BETTER LATE THAN NEVER? ON THE CHANCE TO BECOME HIGHLY CITED ONLY BEYOND THE STANDARD BIBLIOMETRIC TIME HORIZON

Wolfgang Glänzel[a,b], Balázs Schlemmer[a], Bart Thijs[a]

[a] *K.U. Leuven, Steunpunt O&O Statistieken, Dept. MSI, Leuven, Belgium*
[b] *Hungarian Academy of Sciences, IRPS, Budapest, Hungary*

Abstract. According to Garfield (1980), most scientists can name an example of an important discovery that had little initial impact on contemporary research. And he uses Mendel's work as a classical example. Delayed recognition is sometimes used by scientists as an argument against citation based indicators based on citation windows defined for a short- or medium-term initial period beginning with the paper's publication year.

This study is focussed on a large-scale analysis of the citation history of all papers indexed in the 1980 annual volume of the *Science Citation Index*. The objective is two-fold, particularly, to analyse whether the share of delayed recognition papers is significant and whether such papers are typical of the work of their authors at that time.

In a first step, the background of advanced bibliometric models by Glänzel, Egghe and Rousseau, Rousseau and Burrell of stochastic citation processes and first-citation distributions is described briefly. The second part is devoted to the bibliometric analysis of first-citation statistics and of the phenomenon of citation delay. In a third step, finally, delayed reception publications have been studied individually. Their topics and the citation patterns of other papers by the same authors have been studied to uncover principles of regularity or exceptionality of delayed reception publications.

1 INTRODUCTION

One of the frequently used arguments against the use of citation based indicators in research evaluation is that important publications and, among them, also breakthroughs in science are often not cited in the beginning, and only become recognised in a time that is beyond the standard citation windows used in most bibliometric studies. In his essay entitled "Premature discovery or delayed recognition - Why?", Garfield (1980) summarises reasons for what is called *premature discovery*. Although, there is - according to his telephone communication with Stephen Cole - no way to predict whether a paper will become premature or not, delayed recognition papers can be identified on the basis of their citation history (Garfield, 1980). On the other hand, bibliometricians always stress that *from the statistical viewpoint* such papers do not really play an important part. However, there is no systematic bibliometric study on this issue so far.

The present study, therefore, focuses on scientific papers that are lacking significant reception during a certain period immediately following the date of their publication, and are becoming highly cited only in a time being beyond the normal bibliometric horizon. In particular, this phenomenon is basically studied along the following questions:

The final publication is available at Springer via http://dx.doi.org/10.1023/B:SCIE.0000006881.30700.ea

1. In how far is delayed reception typical of scientific communication in general and in given subject fields?
2. Are delayed citations indicating a breakthrough in the reception and thus outstanding future citation frequencies or are they just "shifting" the citation process by several years?
3. Are such papers are typical of the work of their authors at that time, or have they to be considered rather "singular"?

The results of this analysis are expected both, to understand the mechanism of delayed citation and to contribute to the validation of bibliometric indicators.

2 DATA SOURCES AND DATA PROCESSING

In the first part, all papers indexed in the 1980 annual volume of the *Science Citation Index* (SCI) of the Institute for Scientific Information (ISI - Thomson Scientific, Philadelphia, PA, USA) not cited or poorly cited in an initial period of 3 and 5 years, respectively, were analysed for delayed reception in a 21-year period beginning with the year of publication. Citations have been determined on the basis of an item-by-item procedure using special identification-keys for all papers of the type Articles, Letters, Notes and Reviews. Journal impact measures were also calculated on an item-by-item basis for a 21-year period, and then used as reference standard for the observed citation rates.

3 METHODS AND RESULTS

3.1 *Theoretical considerations*

The most convenient way of studying delayed reception of scientific literature is the analysis of both, the number and the succession of citations an individual paper receives. The first simple indicator designed for the measurement of citation immediacy on the basis of citation succession was the Mean Response Time (MRT) introduced by Schubert and Glänzel (1986). The indicator was based on an observation period of five years, beginning with the publication year. Other studies concerned with the theory of first citation have been published by Egghe and Rousseau (1990), Rousseau (1994) and Burrell (2001). In 1992, Glänzel has extended the results by Schubert and Glänzel (1986) through the application of a stopping time approach to (stochastic) citation processes. In that paper, both a continuous and a discrete version of the process have been used. Taking bibliometric counting practice into account, we will restrict this method to the discrete case in the following.

We now take up the random succession of citations the papers have received during a certain time span t. Without loss of generality, we put $t = 0$ for the publication year, and consider citation windows for any $t \geq 0$. To obtain the factual publication and citation year analysed in this study, the value of parameter t has then to be added to the publication year 1980. The citation process, that is, the number of citations received between the publication year and a given year t is denoted by $X(t)$. Let T_i denote the shortest time t during which the papers have received exactly $i = 1, 2, 3, \ldots$ citations, that is, $T_i = \min\{t : X(t) \geq i\}$. Random variables of this type are called

"stopping times". Note that $T_i \geq 0$, where the event $\{T_i = \infty\}$ might have a positive probability, namely, if and only if the paper will never receive *i* citations. The measurability of the events $T_i \leq t$ for each $i > 0$ and every $t = 0, 1, 2, \ldots, \infty$ is obvious. In particular we obtain the following properties (cf. Glänzel, 1992).

1. $P(T_i = t) = P(X(t) \geq i) - P(X(t-1) \geq i)$
2. $P(T_i = 0) = P(X(0) \geq i)$
3. $P(T_i \leq t) = P(X(t) \geq i)$
4. $P(T_i = \infty) = P(X(\infty) < i)$

In what follows, statistical functions on the citation process such as (conditional) mean values and first-citation statistics will be presented and analysed for selected subject fields and the total of about 450,000 papers indexed in the 1980 volume of the SCI.

3.2 *First-citation statistics and the bibliometric approach to citation delay*

The stopping time approach shows that the distribution of papers cited first T_1 years after publication is extremely skewed (see Figure 1). In this figure, we put $t = 0$ for the publication year, that is, for instance, $t = 7$ means that the paper has first been cited in 1987. $P(T_1 = t)$ here denotes the relative frequency for stopping time T_1, that is, the estimate of the corresponding probability. 21.5% of all paper published in 1980 have not received any citation between the publication year and 2000. Most of these papers will probably never be cited. About 28% of all papers (that is, 36% of all *cited* papers) have been cited in the year subsequent to that in which the paper appeared. 60% of all papers have been cited between 1980 and 1982 and, most remarkably, 76% of all papers cited in the 21-year period were cited in an initial periods of three years. In other words, less than one quarter of all papers not cited between 1980 and 1982 are cited in the subsequent 18 years and less than 12% of papers not cited between 1980 and 1984 are cited till 2000. For a period of ten years, the corresponding share is around 3%. The question arises, how many citations these 24%, 12% and 3% of so far uncited papers received in a later period and how many papers of them will become truly *highly cited*.

Before we answer this question, we have a look at the breakdown by selected science fields. Figure 2 presents the same data for three selected fields, namely, Biomedical Research (representing the life sciences), Chemistry (representing the natural sciences) and Mathematics. The skewness of the first-citation distributions in the three fields is significantly different. The share of uncited papers is with 13% the lowest in biomedical research. The distribution is here extremely skewed; almost two thirds of all cited papers are cited in the first two years, more than 80% are cited in an initial 3-year window and more than 90% in an initial 5-year citation window. The first-citation distribution in chemistry is less skewed and the share of uncited papers is at the same time higher than that in the selected life-science area. Practically one third of all papers concerned with mathematics remains uncited. That is by far the highest share of the selected fields. The first-citation distribution is much less skewed. Only 60% of all cited papers receive their first citation within an initial period of three years; the corresponding share for an initial window of five years amounts to not quite 80%. Hence we derive the assumption that the share of delayed recognition papers in mathematics might be higher than that in natural and life sciences if we

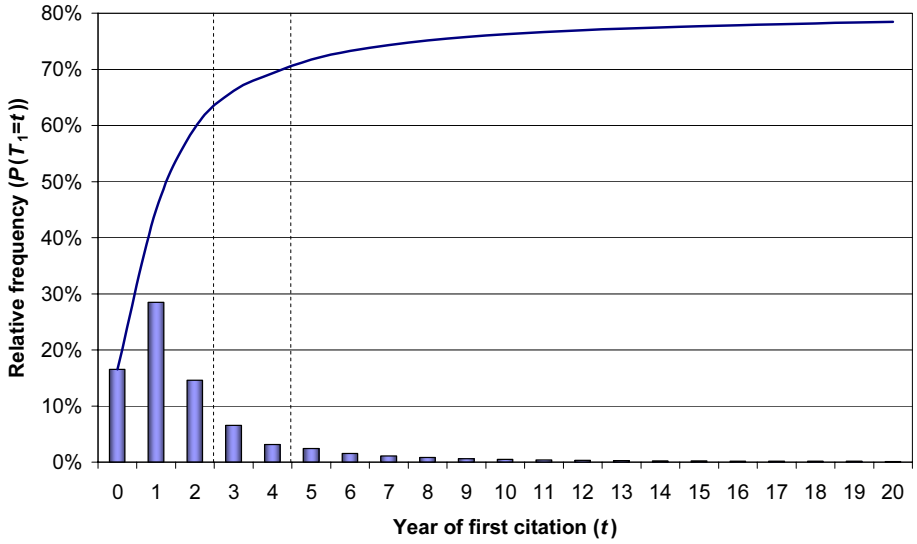

FIGURE 1. Relative frequency of the time of first citation (all fields combined)

base the criteria for delayed response on initial citation windows of thee or five years, respectively. Here we just mention in passing that the field-specific peculiarities in delayed reception should also be interpreted in the context of different ageing literature in these science areas (see Glänzel and Schoepflin, 1995).

In order to decide which papers have become highly cited at the end of the 21-year period, we will analyse the citation impact of papers with delayed response. Figure 3 (top) presents the conditional citation impact $M(X(20)|T_1 = t)$ with $t = 0, 1, 2, \ldots, 20$ years after publication. The impact decreases dramatically; the mean observed citation rate of papers that have been cited first time four years after publication reaches just 17% of the impact of papers that have been cited immediately. The corresponding percentage for the sixth year after publication amounts to 10%. For the eleventh year it already drops to 5%. This is certainly not due to field-specific changes caused by different ageing as show above. However, the remaining period for possible citing decreases in our sample based on a 21-year observation period. In particular, if a paper received its first citation in 1980 or 1981, it can still be cited in the subsequent 20- or 19-year period. For a paper cited first in 1990, the potential period for receiving further citations reduces to 10 years. Although, citation processes are not homogeneous in time (cf., Glänzel and Schoepflin, 1994), normalisation by the remaining period might compensate the bias caused by shrinking citation windows. This is illustrated by Figure 3 (bottom). The shape of the curve for first-citation beyond 1993 is hidden because another bias appears at the end of the period. On one hand, the number of underlying papers decreases dramatically causing an extreme increase on the standard error on mean and, on the other hand, $M(X(20)|T_1 = 20)/1 \approx 1.0$ (cf. chart at top of Figure 3) since the overwhelming share of papers cited first in the last year of our observation period are only cited once. In order to eliminate this bias, we decided to omit the last seven years. The normalisation has only little effect: The mean observed citation rate of papers that have been cited first time three, six and eleven years after publication reaches now 20%, 14% and 11%, respectively, of the impact of

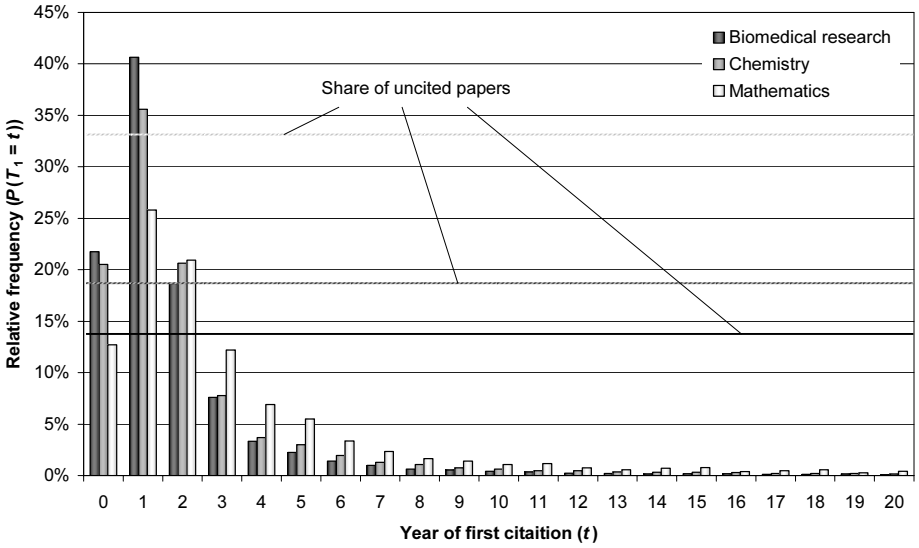

FIGURE 2. Relative frequency of the time of first publication (for three selected fields)

papers that have been cited immediately. Nevertheless, the shape of the curve is otherwise almost identical with that of the unadjusted case (top diagram of Figure 3).

We can conclude that a paper not cited within an initial period of three or five years, respectively, has a much lower expected citation impact than a paper cited immediately or the year subsequent to the publication year. Figure 4 shows the detailed distributions of citations over papers, provided they have not been cited within three or five years, and compares them with that of all papers. The chance of uncitedness increases dramatically with belated first-citation; the chance of being frequently cited later on decreases to the same extent. This is in line with the results in the context of possible prediction of later citation impact on the bases of observed initial citation rates according to Glänzel and Schubert (1995). Figure 4 thus visualises that delayed reception does not result simply in a "shift" of the citation process by several years.

As already discussed in the context of first-citation distribution (see Figure 1), 21.5% of all paper published in 1980 have not received any citation between the publication year and 2000. If the papers have not been cited between 1980 and 1994, then the corresponding share of uncited papers amounts to 70%. In turn, 25% of all papers index in the 1980 SCI volume received more than 15 citations, but only 0.3% of those papers not cited between 1980 and 1984 were cited more frequently than 15 times.

3.3 *Individual analysis of delayed recognition papers*

The results presented above, can be summarised as follows. Delayed reception is not a typical phenomenon. Although the extent of delayed reception deviates in different science fields and mathematics is characterised by reception slower than in other fields, delayed recognition remains the exception to the role in all fields. From the statistical point of view, a paper not yet cited, say, ten years after publications has little chance to become even fairly cited in the future. The shape of the

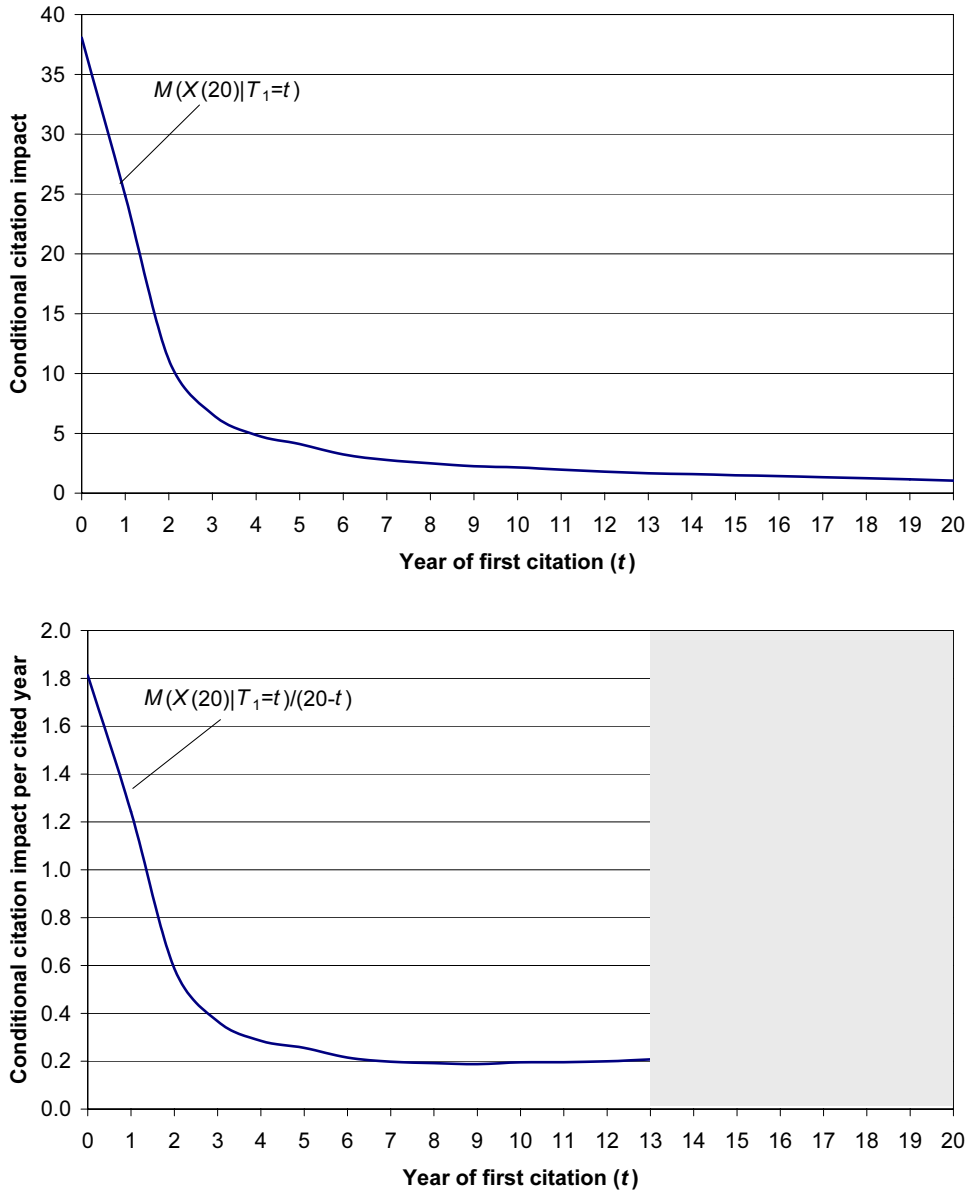

FIGURE 3. Conditional mean citation rates of papers cited first at time *t* after publication in all fields combined (top: Mean observed citation rate, bottom: the same, but normalised by the remaining citation window)

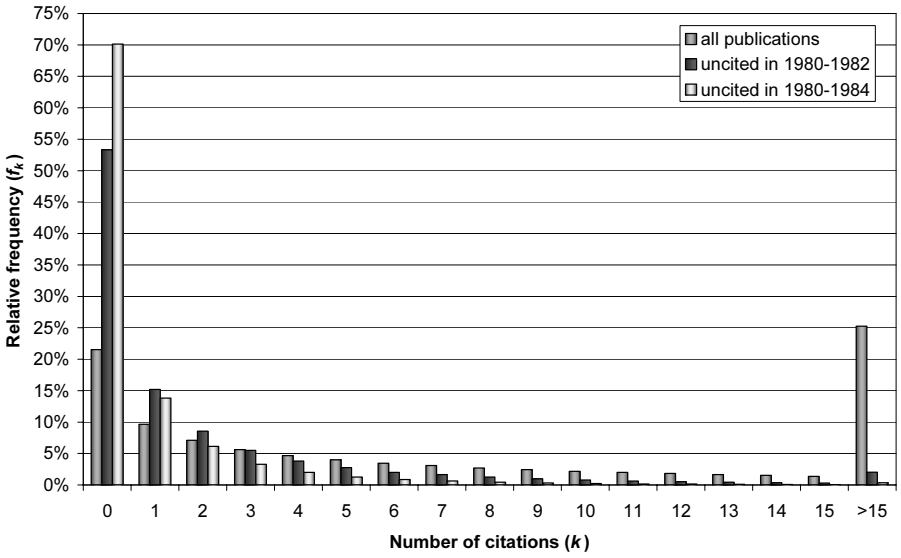

FIGURE 4. Citation distributions of all papers indexed in the 1980 volume of the SCI and of those not cited initially

citation distribution considerably changes if papers are not cited initially. Especially the chance of becoming highly cited decreases with each year of citation delay. Nevertheless, we will make the attempt to search for highly cited papers among those initially uncited or at least poorly cited. We have used a first statistic on this issue applied the following criteria. We consider a paper having delayed reception if it has received (a) only one citation in an initial 3-year period or (b) at most two citations in an initial 5-year period and it is highly cited if it has received at least 100 citations in the remaining period till 2000. The first criterion allows for one or two possible self citation(s). As a result, we have obtained 77 papers out of the almost 450,000 publications index in the 1980 volume under the weak condition and 29 papers under the stronger condition. The second set proved practically to be a subset of the first one although two citations were allowed in this case and these two citations might be received within the first three years. One problem has resulted in reconsidering these selection criteria. In particular, several journals proved to have journal impact measures based on a 21-year citation window that are around 100 or even higher. Several papers cited about 100 times were just close to their expectations on the basis of the impact of the journals in which they have been published. Figure 5 presents the rank statistics for high-impact journal in a log-log diagram. We just mention a passing that the plot of the 1000 journals ranking highest form an almost perfect straight line in the diagram ($r = 0.998$). 100 journals ranking highest have impact measures of about 50 and higher. This resulted in the decision in favour of basing the selection criteria for highly cited papers on both absolute citation rates and journal impact. Finally, we are considering a paper highly cited if it has received at least 50 citations and 10 times the journal impact. This procedure is a modified version of the method introduced by Glänzel and Schubert (1992).

This selection resulted in a set of 60 (weak condition, based on a 3-year initial period) and 16 papers (strong condition, based on a 5-year initial period). Again, the latter set proved to be a

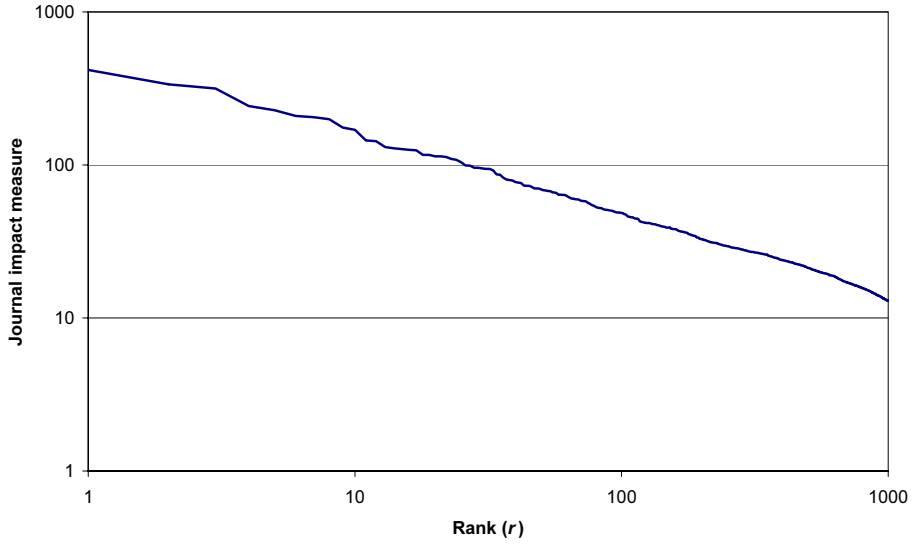

FIGURE 5. Rank frequency of journal citation impact based on a 21-year citation period

subset of the first one. The distribution of papers in the first set over subject categories according Glänzel et al. (2002) is presented in Table 1. Although the numbers in Table 1 are rather small, the relative high shares in mathematics, physics and engineering compared with the subject coverage of the database (cf., Glänzel et al., 2002) are quite striking. In order to answer the third question addressed in the outset, we have selected one quarter of the 60 papers identified as highly cited and having an especially interesting citation history. Eight papers were taken from the second set (strong condition), seven papers from the first set but not covered by the second one have been selected.

Table 2 presents a bibliographic description, the citation rates and journal impact of these papers. Since we do not intend to review the work of individual scientists here, we do not indicate the authors' names and we are just using an internal identification number. However, all papers can be retrieved on the basis of their bibliographic description. The selected papers represent all major fields in the life sciences, physics, chemistry, mathematics and engineering.

Most of the authors of the publications given in Table 2 have only several other papers published in the period 1980–1990 that can be considered having "remarkable" or "outstanding" citation impact. And practically all of those papers had a "regular" citation history with several citations received in the publication year or at least in the subsequent two years. For instance, the author of paper #1 had two additional papers in the mid of the 80 that have received more than 50 citations but those had regular citation history.

The author of paper #2 have about 100 other publications in the period 1980–1990. About 14% of their papers have received more than 50 citations, and can therefore be considered remarkably cited; however, citation rates around 100 are in this topic (biochemistry/neurochemistry) not at all exceptional. All these papers have regular citation history, and would have been identified as at least remarkably cited papers in any standard citation window. However, they do by far not reach the citation rate by the singularly cited paper #2.

TABLE 1. Subject representation of highly cited delayed recognition papers

Science field	Papers
Agriculture & Environment	4
Biology (Organismic & Supraorganismic Level)	3
Biosciences (General, Cellular & Subcellular Biology; Genetics)	6
Biomedical Research	2
Clinical and Experimental Medicine I (General & Internal Medicine)	3
Clinical and Experimental Medicine II (Non-Internal Medicine Specialties)	6
Neuroscience & Behaviour	2
Chemistry	7
Physics	13
Geosciences & Space Sciences	1
Engineering	7
Mathematics	7

The author of paper #3 has one other quite highly cited paper with almost 100 citations the history of which is regular. In the late recognition paper, the author gave a clear physical interpretation and mathematical foundation for the 'grey-world method'. The author's name is already used eponymously.

The author of paper #4 has no other remarkably cited papers in the period 1980–1990. The concept of 'fractional Fourier transforms' was first introduced in this paper. This paper has received one self-citation in 1980, but not cited by other authors till 1988.

Paper #5 presented first in-depth study trying to explain the mechanism of the yellow luminescence (YL) in Gallium Nitride. This paper was practically not cited (<3 citations) till 1992. From 1993 on, this paper has received 20–30 citations per year. YL has become a hot topic that started its dynamic growth in the early 90s. The authors have no further paper in the period 1980–1990 with remarkable or outstanding citation impact.

Paper #6 has become a key publication in numeric analysis. The author has published a generalisation of the results in 1986. Also this paper has become highly cited, however, it has a regular citation history. Nevertheless, paper #6 has just about 5% of its citation received through co-citation with the later publication.

Paper #7 has six co-authors the names of which appear in non-alphabetic order in the by-line of the publication. In order to keep retrieval within reasonable limits, we have retrieved publications of the first author where he might occur also as co-authors. The citation-context analysis of this Soviet paper proved difficult. We have, however not found any other highly cited publication of the first author.

The author of paper #8 has only two publications indexed by the SCI database. The second paper published in 1984 in the same journal and was concerned with the same topic. In paper #8, the author suggested the idea of a temporary urethral stent, and gave a description of the first intraprostatic partial catheter that has become known as the "urological spiral". The paper has received not more than two citations till 1988; although it has been published in German language, it became quite highly cited in the 90s. The second paper in 1984 (Urological spiral II) also published in German has attracted less citations. This paper has not been cited between 1984

TABLE 2. Citation data on 15 selected highly cited delayed recognition papers

ID#	Citations till			Impact	Journal Title	Year	Vol	1st page
	1982	1984	2000					
1	0	1	148	13.04	Acta Crystallographica Section B – Structural Science	1980	36	2042
2	1	1	732	63.73	Journal of Biological Chemistry	1980	255	896
3	0	0	77	4.23	Journal of the Franklin Institute	1980	310	1
4	1	1	127	6.57	Journal of the Institute of Mathematics and Its Applications	1980	25	241
5	0	0	165	10.62	Japanese Journal of Applied Physics	1980	19	2395
6	2	2	281	13.30	Numerische Mathematik	1980	35	315
7	0	0	71	3.04	Soviet Physics Semiconductors – USSR	1980	14	1078
8	2	2	87	3.47	Urologe-Ausgabe A	1980	19	236
9	0	19	51	3.26	Deutsche Medizinische Wochenschrift	1980	105	1779
10	1	8	257	16.12	European Journal of Medicinal Chemistry	1980	15	363
11	1	3	179	15.30	Journal of Theoretical Biology	1980	87	237
12	1	12	247	19.99	Polymer	1980	21	595
13	1	13	239	21.14	Psychological Bulletin	1980	88	588
14	0	8	1151	24.93	Soil Science Society of America Journal	1980	44	892
15	0	1	61	5.79	Veterinary Record	1980	107	390

and 1998, either. The two papers have been frequently, but not always been co-cited. There seems to be a trend in favour of citing the "original" one.

Paper #9 has four co-authors indicated in non-alphabetic order. We have retrieved only the papers of the first author published in the period 1980–1990. The paper has been published in German, and is concerned with clinical and experimental medicine. Even in this field, the citation rate of 51 is on the borderline of what can be considered highly cited. However, we have selected this paper because of its remarkable citation history. The paper has received its first citation in 1983, and then attracted all citations till 1991. After 1991, the paper has not been cited any more. The paper might thus be considered already "obsolete"; moreover, the author has more recent and highly cited papers on topics related with tick-borne diseases (lyme-disease, borreliosis , erythema migrans, meningoencephalitis) having regular citation history. An interesting detail is worth mentioning, namely that the author has published the majority of his papers in German. About 40% of citations paper #9 received in till 1984 stems from papers in German; this share decreased to about 20% of citations given after 1984.

Paper #10 has four co-authors in non-alphabetic order. Two other papers of the first author published in 1988 and 1990 have regular citation history. It is quite interesting that paper #10 is frequently co-cited with two papers published in 1987 and 1989 the author of which never acted as co-author of the first author of #10. The co-citations result in about half of all citations.

The fate of paper #11 is utmost interesting. It is a mathematical paper published in a journal of theoretical biology, and has been frequently cited in physics journals. In this paper, the author suggested a model to describe the adsorption of proteins on solid surfaces. The Random Sequential Adsorption (RSA) model has become very popular. The Diffusion RSA (DRSA) introduced by colleagues in 1992 is an extension of this model. Five other papers of the author published between 1980 and 1990 on the same and related topics proved to be highly cited, too. They have, however, a regular citation history.

Paper #12 is on polymer synthesis. The paper has two co-authors. Above all, the second co-author has several other highly cited papers whether in this topic or those were concerned with polymeric and non-polymeric materials with (potential) application to molecular electronics. All these papers have a regular citation history.

Paper#13 would not be an element of our set if the *Social Sciences Citation Index* were used in addition. In this paper on goodness-of-fit and model choice for covariance structure models for psychologists, the authors have proposed measures (*fit indices*) for comparing the goodness of fit of model as compared to a baseline model. It has received more than 2000 citations from papers indexed in the SSCI in the period 1980–2000; and it has already been frequently cited since 1981. Although it has been found highly cited with a certain delay on the basis of the SCI, too, it is above all standards if citations are retrieved from both databases, the SCI and the SSCI.

Paper #14 is by far the most cited paper by the author in the period 1980–1990. Other papers have received remarkable citations rates, too, but their impact is by one order lower. Nevertheless, except for one paper those publications have a regular citation history. The author of paper #14 introduced in this paper a 'soil water retention function' which has become a standard numerical model in soil and groundwater hydrology.

Paper #15 has three co-authors. The first author has dealt with research in 'listeriosis' in the 80s; the co-authors were above all dealing with cattle diseases, most recently with Bovine Spongiform Encephalopathy (BSE). One of them has actually co-authored the first paper reporting on this new form of progressive spongiform encephalopathy in the United Kingdom previously unknown in cattle. This paper appeared in 1987, and has become almost immediately highly cited.

Although the subject coverage of the fifteen selected papers seems at first sight to be representative for the SCI database, the high share of mathematical, numerical and statistical papers in part published in non-mathematics journals was quite striking.

3.4 *On possible errors in determining individual citation rates*

Whenever citation rates are determined by computerised data processing, the resulting figures are affected by measurable errors. These errors are only in part a consequence of spelling variances or misspellings. However, at the meso and macro level one can assume that all units are large and by to the same extent subject to such errors. The extent of errors is, however, unpredictable and differs among individual papers. At the level of individual papers, such errors might result in certain distortions. We will, therefore, point to this problem here. Basically, we can distinguish the following four main sources of errors.

1. Errors caused by citing authors;
2. Errors caused by the editors of the journals;
3. Errors caused by the database producer;
4. Errors caused by the user of the bibliographic database.

The authors themselves are responsible for probably most errors having effect on citation analyses. We will therefore restrict our comments to this type of errors among which we find, above all,

TABLE 3. Example for an incorrectly cited publication

Cites	1st Author	Journal	VOL	BP	PY
113	SCHUBERT A	SCIENTOMETRICS	16	3	1989
3	SCHUBERT A	SCIENTOMETRICS	16	3	*1988*
1	SCHUBERT A	SCIENTOMETRICS	16	3	*1987*
2	*BRAUN T*	SCIENTOMETRICS	16	3	1989
12	SCHUBERT A	SCIENTOMETRICS	16	*1*	1989
1	SCHUBERT A	SCIENTOMETRICS	16	*8*	1989
1	SCHUBERT A	SCIENTOMETRICS	16	*18*	1989
1	SCHUBERT A	SCIENTOMETRICS	16	*218*	1989
1	SCHUBERT A	SCIENTOMETRICS	16	*239*	1989
1	SCHUBERT A	SCIENTOMETRICS	16	*432*	1989
1	SCHUBERT A	SCIENTOMETRICS	16		1989
1	*SCUBERT A*	SCIENTOMETRICS	16	3	1989

misspelling of author names and incorrect citation of publication year, volume and first page of cited work.

The following example for errors caused by citing authors might illustrate this quite dramatic effect. The paper by

> SCHUBERT, GLÄNZEL and BRAUN, "Scientometric datafiles. A Comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields, 1981–1985" published in *Scientometrics*, vol. 16, 1989, pp. 3–478,

has received 138 citations till March 2003. These data have been retrieved from the on-line version of the *Web of Science* (WoS) of the Institute for Scientific Information (ISI - Thomson Scientific, Philadelphia, PA, USA). Among those 138 citations are 113 correct citations, whereas 25 citations were incorrect. The error caused by citing authors amounts to 18.1%. We can detect three particular types of errors; the most frequent was the incorrect or missing page number, followed by an incorrect publication year and the incorrect first author. The head of a research group or institute is sometimes erroneously assumed to be the first authors. All variances of the cited work that occurred in the *Web of Science* database are presented in Table 3.

We have found such errors also in the set of delayed recognition papers. Since several incorrect citation variants were quite frequent, we assume that some authors might have copied them from the reference lists of other publications. Incorrect citations occur over the full citing period till 2001. Therefore, we have not corrected citation counts manually. Nevertheless, we wish to draw the readers' attention to the fact that the true citation rates of the above papers might be somewhat higher than reported in this study.

4 CONCLUSIONS

The statement that bibliometric indicators might lack validity because of necessary technical and methodological restrictions through limited citation windows does only hold for exceptional cases. The sometimes-heard argument that papers have not been cited for a long time and than,

ten or more years later become highly cited may only bear upon exceptions to the rule. This study gives evidence on the validity of citation-based indicators for the overwhelming share of papers published in scientific journals. Beyond this, a new insight is gained into general and field-specific rules of citation processes on the basis of long-term observations. This study has shown again that in theoretical fields, above all, in pure and applied mathematics a citation window larger than three years should be applied. This observation is in line with results by Rousseau (1988) on journals in pure mathematics

The individual analysis of selected delayed reception papers has shown that these publications have presented fundamental results or are related to important discoveries. The high citation rates reflect reception by the community in an adequate manner. More regular citation history of related papers of the same authors show that the analysis of publication sets using standard citation windows would have indicated outstanding citation impact even at the level of individual scientists if a the analysis is based on a longer publication period and/or a set of selected publications by the same author.

REFERENCES

Burrell, Q.L. (2001), Stochastic modelling of the first-citation distribution. *Scientometrics*, 52 (1), 3–12.

Egghe, L., Rousseau, R. (1990), *Introduction to informetrics. Quantitative methods in library, documentation and information science*. Elsevier Science Publisher. Amsterdam.

Garfield, E. (1980), Premature discovery or delayed recognition – Why? *Current Contents*, 21, 5–10.

Glänzel, W., Schubert, A. (1992), Some facts and figures on highly cited papers in the sciences, 1981–1985. *Scientometrics*, 25 (3), 373–380.

Glänzel, W. (1992), On some stopping times on citation processes. From theory to indicators. *Information Processing and Management*, 28 (1), 53–60.

Glänzel, W., Schoepflin, U. (1994), A stochastic model for the ageing analyses of scientific literature. *Scientometrics*, 30 (1), 49–64.

Glänzel, W., Schoepflin, U. (1995), A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science*, 21 (1), 37–53.

Glänzel, W., Schubert, A. (1995), Predictive aspects of a stochastic model for citation processes. *Information Processing & Management*, 31 (1), 69–80.

Glänzel, W., Schubert, A., Braun, T. (2002), A relational charting approach to the world of basic research in twelve science fields at the end of the second millennium. *Scientometrics*, 55, 335–348.

Rousseau, R. (1988), *Citation distribution of pure mathematics journals*. In: L. Egghe & R. Rousseau (Eds), Informetrics 87/88. Elsevier Science Publishers B.V., 249–260.

Rousseau, R. (1994), Double exponential models for first-citation processes. *Scientometrics*, 30 (1), 213–227.

Schubert, A., Glänzel, W. (1986), Mean Response Time. A new indicator of journal citation speed with application to physics journals. *Czechoslovak Journal of Physics*, B 36, 121–125.