robert, J. de Leeuw, A. F. J. van Raan (1987). Quasi-Correspondence Analysis on Square Scientometric Transaction Matrices, *Scientometrics, 11,* 347-361.

Tijssen, R. J. W. (1992). *Cartography of Science: scientometric mapping with multidimensional scaling methods*. Leiden: DSWO Press/ Leiden University.

Van den Besselaar, Peter & Loet Leydesdorff (1996). Mapping Change in Scientific Specialties: A Scientometric Reconstruction of the Development of Artificial Intelligence, *Journal of the American Society for Information Science, 47,* 415-36.

Wouters, Paul (1999). *The Citation Culture*. Unpublished Ph.D. Thesis, University of Amsterdam.