

Is Scientific Literature Subject to a ‘Sell-By-Date’? A General Methodology to Analyze the ‘Durability’ of Scientific Documents

Rodrigo Costas, Thed N. van Leeuwen, and Anthony F.J. van Raan

Centre for Science and Technology Studies (CWTS), Leiden University, Wassenaarseweg 62A, P.O. Box 905, 2300 AX Leiden, The Netherlands. E-mail: {rcostas,leeuwen,vanraan}@cwts.leidenuniv.nl

The study of the citation histories and ageing of documents are topics that have been addressed from several perspectives, especially in the analysis of documents with “delayed recognition” or “sleeping beauties.” However, there is no general methodology that can be extensively applied for different time periods or research fields. In this article, a new methodology for the general analysis of the ageing and “durability” of scientific papers is presented. This methodology classifies documents into three general types: delayed documents, which receive the main part of their citations later than normal documents; flashes in the pan, which receive citations immediately after their publication but are not cited in the long term; and normal documents, documents with a typical distribution of citations over time. These three types of durability have been analyzed considering the whole population of documents in the Web of Science with at least 5 external citations (i.e., not considering self-citations). Several patterns related to the three types of durability have been found and the potential for further research of the developed methodology is discussed.

Introduction

One topic of high concern of researchers is how many years they need to wait until their papers can be properly acknowledged and accepted (and therefore cited) by their scientific community. Previous authors have studied the effects of the “scientific prematurity” (Stent, 1972) or “delayed recognition” (Garfield, 1980) related to all cases of papers that have been cited later in comparison with the average papers in their research field.

The effects of the ageing of scientific publications are important for the study of scientific communication (Pollmann, 2000). Policy makers and research managers are often inclined to think that science is primarily a “hectic”

business in which research results will prove their impact soon after publication. This attitude underestimates the crucial importance of durability of research results. Therefore, the study of the ageing and delayed reception of scientific documents is certainly also an important topic for the application of bibliometrics in research assessment, for instance, the correct establishment of citation windows and proper calculation of impact indicators.

The study of ageing and citation histories of documents has been addressed from several perspectives (Aversa, 1985; Glänzel & Schoepflin, 1995; Moed, van Leeuwen, & Reedijk, 1998; Aksnes, 2003), especially the analysis and detection of documents with “delayed recognition” (Garfield, 1980; Glänzel, Schlemmer, & Thijs, 2003) or even “sleeping beauties” (van Raan, 2004). The opposite effect of delayed recognition was also described as “flash in the pan” (van Dalen & Henkens, 2005): documents that are noticed immediately after publication and frequently cited but do not seem to have a lasting impact and die early in life. Both aspects (delayed recognition and flashes in the pan) deal with the more general ideas on durability and obsolescence of documents, as well as with the ageing of scientific literature, which is strongly related to the impact of documents over time.

Earlier studies have focused mainly on the ageing and citation patterns of highly cited papers (Cano & Lind, 1991; Aversa, 1985; Aksnes, 2003; Levitt & Thelwall, 2008, 2009) trying to establish their main determinants and properties. However, until now there is no proper methodology for a global durability analysis of documents regardless degree of citedness, year of publication, or research field. Previous attempts were based on fixed years of publication and the thresholds for the classification of documents were based on the selection of the authors (Aksnes, 2003; Moed et al., 1998). Therefore, the proposal of an integral methodology with classification of documents according to their durability regardless of publication year or total number of citations is necessary, as well as a general “technical” definition of different types of durability of documents.

Received July 8, 2009; revised September 3, 2009; accepted September 3, 2009

© 2009 ASIS&T • Published online 17 November 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21244

TABLE 1. Example citation history document A.

Document	Year of publication	Total external citations	2002	2003	2004	2005	2006	2007	2008
A	2002	25	0	3	6	7	4	3	2

In this article, we propose a methodology to classify documents according to the “durability” of their citations. Here, durability is understood as the different patterns of the citation history of documents. The importance of such a methodology for the identification of papers deviating from the typical citation patterns was already suggested by Garfield (1980), but until now, there is no general methodology that can be used for separate years, disciplines, and document types.

Objectives

Our main objective is to develop a general methodology for the classification of research publications according to the durability of their citations. We aim at a classification of all documents regardless of the year of publication or degree of citations (i.e., taking into account all and not only highly cited documents). This methodology considers the different Web of Science (WoS)¹ Journal Subject Categories (JSC)² (fields) where documents are classified.

After the development of the methodology, we analyze the main characteristics of documents according to their durability. The analysis focuses on a general picture of the different types of durability and how citations histories of documents evolve over time. This provides us with a general description and, particularly, citation patterns for the different types of durability that can be used in the analysis of different fields as well as of different citation windows.

Methodology

Development of a General Methodology for the Study of Durability

Our methodology for the classification of the durability of scientific papers aims at the classification of documents according to their citation histories in three general types:

- Normal: documents with the typical distribution in their citations over time, i.e., reaching a maximum in, say, 3 to 4 years after publication and then followed by an exponential decay.
- Flash in the pan: documents that tend to receive citations immediately after their publication but are not cited in the longer term.

¹Thomson Reuters is the producer and publisher of the Web of Science (WoS) that covers the (former) Science Citation Index (-extended), the Social Science Citation Index, and the Arts & Humanities Citation Index. Throughout this paper we use the term ‘WoS’ for the above set of databases.

²We use the definition of fields based on a classification of scientific journals into *categories* developed by Thomson Reuters. Although this classification is not perfect, it provides a clear and ‘fixed’ consistent field definition suitable for automated procedures within our data system.

- Delayed: documents that receive the main part of their citations later than normal documents; sleeping beauties are included within this type.

This classification is similar to that used by van Dalen and Henkens (2005); however, they called the third type sleeping beauties. We prefer the term “delayed” because sleeping beauties are a rare and rather extreme phenomenon in science (van Raan, 2004). Other authors also observed similar classifications. For example, Aversa (1985) found two general patterns: “delayed rise – slow decline,” which coincides with our delayed type, and “early rise – rapid decline,” which is basically our flash in the pan type. Aksnes (2003) also considered a third type, “medium rise – slow decline,” which can be considered as strongly related to our normal-type documents.

This article is based on the data of documents covered by the WoS, including more than 8,700 journals and covering all scientific disciplines (Costas & Iribarren-Maestro, 2007), thus being one of the most important databases for bibliometric studies and research assessment purposes. We have also focused on “external citations,” i.e., citations received by documents after the exclusion of self-citations.³ We take the position that this type of citations represents the real impact and transfer of knowledge of documents beyond their original producers (Costas, van Leeuwen, & Bordons, 2008). Nevertheless, future research on the evolution and effects of self-citations on durability types will be of high importance.

Our approach, which is based on the distribution of the percentage of citations that documents receive each year (citation history), comprises the following steps:

1. For each document in the WoS, its *citation history* over time is analyzed (self-citations excluded). This provides us with the evolution of citations since the year of publication until the last year considered in the analysis; see example in Table 1.

According to Table 1, document A was published in 2002, received a total of 25 external citations until 2008, and the evolution of the citations is shown in the following years.

2. For all documents with at least one citation, the percentage of citations received each year has been calculated on the basis of their citation history and the cumulative value of the percentage of citations; see example in Table 2.

³A citation is a self-citation if any of the authors of the citing paper is also an author of the cited paper. In other words, self-citations are calculated at the document level, where source documents are matched to their citing papers and those matches that at least share one author are considered self-citations. Thus, self-citations are citations given by any of the authors of the source paper (regardless the position of the author in the papers).

TABLE 2. Example percentage of citations and cumulative values of document A.

Document	Year of publication	Total external citations	2002	2003	2004	2005	2006	2007	2008
A	2002	25	0	3	6	7	4	3	2
% citations per year		100	0.00	12.00	24.00	28.00	16.00	12.00	8.00
Cumulative %		100	0.00	12.00	36.00	64.00	80.00	92.00	100.00

TABLE 3. Example of document A and the P25 and P75 of its field.

Document	Year of publication	Year 50%	P25-JSC	P75-JSC
A	2002	4	3	6

Aksnes (2003) used a similar approach: calculating the yearly percentage of citations.

In Table 2, we see the evolution of the percentage of citations for document A and its cumulative value.

3. For each document, we identified the year after publication in which the document received for the first time *at least* 50% of its citations (“Year 50%”) during the time period up to and including 2008. This provides us with an ordinal number for the document (e.g., for documents published in 1994, the number 1 is given when documents received at least 50% of their citations in 1994, 2 when 50% of citations is reached in 1995, and 3 if this happens in 1996). In our example for document A, this year is 2005 (shaded in Table 2), which means that in our example, the value for “Year 50%” of document A is 4.
4. For the whole population of documents (with at least one external citation), taking into account all document types and considering the different research fields (i.e., WoS JSC), we calculated, for all documents *of the same year of publication*, the percentiles 25 and 75 of the distribution function of the value of the new indicator “Year 50%.” As noticed by Glänzel & Schoepflin (1994), the ageing behavior of documents is influenced by field characteristics, and we, therefore, consider the scientific field as the best reference for the classification of each document. Taking again our example, document A is classified in one field (JSC) with P25 and P75 scores as given in Table 3. According to Table 3, document A is classified in a field where the P25 and 75 of the “Year 50%” of all documents published in 2002 is 3 and 6, respectively.
5. The general criterion for the classification of documents is now as follows:
 - a. Flashes in the pan: the value for “Year 50%” is smaller than the P25 value for that field ($<P25$).
 - b. Delayed documents: the value for “Year 50%” is larger than the P75 value for that field ($>P75$).
 - c. Normal documents: the value for “Year 50%” is between the P25 and the P75 value for that field ($\geq P25$ and $\leq P75$).

This criterion is based on a personal communication of Derek de Solla Price to Aversa (1985), in which he claimed that “papers exhibit three basic citation patterns,” distributed among the following: “25 per cent of papers cited at a constant rate without declining” (roughly, the equivalent of our delayed type); “25 per cent of gradually increased in citedness and then declined at a similar rate” (similar to our flash in the pan); and finally “50 per cent

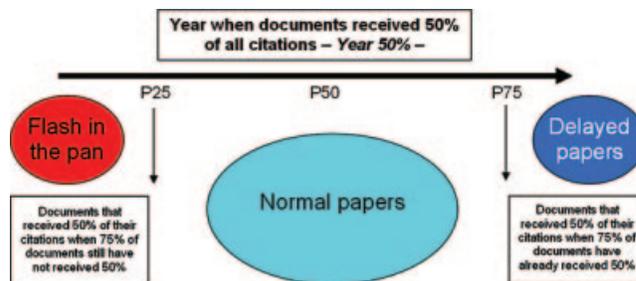


FIG. 1. General scheme of classification of papers by durability type.

cited at a constant rate for several years” (the equivalent of our normal-type). In Figure 1, a general scheme of the criteria proposed is shown.

According to the above-mentioned criteria, flashes in the pan can then be defined as those documents that have received 50% of their citations when the 75% of other documents still have not received 50% of their citations. Normal documents are all documents that receive the 50% of their citation around the year of P50 (between P25 and P75). Finally, delayed documents are those papers that have received 50% of their citations after P75 years in their fields. Notice that the criteria of step 5 tend to favor normal documents, because we assumed that this pattern should be preferred instead of the other two when they are equal to the thresholds (P25 and P75). Thus, in a hypothetical situation of all documents gathering 50% of their citations in the same number of years after publication, P25 and P75 are equal to P50, and according to our classification, all documents would be classified as normal.

This methodology is originally based on the JSC distribution of documents according to the classification of journals in the WoS (more specifically, Thomson Reuters’ Journal Citation Reports). However, journals can be classified into more than one field (JSC). Consequently, documents in such journals will be assigned to more than one field and, therefore, to more than just one durability type: a document can be in one of the normal document categories and in a delayed document category. To avoid this “document schizophrenia” and to characterize each document with just one single type, the following steps are proposed:

6. For documents in more than one field (i.e., more than one JSC) and with more than one “durability class,” different class values are assigned to different types:
 - a. Flash in the pan → 1
 - b. Normal → 2
 - c. Delayed → 3

TABLE 4. Document A in more than one field.

Document	Year of publication	Year 50%	JSC	P25	P75	Durability type	Class value
A	2002	4	Entomology	3	6	Normal	2
			Biochemistry	2	3	Delayed	3

TABLE 5. Final durability type for document A.

	Year of publication	Year 50%	Tot. JSC	Total sum class values	Average class value	Final durability type
A	2002	4	2	5	2.5	Delayed

TABLE 6. General values for documents in different durability types.

Durability	Total publications	%	Authors/doc.	Addresses/doc.	Countries/doc.	Keyword/doc.	Pages/doc.	References/doc.
Delayed	1,665,712	19.97	3.37 ± 3.41	1.99 ± 1.36	1.14 ± 0.44	4.82 ± 1.80	9.22 ± 37.05	28.56 ± 28.82
Flash in the pan	814,098	9.76	3.74 ± 9.57	2.07 ± 1.77	1.16 ± 0.54	4.86 ± 1.79	8.11 ± 39.69	29.31 ± 30.09
Normal	5,860,703	70.27	3.75 ± 6.48	2.25 ± 1.66	1.18 ± 0.53	4.91 ± 1.82	9.29 ± 42.66	32.96 ± 33.18

Note. "Keywords/doc." refers to the number of author-given keywords per document. Values are indicated by the mean value ± standard deviation.

7. The number of different fields (JSC) of each document is also calculated and the total durability class values are summed. In Table 4, the example of document A is presented. This document is classified in two fields (entomology and biochemistry) with different P25 and P75 and different durability types.

We see that document A, assigned to two different fields, has two different durability types in these fields and, thus, one different class value for each field.

8. Finally, the total durability class values of each document are divided by the number of fields. Thus, all documents get values between 1 and 3. The final criteria for classification is as follows:

- a. Flash in the pan → ≤ 1.5
- b. Normal → > 1.5 and < 2.5
- c. Delayed → ≤ 2.5

This second classification is based on the idea that the most frequent durability type among the different fields of papers is the most important. This means that in this second classification, implicitly deviant types (i.e., flash in the pan and delayed) prevail when they appear with the same frequency as the normal type, but they are mutually balanced when they appear together in the same document, in favor of the normal type. In other words, if a paper is assigned to two fields and is a delayed in the first field and a flash in the pan in the other, then it is finally considered normal. Further analyses and refinements of this aspect of the methodology will be developed in the next stage of this study.

As a final result of this methodology, all documents are classified in just one final durability type: in our example in Table 5, document A gets a final classification as delayed, which makes the analysis of all documents more straightforward. Thus, with our methodology, we are able to classify all documents regardless of their degree of citedness (even documents with just one external citation can be classified

depending on the year they receive their citation), taking into account both their year of publication and their (if this is the case) different fields.

Results

The above discussed methodology has been applied to all documents in the WoS database published between 1980 and 2008 (in total 30,445,406 documents), including all document types, languages, years of publication, etc. Essential elements of the analysis to identify the durability type of each publication are year of publication, citation trend up to and including 2008, and field (JSC). In this section, we discuss our results concerning the main characteristics of the three types of durability.

First, we present a general description of the three types of durability and their properties. This analysis provides an impression of the documents behind the three types. Although all documents with at least one external citation were identified in the WoS, we considered in our analysis only those publications with a minimum of five external citations and published between 1980 and 2003 (8,340,513 documents) to avoid the influence of hardly cited documents and documents published at the very end of the period. Thus, we have "source publications," which are the cited publications for the period 1980–2003, and for the citing publications to count citations, we use the period 1980–2008.

General Properties of Documents According to Their Durability

In Table 6, we present a general description of documents, considering different bibliometric properties of documents classified in the three durability classes.

TABLE 7. Distribution of document types among durability types.

Durability	Document type	Total documents	%
Delayed	Articles	1,496,444	89.84
	Notes	74,104	4.45
	Reviews	52,341	3.14
	Letters	23,707	1.42
	Editorials	10,950	0.66
	Others	8,166	0.49
Flash in the pan	Articles	677,461	83.22
	Notes	39,035	4.79
	Reviews	29,713	3.65
	Letters	20,454	2.51
	Editorials	28,708	3.53
	Others	18,727	2.30
Normal	Articles	5,202,724	88.77
	Notes	205,920	3.51
	Reviews	256,280	4.37
	Letters	84,354	1.44
	Editorials	74,063	1.26
	Others	37,362	0.64

Statistically significant differences ($p < 0.000$) have been observed among all durability types in all indicators (test *U* Mann–Whitney). We also observed significant differences for the direct comparison of delayed and flash in the pan ($p < 0.000$). A further important interesting observation is that the distribution of documents among the three types does not come up to a 25-50-25 distribution. The reason for this is that we discarded very lowly cited documents and that the “threshold equals,” discussed in step 5, tends to favor normal documents. According to Table 6, delayed documents have a slightly lower number of authors per document and institutes and countries per document, and they also have fewer author-keywords and references per document.

On the other hand, comparing flashes in the pan with delayed documents, we observe that the first are generally documents with more authors, more countries, and more addresses per document than the delayed documents. They also have more references per document than delayed. It is important to stress that flashes in the pan are, on average, the shortest documents (i.e., less pages per document). The distribution of all three durability types by document type was also analyzed from a general perspective; see Table 7 and Figure 2.

We immediately observe that flashes in the pan are published proportionally more as notes, letters, editorials, etc., than the other two types. On the other hand, delayed documents are more represented in the articles document type. These results already provide a first explanation of why flashes in the pan are generally shorter documents, because this type contains proportionally more notes, letters, and editorials than the other two. In this line, a closer look at two letter journals (which include mainly rapid communications; Moed et al., 1998), such as *Science* and *Nature*, shows that these journals cover 10.9% and 10.5% of flashes in the pan, respectively. This means that they publish proportionally more papers of this type than the observed average for the whole database (9.8%). These results corroborate previous results by Glänzel & Schoepflin (1994, 1995) in that

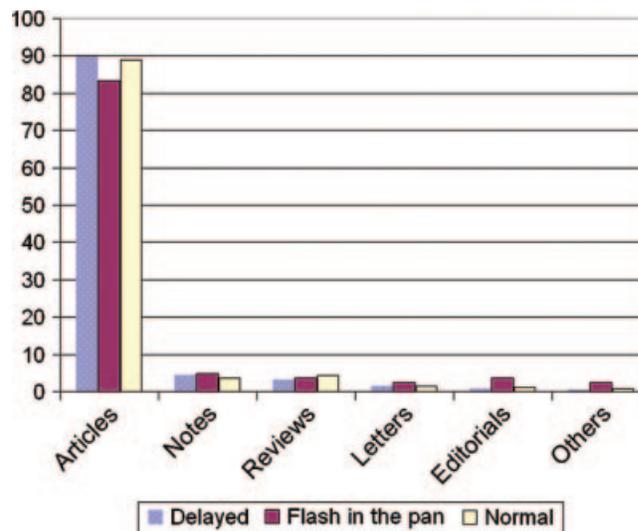


FIG. 2. Durability classes by document types.

document and journal types influence the ageing behavior of documents. This implies that letters, short communications, and occasionally short review articles represent a much faster communication, in other words, a higher degree of flashes in the pan.

The performance of the three types of durability was also studied on the basis of the Centrum voor Wetenschap en Technologische Studies (Center for Science and Technology Studies) standard indicators⁴ (Moed, De Bruin, & van Leeuwen, 1995). For the analysis presented in this article, only articles, notes, letters, and reviews were considered (Table 8). A variable citation window was used, which means that citations are counted from the year of publication of documents up to and including 2008.

Standard bibliometric indicators are as follows:

- Number of publications *P* in WoS-covered journals.
- Number of citations *C* received by *P* during the specified period, without self-citations.
- Average number of citations per publication, without self-citations (CPP).
- Journal-based worldwide average impact as an international reference level (JCS, journal citation score), without self-citations; as many journals are involved in the different classes, we use the average JCS_m; for the calculation of JCS_m, the same publication and citation counting procedure, time windows, and article types are used as in the case of CPP.
- Field-based, worldwide average impact as an international reference level (FCS, field citation score), without self-citations; as many fields are involved, we use the average FCS_m; for the calculation of FCS_m, the same publication and citation counting procedure, time windows, and article types are used as in the case of CPP; we refer in this article to the FCS_m indicator as the “field citation density.”

⁴We stress again that self-citations have been removed from the publications used in this analysis and also from all publications used as an international reference. Thus, only ‘external citations’ (i.e., citations given by authors different from the co authors of the original paper) have been considered for the calculation of all indicators.

TABLE 8. Standard indicators by durability type.

Durability	Period citations	P	C	CPP(*)	FCSm	JCSm	%sc	CPP/FCSm	CPP/JCSm	JCSm/FCSm
Delayed	1980–08	1,646,596	56,258,641	34.17	17.71	21.42	0.12	1.93	1.60	1.21
Flash in the pan	1980–08	766,663	10,892,718	14.21	18.49	22.31	0.19	0.77	0.64	1.21
Normal	1980–08	5,749,278	169,749,327	29.53	17.81	24.42	0.14	1.66	1.21	1.37

Note. Period 1980–2003 for documents and 1980-2008 for citations. For an explanation of the indicator symbols, see the text box below this table.

(*) The total number of external citations of documents in each durability type have also been analyzed through the Mann-Whitney test which indicated significant differences among the three types of durability ($p < 0.000$).

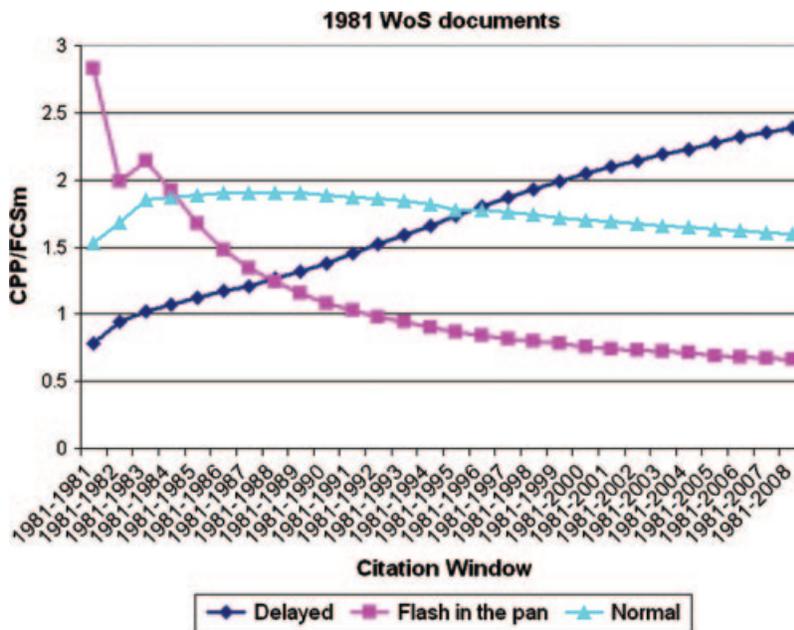


FIG. 3. Evolution of CPP/FCSm, with increasing citation window.

- The percentage of self-citations %sc in total number of citations.
- Comparison of the CPP with the worldwide average based on JCSm as a standard, without self-citations, indicator CPP/JCSm.
- Comparison of the CPP with the worldwide average based on FCSm as a standard, without self-citations, indicator CPP/FCSm; this is our “crown indicator.”
- Ratio JCSm/FCSm is the relative, field-normalized journal impact indicator.

According to Table 8, delayed documents present the highest average impact in citations per paper (CCP) compared with the other types of documents. However, they are published in fields (FCSm) and journals (JCSm) with the lowest level of citation density as compared with the other durability types. Also, for the (long life) field-specific impact CPP/FCSm, they show the highest level, as well as the highest CPP/JCSm. Another interesting finding is that delayed documents relate to the lowest percentage of self-citations. This finding is in agreement with the general observation that the share of self-citations is decreasing as a function of time; in other words, the longer after publication we measure the total number of citations, the smaller the fraction of self-citations.

Flashes in the pan show the lowest CPP, but these documents are normally published in fields with high citation

density (FCSm), although their journals are not cited at the same level (intermediate JCSm). They present the lowest CPP/FCSm and CPP/JCSm and the highest percentage of self-citations.

Effects of the Durability on Citation Windows

We also analyzed the effect of different citation windows for calculation of indicators. First, the evolution of the CPP/FCSm of the three different durability types has been analyzed with variable citation windows. For a better understanding of the results, documents published only in 1981 were considered in this analysis; see Figure 3.

Figure 3 presents the evolution of CPP/FCSm for documents published in 1981, with a yearly increasing citation window. The graphs clearly show that in case of a lengthening of the citation window, the CPP/FCSm values of the different durability types follow quite different patterns. More specifically, flashes in the pan have the highest scores with the shortest citation window, but they display a decreasing pattern as the citation window becomes longer. Delayed documents have the opposite pattern; an increasing citation window implies that their CPP/FCSm also increases over time, and this trend holds during the entire period. Finally, normal documents show a quite stable pattern, increasing

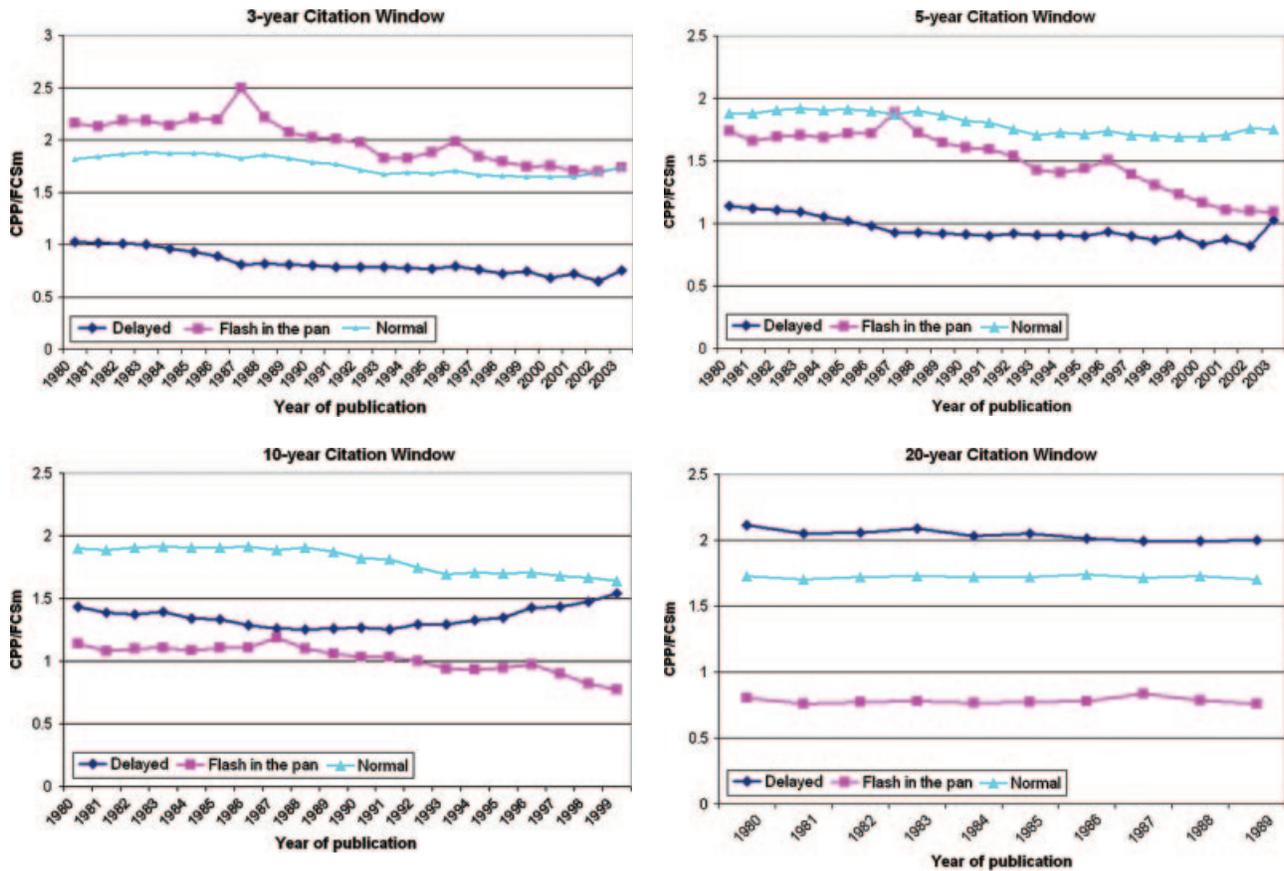


FIG. 4. Evolution of CPP/FCSm with different fixed citation windows.

during 3 to 4 years after publication and slightly decreasing (but quite stable) afterwards.

According to these findings, it is clear that the selected citation window has an effect on the final results. It is also evident that this first indication of differences among the three types of durability as a consequence of changing citation patterns is very relevant for the use of bibliometric indicators in research assessment procedures. To test this effect, different *fixed* citation windows (3, 5, 10, and 20 years) have been considered for the documents in the three durability types. In Figure 4, the evolution of the CPP/FCSm with different fixed citation windows is presented for the three types of durability.

With the shortest citation window, 3 years, flashes in the pan have the highest CPP/FCSm scores and delayed documents the lowest CPP/FCSm scores (even below 1!). When we increase the length of the citation window up to 5 years, normal documents are the ones with the highest scores, but flashes in the pan still present higher scores in comparison with delayed documents. If, however, the citation window is increased up to 10 years, normal documents are still the ones with the highest scores but now delayed documents outperform flashes in the pan. Finally, with the longest citation window (20 years), delayed documents show the highest CPP/FCSm (mainly above 2!), normal documents are second in the rank, and, finally, flashes in the pan have the lowest scores. Normal documents exhibit a very similar level of impact in the four figures regardless the citation window used,

being always above 1.5 and below 2. This may seem high, but remember that we discarded all papers with less than 5 external citations, and these papers count for about 50% of all papers with at least 1 external citation. If we would include all papers and the not-cited papers, then the CPP/FCSm value for this whole set would be close to 1, as expected.

Evolution of CPP in Relation to Durability

The impact of documents in relation to their different types of durability has been analyzed. The average “simple” impact (i.e., impact not normalized by field, such as in case of the CPP/FCSm; CPP values) as well as the percentage of the citations received on a yearly basis, has been studied; see Figure 5.

According to Figure 5, the evolution of the CPP presents characteristic trends. All three types of durability have an increasing pattern at the beginning and after they have reached a peak they start to decline. However, the shapes of these CPP patterns are quite different for the three types of durability. Flashes in the pan have a very high impact (CPP) during the immediate years after publication, reaching their maximum peak after 3 years and then showing a rapid decline. In contrast, delayed documents are not very highly cited at the beginning, but their impact increases steadily during several years, reaching a peak after 10–11 years after publication and showing a slow decrease during the rest of

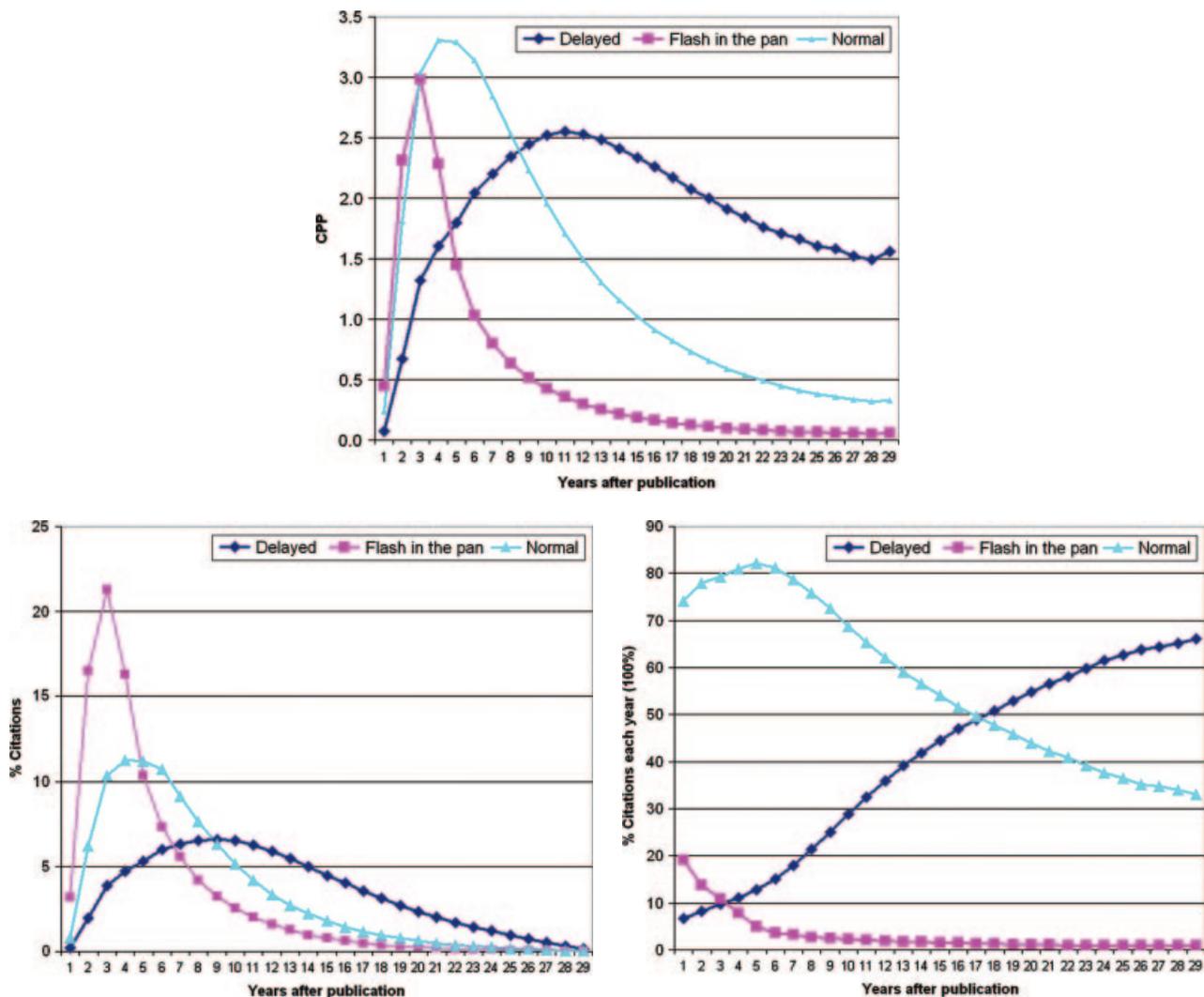


FIG. 5. Evolution of CPP documents in relation to year of publication.

the period. It is also interesting to stress that it is not until the 5th year that delayed documents surpass flashes in the pan in CPP scores.

Normal documents present a pattern of increasing impact during the first years, reaching a peak in the 4th to 5th year after publication, followed by a decrease faster than delayed documents but slower than flashes in the pan. As in the left bottom-corner of Figure 5 shows, flashes in the pan accumulate the highest percentage of their citations during the first years following publication, while delayed documents have a more even distribution of the percentage of their citations (always below 10%).

Finally, in the right bottom-corner of Figure 5, we see how the main part of the citations during the first years comes predominantly from normal documents, but these documents show a decreasing trend after 5–6 years following publication. On the other hand, the percentage of citations represented by delayed documents increases during the whole period.

Another interesting aspect is a possible country dependence of the different durability types. Thus, we analyzed the

TABLE 9. Presence of U.S. addresses in the documents by durability.

Final durability	Total WoS	Total USA	%USA	%Non USA
Flash in the pan	814,098	364,856	44.8	55.2
Normal	5,860,703	2,640,768	45.1	54.9
Delayed	1,665,712	692,852	41.6	58.4
Total	8,340,513	3,698,476	44.3	55.7

presence of addresses from the United States in documents from the perspective of their durability and show the results in Table 9.

We find that delayed documents have a somewhat lower presence of U.S. institutions because 58.4% of delayed documents have no U.S. participants (for the total this is 55.7%), while flashes in the pan have a higher U.S. participation as compared with delayed documents (44.8% vs 41.6%). A possible explanation is that the United States is generally considered to be more “central” in science and, therefore, more rapidly noticed and accepted by the scientific community, while ideas from other countries may need more time to be accepted and acknowledged.

Discussion

In this article, we present a general methodology for the bibliometric description and classification of documents, describing the phenomenon of durability as measured by their citations. This methodology, and especially the technical definition of durability types, implies an important step in the analysis of the ageing and durability of documents because large sets of documents can be studied (in fact, all documents contained in the WoS were classified). It also has policy implications with respect to the measurement of research performance. Moreover, the delimitation of the durability types is more flexible and general than in earlier studies on more extreme cases such as sleeping beauties or delayed recognition. Furthermore, the possibility of studying not only highly cited documents but also moderately cited documents implies an important improvement in the analysis of the citation histories of documents, an issue that was suggested before by Aversa (1985).

The methodology discussed in this article presents several advantages:

- It can be applied to *large sets* of documents;
- it can also be applied to documents published in *different years* (although we consider 5 years of citation history as a minimum threshold for reliable results);
- it takes into account that documents can be classified in *more than one field*; and
- it can also be *updated yearly or monthly* and, thus, improved as more information is available.

Another important advantage of the methodology is its simplicity as only three general types of durability are considered, compared with the broader classifications provided by other authors (Avramescu, 1979; Moed et al., 1998). Our approach of three types of durability is also supported by the results of Aksnes (2003), who proposed an initial classification for highly cited papers of nine different types based on the rise and decline of citations of documents. However, according to the results of this author, only three main types account for the larger majority of the papers: the “early rise – rapid decline” (roughly, our flash in the pan); the “medium rise – slow decline” (similar to our normal type); and the “delayed rise – no decline” (very close to our delayed type).

The analysis of documents classified by our methodology shows the following main results. By comparing delayed documents with flashes in the pan, we find that the first receive more citations (CPP, which is equal to C for an individual document) and that they have a higher field-specific impact (CPP/FCSm). This is in agreement with earlier more general results by Aversa (1985), Aksnes (2003), and Levitt and Thelwall (2009), who showed that slower ageing rates and lateness in citations are correlated with higher citation counts.

With respect to the bibliographic properties of the three types of durability, our first observations necessitate further analysis in follow-up work. To start with, delayed papers are characterized by, on average, less authors, centers, and

countries per document than flash in the pans and normal documents. These results suggest that delayed documents are papers with *lower* levels of collaboration. This is an interesting finding because it could imply that less collaborative papers, which are generally considered as papers with lower impact (Glänzel & Schubert, 2001; Persson, Glänzel, & Danell, 2004), do not appear to be necessarily of less quality in the long run. In other words, it could be hypothesized that less collaborative papers could also gain high impact, but they need more time to become fully recognized and cited accordingly.

A complementary explanation to this idea is that papers in collaboration and especially in international collaboration have larger audiences (van Raan, 1998), thus benefiting from more direct visibility and impact. Also in this line of arguing, papers in collaboration improve their potential visibility by using the network contacts of their collaborators (Katz & Martin, 1997), and, thus, they can diffuse their findings wider and faster among larger audiences. According to this, less collaborative but still high-quality papers (such as delayed documents) may need more time to be distributed, known, and acknowledged by their peers. Our finding can also support this idea that delayed documents also have fewer keywords and references per document. This means that their retrieval through bibliographic databases is more difficult (there are less retrieval elements as compared with flash in the pan and normal documents).

On the other side, flashes in the pan are relatively more frequent in the “shorter” document types (notes, letters, editorials, etc.). This explains why they have the smallest number of pages. This is also linked to the idea that these documents represent an immediacy value rather than an archival value, as they are published in journals that try to present research results as quickly as possible. This is in agreement with Moed et al. (1998) who observed that journals containing rapid communications show a fast decline and “mature” patterns, particularly journals such as *Nature* and *Science*. In the same line, van Dalen & Henkens (2005) found that flashes in the pan are related with high-reputation journals, considering that top journals are the ones where the broader academic debate takes place, while other more specialized journals mostly do not publish work with a broad impact. In any case, these results also point out again the weakness of the journal impact factor as a measure of scientific impact, because it favors journals publishing more flash in the pan than delayed documents (even within the same field). This supports the criticism of Seglen (1997), who found that high impact factor journals are likely covering areas of basic research with a rapidly expanding but short-lived literature that use many references per article.

It is also interesting to notice that flashes in the pan are documents published mainly in fields with high citation density (relatively high *FCSm*), but their overall impact in the long run is much lower than for delayed documents. Van Dalen & Henkens (2005) also found that flashes in the pan are less-cited documents. With the observation that flashes in the pan have more self-citations than delayed documents, and taking

into account that self-citations come mainly during the first years after publication (Glänzel, Thijs, & Schlemmer, 2004; Schubert, Glänzel, & Thijs, 2006), we suggest that flashes in the pan are boosted and advertised by their own authors in the years following publication (Medoff, 2006; Fowler & Aksnes, 2007), and thus gaining immediate attention by other researchers, but losing interest in the long term. In other words, self-citations can boost significantly the impact of documents, because they mainly influence the impact during the first years after publication, which are crucially important years for flash in the pan but not for the more durable papers (i.e., normal and delayed documents). Nevertheless, these arguments are based on general self-citation characteristics, as if “technical” arguments can explain the entire impact history of papers. It is, however, very well possible that delayed documents represent research work of another kind than the work published in flash in the pans and in normal papers, namely, related to the idea of *sleeping beauties*, work that is ahead of time, so that conceptual and intellectual characteristics provide an explanation of the impact history. In this line, McCain and Turner (1989), following the classification of Aversa (1985) and studying papers on molecular genetics, found that articles presenting research methods or theoretical contributions fell into the latter-peaking pattern (our delayed type), while those cited for specific results fell into the early-peaking pattern (our flash in the pans). Future research on the content and contextual properties of the three durability types will shed more light on the citation patterns of documents regarding their intellectual and conceptual approaches.

Considering the evolution of citations over time for the three types of durability, we observe how delayed documents present a rise and decline in their impact which is much slower than for the other two types, i.e., the delayed documents are the more frequently cited documents 10 years after publication. In contrast, flashes in the pan show a very strong growth in impact immediately after publication and then they rapidly decrease after reaching a peak. One possible explanation for this fast declining pattern of flashes in the pan is that they are often ‘followed up’ by normal papers, which will ‘take over’ and present the same results but more in detail and more thoroughly treated. Thus, the impact shifts from the flash in the pan to the normal paper, leaving the role of the flash in the pan as ‘early warnings’ in order to get as quick as possible the credits for the results. This explanation reinforces the idea of sets of publications being in fact just ‘one oeuvre’. We will focus further research on the identification of such links between of strongly related papers.

With regard to the different citation windows, we find that after 20 years delayed documents show the highest impact in their fields (surpassing both normal and flashes in the pan). This finding implies that very short citation windows mainly favor flash in the pan and, thus, we were not surprised by the fact that flash in the pans are published in journals with the highest impact factors—, while the opposite is found for normal and delayed documents. According to our results, it is clear that establishing the percentage of delayed documents is a crucial topic to test if indicators based on shorter

citations windows can affect those units of analysis, such as research groups and institutes, with higher degrees of delayed papers.

Finally, the methodology presented in this article raises many questions about the ageing of documents and how documents can be considered for bibliometric analysis. Several intriguing lines of future research are suggested:

- Better understanding of the three general types of durability, *improving their characterization*, and providing more explanations for the general patterns found in this article, as well as the consideration of other aspects of interdisciplinarity, self-citations, etc.
- Study of *different scientific fields* to detect main differences among them; also deviations from the general pattern according to durability are an important issue of future research.
- Analysis of the *effects* of these durability types for the calculation of general *bibliometric indicators* is very important as well as the establishment of more appropriate citation windows.
- Importance of the durability types for the study of *different units of analysis* and aggregation levels is highly important to find out whether the impact assessment of groups, institutes, or even of individual researchers can be affected by these types of documents. In this line, important questions are how these durability types are distributed among different institutes and who are the researchers behind these types; these issues are directly related to a better understanding of the ageing of scientific literature.
- Identification of strong links among papers with different classifications, in particular, as discussed above, links between flash in the pan and normal papers.
- Finally, the *prediction of impact* is an important topic of research nowadays (Geller, de Cani, & Davies, 1986; Feitelson & Yovel, 2004; Mingers & Burrell, 2006; Castillo, Donato, & Gionis, 2006).

These issues illustrate the potential for further research of the developed methodology.

References

- Aksnes, D.W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159–170.
- Aversa, E.S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics*, 7(3–6), 383–389.
- Avramescu, A. (1979). Actuality and obsolescence of scientific literature. *Journal of the American Society for Information Science*, 30(5), 296–303.
- Cano, V., & Lind, N.C. (1991). Citation life cycles of ten citation classics. *Scientometrics*, 22(2), 297–312.
- Castillo, C., Donato, D., & Gionis, A. (2007). Estimating number of citations using author reputation. In N. Ziviani & R. Baeza-Yates (Eds.), *String Processing and Information Retrieval (SPIRE)*, LNCS 4726, pp. 107–117.
- Costas, R., & Iribarren-Maestro, I. (2007). Variations in content and format of ISI databases in their different versions: The case of Science Citation Index in CD-ROM and the Web of Science. *Scientometrics*, 72(2), 167–183.
- Costas, R., van Leeuwen, Th.N., & Bordons, M. (2008). Considering self-citations: Different methodologies of calculation and their effect at micro-level. 10th International Conference on Science and Technology Indicators. Vienna (Austria): Austrian Research Centres GmbH.

- Feitelson, D.G., & Yovel, U. (2004). Predictive ranking of computer scientists using CiteSeer data. *Journal of Documentation*, 60(1), 44–61.
- Fowler, J.H., & Aksnes, D.W. (2007). Does self-citation pay? *Scientometrics*, 72(3), 427–437.
- Garfield, E. (1980). Premature discovery or delayed recognition—Why? *Essays of an Information Scientist*, 4, 488–493.
- Geller, N.L., de Cani, J.S., & Davies, R.E. (1981). Lifetime-Citation rates: A mathematical model to compare scientists' work. *Journal of the American Society for Information Science*, 32(1), 1–15.
- Glänzel, W., Schlemmer, B., & Thijs, B. (2003). Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon. *Scientometrics*, 58(3), 571–586.
- Glänzel, W., & Schoepflin, U. (1994). A stochastic model for the ageing of scientific literature. *Scientometrics*, 30(1), 49–64.
- Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception process of scientific literature. *Journal of Information Science*, 21(1), 37–53.
- Glänzel, W., & Schubert, A. (2001). Double effort = Double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50(2), 199–214.
- Glänzel, W., Thijs, B., & Schlemmer, B. (2004). A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics*, 59(3), 281–310.
- Katz, J.S., & Martin, B.R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.
- Levitt, J.M., & Thelwall, M. (2008). Patterns of annual citations of highly cited articles and the prediction of their citation raking: A comparison across subject categories. *Scientometrics*, 77(1), 41–60.
- Levitt, J.M., & Thelwall, M. (2009). The most highly cited Library and Information Science articles: Interdisciplinarity, first authors and citation patterns. *Scientometrics*, 78(1), 45–67.
- McCain, K.W., & Turner, K. (1989). Citation content analysis and aging patterns of journal articles in molecular genetics. *Scientometrics*, 17(1–2), 127–163.
- Medoff, M.H. (2006). The efficiency of self-citations in economics. *Scientometrics*, 69(1), 69–84.
- Mingers, J., & Burrell, Q.L. (2006). Modeling citation behavior in Management Science journals. *Information Processing & Management*, 42, 1451–1464.
- Moed, H.F., De Bruin, R.E., & van Leeuwen, T.N. (1995). New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics*, 33(3), 381–422.
- Moed, H.F., van Leeuwen, T.N., & Reedijk, J. (1998). A new classification system to describe the ageing of scientific journals and their impact factors. *Journal of Documentation*, 54(4), 387–419.
- Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.
- Pollmann, T. (2000). Forgetting and the ageing of scientific publications. *Scientometrics*, 47(1), 43–54.
- Schubert, A., Glänzel, W., & Thijs, B. (2006). The weight of author self-citations. A fractional approach to self-citation counting. *Scientometrics*, 67(3), 503–514.
- Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 498–502.
- Stent, G.S. (1972). Prematurity and uniqueness in scientific discovery. *Scientific American*, 227(6), 84–93.
- van Dalen, H.P., & Henkens, K. (2005). Signals in science—On the importance of signaling in gaining attention in Science. *Scientometrics*, 64(2), 209–233.
- van Raan, A.F.J. (1998). The influence of international collaboration on the impact of research results: Some simple mathematical considerations concerning the role of self-citations. *Scientometrics*, 42(3), 423–428.
- van Raan, A.F.J. (2004). Sleeping beauties in science. *Scientometrics*, 59(3), 467–472.