Wolfgang Glänzel[*,], Koenraad Debackere[*], Bart Thijs[*], András Schubert[**]**

[*] **SOOS, KU Leuven, Belgium**
[**] **IRPS, HAS, Budapest, Hungary**

A concise review on the role of author self-citations in information science, bibliometrics and science policy

Abstract

The objective of the present study is twofold: (1) to show the aims and means of quantitative interpretation of bibliographic features in bibliometrics and their re-interpretation in research policy, and (2) to summarise the state-of-art in self-citation research. The authors describe three approaches to the role of author self-citations and possible conflicts arising from the different perspectives.

Résumé

L'objectif de cette étude est de caractériser les effets de quantification comme appliqués dans la bibliométrie et leur interprétation en politique de recherche. Nous présentons un sommaire de la recherche en ce qui concerne le phénomène des auto-citations et nous décrivons trois possibilités d'expliquer les auto-citations aussi bien que les conflits qui résultent de ces points de vues.

1. Introduction

In research evaluation, citation indicators are among the most important impact measures of scientific literature. The application of bibliometric citation data, however, has always been somewhat controversial, sometimes even heavily disputed. The discussion about the use of citation-based indicators gained a new dimension when bibliometric indicators were set not only to be used for monitoring national or institutional research performance, but when they also became components of formulae for the funding of scientific research. In this context the question arose to what extent citations might reflect the "quality of research" and, if so, whether those measures are also reliable in the sense that the authors themselves might deliberately influence or even manipulate the measurable impact of their publications. In particular, this might be possible by forming of so-called citation cliques or by exaggerating citing of the authors' own work. Thus, users of bibliometric results are sometimes condemning author self-citations as a possible means of artificially inflating citation rates and thus of strengthening the authors' own position in the scientific community. According to a recent Flemish article (*Knack,* 2004) author self-citations are highly problematic and suspect in determining the quality of scientific journals. The article was entitled *Eigen lof stinkt* ("Self- praise is no recommendation"). Although authors of scientific publications are somewhat more cautious in drawing conclusions, the current controversial discussion in the scientific literature is equally prominent and strong (see *Science,* 2003).

In order to understand the role of author self-citations, one has first to consider it as part of the citation behaviour within the framework of scientific communication. In a second step, the role of citations in general as considered in information science will be discussed. This will be used as the groundwork to understand what problems might arise if citation data are used by science policy in decision making. Furthermore, we will see what bibliometric research can contribute to the reasonable use of these data in practice.

From the bibliometric perspective, basic regularities of author self-citations have been studied at different levels of aggregation (cf. *Aksnes* 2003, *Glänzel* et al., 2004). Regularities related to the ageing, to the relation between self-citations and foreign citations and to the interdependence of self-citations with other bibliometric indicators have been found which allow the conclusion that at high (macro-) levels of aggregation self-citations can be considered a natural part of scientific communication. However, meso-studies (e.g., *Aksnes* 2003, *Nederhof* et al., 1993, *Thijs* and *Glänzel*, 2004) have shown that inclusion of self-citations might form a source of error, for instance, in the ratio of observed/expected citation impact. The Centre for Science and Technology Studies (CWTS) at Leiden University (The Netherlands), for instance, uses self-citation rates to detect departments with deviant levels of self-citation. Although information scientists and bibliometricians have – as shown by these examples – paved the way for a pragmatic discussion, the policy-driven approach to author-self citations as used in research evaluation and calculation of funding formulas results in the interpretation as a source of distortion of the measurable impact of scientific research. The deviating interpretation of one and the same phenomenon in different contexts is from the understanding of the three approaches (information science – bibliometrics – science policy) in fact plausible, but repercussions caused by policy use on the scientists' communication behaviour might become measurable (e.g., *Glänzel* and *Debackere*, 2003, *Butler*, 2004).

The main objective of the present study is twofold: (1) to show the aims and means of quantitative interpretation of bibliographic features in bibliometrics and their re-interpretation in research policy, and (2) to summarise the state-of-art in self-citation research in bibliometrics. In order to be able to understand the role of author self-citations, self-citations have to be considered a part of the authors' citation behaviour within the framework of documented scientific communication.

2. The the social network of scientific communication

2.1. The citation as considered in information science

Citation behaviour is a complex phenomenon studied both by information science and sociology of science. Whilst the first discipline regards the citation as means of dissemination and identification of information, the latter one sees the social component of the scientific information in the citation. This approach is clearly expressed, for instance, by studies of the motivation and preferences in citation behaviour of authors. In his book "The Citation Culture" *Paul Wouters* (1999) treats both mentioned approaches. About at the same time the discussion issue of the journal *Scientometrics* about *citation theories* appeared. The discussion about reappraisal of existing theories of citation was initiated by *Loet Leydesdorff* (1998). In what follows, we will give a concise overview of the interpretation of literature citation in the sciences reflecting different arguments from both, information scientists and sociologists of science. Scientists in these fields still distinguish between the function of individual citations and the reasons for giving them.

Susan Cozzens (1989) argued that citation is only secondarily an element of the reward system. Primarily, it is rhetorical part of persuasively arguing for the knowledge claims of the citing document. Her approach is therefore called *rhetoric-first model*.

By contrast, *Linda C. Smith* (1981) observes in line with the information science approach tersely that *"citations are signposts left behind after information has been utilized"*. *Blaise Cronin* (1981) develops this idea further by observing that citations are *"frozen footprints in the landscape of scholarly achievement … which bear witness to the passage of ideas"*. However, he also points to certain problems with regard to reference practices as he concludes: *"If authors can be educated as to the informational role of citations and encouraged to be more restrained and selective in their referencing habits, then it should be possible to arrive at a greater consistency in referencing practice generally."*

Glänzel and *Schoepflin* (1999) gave a more pragmatic interpretation that allows also bibliometric applications. They consider the citation "*one important form of use of scientific information within the framework of documented science communication*," however, in general, neither form of nor reason for the concrete information use is specified. Although citations cannot describe the totality of the reception process, they give, according to *Glänzel* and *Schoepflin* (1995) "*a formalised account of the information use and can be taken as a strong indicator of reception at this level*". This view applies also to author self-citations. A self-citation indicates the use of own results in a new publication. Authors do this quite frequently to build upon own results, to limit the length of an article by referring to already published methodology, or simply to make own background material published in "grey" literature visible. Acknowledging the own contribution to the advancement of the research topic in question is, of course, an important reason for giving self-citations, too. Further reasons for author self-citations are briefly summarised by *Pichappan* and *Sarasvady* (2002). More generally, *Garfield* and *Weinstock* (cf. *Weinstock*, 1971) listed fifteen reasons why authors give citations to others' work. These reasons apply, of course, partially to self-citations, too. The reasons listed by Garfield and Weinstock can be sub-divided into three groups according to the *reception* of others' results as well as according to the *relevance* of citations. On one hand, there are citations expressing 'positive' reception (e.g., 'paying homage to pioneers'), 'neutral' citations (e.g., 'providing background reading') and citations reflecting 'negative' reception such as 'disclaiming work or ideas of others' or 'disputing priority claims of others'. On the other hand, references can be considered 'relevant' (for instance, the cited work was the essential groundwork for the own research), 'less' relevant or even 'redundant' if, for instance, a document is cited along with another one that already covers all information relevant for the citing publication. These differences might visualise that not all citations have the same weight. Of course, this statement applies to individual documents and individual citations. However, bibliometrics is based on frequency distribution and statistical functions. The concrete citation context is thus replaced by bibliometric indicators. This view essentially changes the picture described by sociology of science and information science. We will have a closer look at this below.

2.2. The citation as considered in bibliometrics

From the historical perspective, information scientists and bibliometricians proceeded from the same approach to citations, particularly from citations as expression of and possibly also measure for the use of information. The first citations studies consequently attempted to use citation measures in library science. *Gross* and *Gross* (1927) published their classic citation-based study on determining the most important journals of chemistry in order to aid the decision which chemistry periodicals should best be purchased by small college libraries.

The *journal impact factor* (JIF) according to Eugene Garfield (e.g., *Garfield*, 1979) had originally a similar function. JIF was the first standardised statistical journal citation measure. Its interpretation has completely changed during the last three decades: the utilisation measure originally designed for the use in libraries, information retrieval and scientific information has evolved to a "quality" measure in the sense of research evaluation for systematic application to science policy. This "perspective shift" is at least in part a consequence of bibliometric research and indicator engineering.

According to *Pritchard* (1969) bibliometrics is "*the application of mathematical and statistical methods to books and other media of communication*". Thus bibliometrics is concerned with *sets of publications*. Sceptics state that citations are subject to the intentions of authors and that information sources are therefore arbitrarily filtered. This might possibly the case in individual documents. However, if publication sets of a large number of authors are studied, this phenomenon is not longer typical of the citation process. Several reasons in Weinstock's list lose their importance if large sets of papers and large citation windows are used. Criticism, priority claims and correcting the work of others is practically never expressed through repeated and frequent citations. References to frequently cited publications are above all given in a positive or neutral context. Evaluative citation analyses may therefore be based on citation frequencies. According to *Westney* (1998), citations are useful indicators of scholarly impact: "*Despite its flaws, citation analysis has demonstrated its reliability and usefulness as a tool for ranking and evaluating scholars and their publications. No other methodology permits such precise identification of the individuals who have influenced thought, theory, and practice in world science and technology.*"

The fact that a document is less frequently cited or even (still) uncited several years after publication provides information about its reception by colleagues but does not reveal anything about its quality or the standing of its author(s) in the community. Uncited papers by Nobel Prize winners may just serve as an example. However, "*if a paper receives 5 or 10 citations a year throughout several years after its publication, it is very likely that its content will become integrated into the body of knowledge of the respective subject field; if, on the other hand, no reference is made at all to the paper during 5 to 10 years after publication, it is likely that the results involved do not contribute essentially to the contemporary scientific paradigm system of the subject field in question*" (*Braun* et al., 1985).

These interpretations substantiate that citations may be used to directly or indirectly measure the impact of scientific literature in the framework of documented scientific communication that takes place within the scientific community. Of course, this model is restricted to the sciences since citations and references in the arts and humanities and in part also in the social sciences have, besides the above-mentioned functions, other reasons and functions, too.

Several authors in bibliometrics consider the number of citations received automatically an expression or even a measure of 'quality'. As we will see in the following section, this view does not only results in restrictions concerning the application of bibliometric methods, but might cause also undesired side-effects when citation analysis is applied to the micro level.

From the bibliometric-statistical perspective the author self-citation is utmost interesting. Already *Lawani* (1982) has pointed to the fact that there are two types of self-citations, particularly, synchronous (retrospective) and diachronous (prospective) self-citations. The first ones refer to the share of self-citations in the total of cited work, the latter ones to the share in the citing literature. Although the first type has little to do with the citation impact, it is nonetheless fallen in disrepute because of the general negative opinion on self-citations as

well as because of the fact that missing "foreign" references might express certain isolation of the research in question or at least lacking communication. On the other hand, the almost complete lack of self-citation is also problematic: It might be considered lack of originality in research. In any case, author self-citations are part of the system of science communication since they indicate also a successful and dynamic publication activity in refereed journals. Only the almost absolute lack of self-citations over a longer period as well as an always overwhelming share of self-citations must be considered "pathological" cases.

2.3. The citation as considered in science policy

Unlike in bibliometrics, where citations are used as measure of reception of scientific results, science policy and research management regards citations as an expression of impact or even quality. If literature is cited frequently then this approach seems to be reasonable: good reception and considerable citation impact can be considered an expression of quality, too, as outstanding bibliometric indicators in general reflect a good state of the corresponding science system. On the other hand, if the citation rate of a given paper set is low, bibliometrics cannot immediately conclude on the quality of underlying research. Of course, the situation becomes problematic if at the institutional or even national level citation indicators remain constantly low in a given science field. However, to draw valid conclusions, further research on the causes is then necessary. Although *Holmes* and *Oppenheim* (2001) showed that citation measure significantly correlate with other quality measures, the science policy re-interpretation of citations as an element of the reward system has severe consequences.

The role of self-citations is perhaps the most striking example for the mentioned consequences. If the citation expresses reward, self-citations distort necessarily the system as such. Consequently, self-citations are considered potentially falsifying the impact of research. Possible repercussions on the authors' citation behaviour are that they might feel urged to avoid self-citations which, in turn, might distort 'natural' communication behaviour in a self-organising system. Figure 1 presents the relationship of different interpretations of the citation in information sciences and science policy schematically.

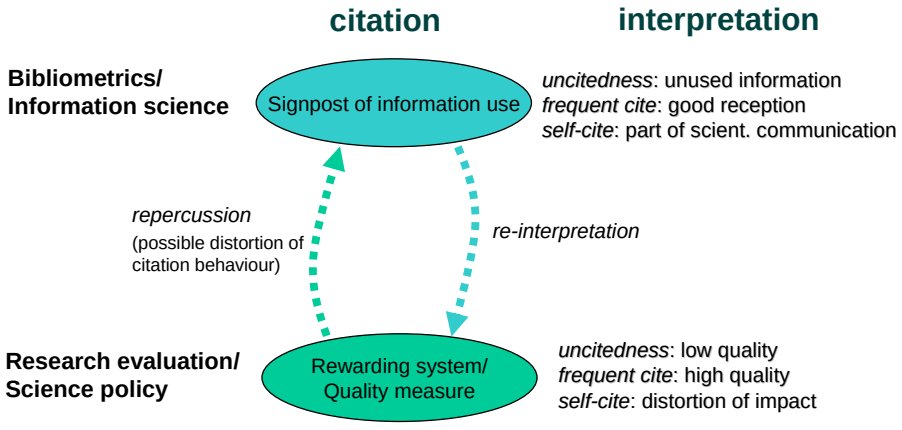


Figure 1 The process of re-interpreting the notion of citation and its consequences

Glänzel and *Debackere* (2003) have already discussed possible distorted behaviour based on policy use and misuse of bibliometric data. Consistent policy use of bibliometric indicators might potentially induce changes in the publication, citation and collaboration behaviour of scientists (both positive and negative): If bibliometric tools have an effect on decision-making

in science policy and research management and the scientific community recognises the feedback in terms of their funding, then there might be measurable repercussions on their behaviour, too. Re-interpretation and "perspective shift" as mentioned above might even catalyse this process.

3. The role of author self-citations in scientific communication

In this section, an overview of bibliometric research on author self-citations is presented. We briefly summarise, among others, questions tackled by *MacRoberts* and *MacRoberts* (1989), first quantitative results published by *Snyder* and *Bonzi* (1998), methodological and mathematical considerations by *Lawani* (1982), *Rousseau* and *van Raan* (1998) and more recent results by *Pichappan* and *Sarasvady* (2002), *Aksnes* (2003), *Glänzel* et al. (2004) and *Glänzel* and *Thijs* (2004a, 2004b) as well as new, unpublished material by the authors (*Thijs* and *Glänzel*, 2004, *Glänzel* and *Debackere*, 2005, *Glänzel* et al., 2005). Both the above review and the state-of-the-art report are aiming at paving the way to a more pragmatic and realistic approach to an important form of scientific communication.
First *Garfield* and *Sher* used a quantitative approach to author self-citations; they found that the share of authors self-citations in basic research amounts to 20% on average (*Garfield* and *Sher*, 1963). *MacRoberts* and *MacRoberts* have given a first overview of the unsolved problem of self-citations in their critical review on problems of citation analysis in 1989. At the same time they approximated this share by 10%-30%. However, these are synchronous self-citations. Again, *Lawani* was the first who distinguished between these two types of self-citations. Beside the principle discussion on the role of author-self citations, there is no real consensus concerning how this type of self-citations should be defined operatively. In practice, two different approaches to direct self-citations are in use.

At the micro level, that is, on the level of individual authors, a direct self-citation for an author A occurs whenever A is also (co-)author of a paper citing a publication by A. This definition cannot, however, be applied to higher levels of aggregation, that is, when publications and citations are aggregated over sets of different (co-)authors, and the notion of self-citations is uncoupled from an individual author A. At the meso and macro level, other criteria have to be used to determine what is considered a self-citation. Normally the definition by *Snyder* and *Bonzi* (1998) is used. According to this approach, a self-citation occurs whenever the set of co-authors of the citing paper and that of the cited one are not disjoint, that is, if these sets share at least one author. Although the reliability of this methodology is affected by homonyms (resulting in Type II errors by erroneous self-citation counting) and spelling variances/misspellings of author names (resulting in Type I errors by not recognising self-citation), at high levels of aggregation, for instance at the meso and macro level, there is no feasible alternative to the method in this study.

Aksnes (2003) studied the role of self-citation in the scientific production of Norway in the quite long period (1981-1996). He found that share of self-citations decreases with growing time window. He found also a strong positive correlation between the number of self-citations and the number of co-authors of the publications as already observed earlier by *van Raan* (1998). Self-citations are characterised by large variations among different scientific disciplines. Aksnes concluded that at lower levels of aggregations, such as the meso-level, self-citations might represent a serious problem, and recommended to preferably removing self-citations before making comparisons or, at least, to carefully consider by-effects caused by self-citations before using citations as indicators of scientific impact.

Glänzel et al. (2004, 2005) have given a mathematical model for the phenomenon of self-citations, and extended the above-mentioned results to all science fields and all countries at the macro level, and to six European countries at the meso level. Before we briefly summarise the most important results we give some basic figures on author self-citations. In most bibliometric studies four document types are used: articles, notes, letters and reviews. In 2001 were 27.2% of all citations received by articles and notes indexed in the Science Citation Expanded® of the Institute for Scientific Information (Thomson – ISI, Philadelphia, PA, USA) author self-citations in the citation window 2001–2003. The share of self-citations for letters amounted to 25.5% which is in line with that of articles and notes. The share of self-citations for reviews in the same period with 13.9% was significantly lower. These figures substantiate the validity of the information-science approach explaining (self-)citations as information use in the context of science communication: Working in research projects often results in publishing follow-ups which quite necessarily increase the number of self-citations. Follow-up to reviews are much less frequent. On the other hand, reviews attract statistically more citations than other document types. The latter two reasons explain the statistically low self-citation share for these documents.

According to the model by Glänzel et al., the random variable $\xi(t)$ denotes the number of self-citation a paper published at time $s = 0$ has received in the period $[0, t]$. Putting $s = 0$ does not result in any restriction of generality. The rate of foreign citations (non-self-citations) is then denoted by $\zeta(t)$. Thus we have $\eta(t) = \xi(t) + \zeta(t)$ for the total citation rate of a paper published at time 0. $\xi(t)$ and $\eta(t)$ are not independent random variables since $0 \leq \xi(t) \leq \eta(t)$ and $\eta(t) = 0$ implies $\xi(t)) = 0$. In order to simplify notation we will use ξ and ζ without indicating the time parameter. Table 1 presents the empirical probability of receiving both foreign and self-citations. All documents of the above four types indexed in the 1994 volume of the *Web of Science* (WoS) of the Institute for Scientific Information (Thomson – ISI, Philadelphia, PA, USA) have been analysed in two different citation windows, particularly, in a three-year and a ten-year window. The chance of receiving foreign citations $P(\zeta>0, \ldots)$ considerably increases with growing citation window. This effect is even stronger if one assumes the case if a paper receives both foreign and self-citations ($P(\zeta>0, \xi>0)$). The share of uncited papers $P(\zeta=0, \xi=0)$ practically halves. The most interesting trend concerns the case if a paper receives only self-citations ($P(\zeta=0, \xi>0)$). Ten years after publication this share is marginal but it is already small in the three-year window. This example might visualise that self-citations are normally companions of foreign citations; the chance of compensating lacking foreign "impact" through self-citing is very small.

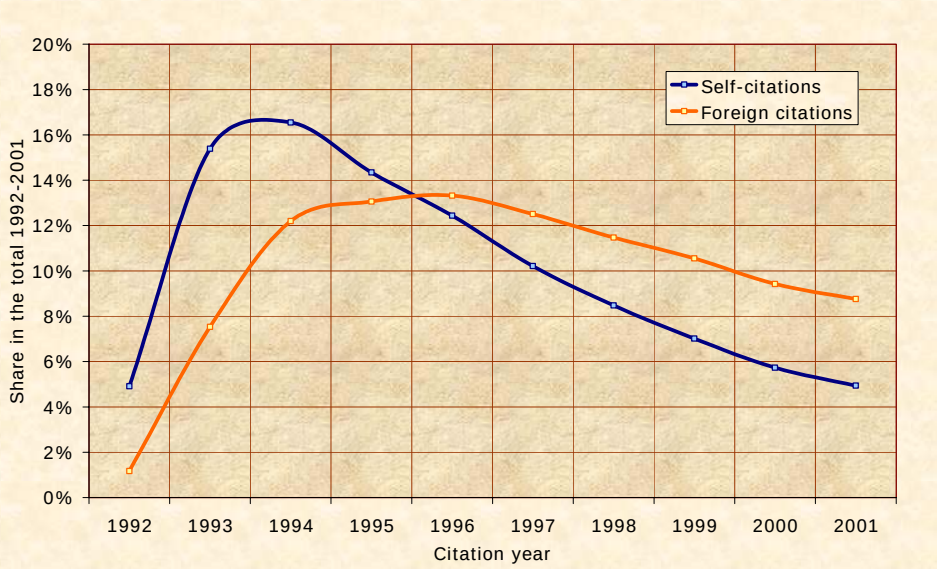
Tabel 1 Relative frequency of self-citations and foreign citations to papers published in 1994 using two different citation windows

Rel. frequency	**1994-1996**	**1994-2003**
$P(\zeta = 0, \xi = 0)$	38.0%	19.7%
$P(\zeta = 0, \xi > 0)$	11.6%	5.6%
$P(\zeta > 0, \xi = 0)$	22.3%	26.2%
$P(\zeta > 0, \xi > 0)$	28.1%	48.5%

Source: Glänzel and Debackere (2005)

The figures in Table 1 substantiate that the time factor plays and important part in the share of self-citations. Figure 2 presents the life-time curve of both foreign and self-citations of all articles, notes, letters and reviews indexed in the 1992 annual volume of the WoS. The ageing

of self-citations is much faster than that of foreign citations. This observation applies to all science fields, and is in line with the results by *Aksnes* (2003). However, the deviation of the ageing patterns of individual subject fields from each other and from that of all fields combined is considerable, but ageing of self-citations is somewhat less field-specific than that of foreign citations (cf., *Glänzel* et al., 2004). The strong increase of self-citations during the first three years after publication can at least in part be explained with follow-up in the framework of continuous project work, but own results lose their attractiveness for own research soon whereas they still might provide impulses for ideas and work of others.



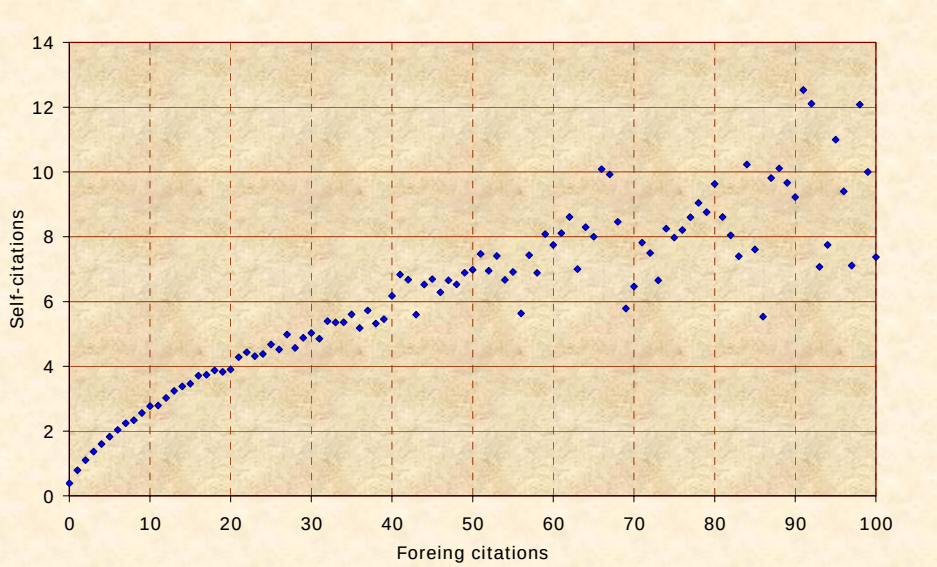
Source: Glänzel et al. (2004)

Figure 2 Empirical density of the life-time distribution of foreign citation and self-citations for all science fields combined

Further analyses have shown that self-citation indicators become quite stable in a period of 3-4 years after publication. A citation window of three years is thus sufficient for self-citation studies (cf. *Glänzel* et al., 2004).

The second important question concerns the possible interdependence of the two forms of citations. Theoretically, the two variables ξ and ζ might be independent. However, if self-citations prove to be independent of the number of foreign citations then citation indicators should be cleaned from self-citations, indeed, since in that case they were really potential means of manipulation. On the other hand, if the number self-citations is uniquely determined by the number of foreign citations, that is, if the number of self-citations is a mathematical function of foreign citations, then self-citations can be omitted as they are already implicitly expressed by foreign citations. The study by *Glänzel* et al. (2004) has shown that none of the two cases holds. They are not independent but the correlation between them is relatively weak. In order to gain a deeper insight in the possible inter-dependence of the two variables, the conditional expectations $E(\xi|\zeta = k)$ have been calculated for all fields combined and for all possible citation windows. The condition $E(\xi|\zeta) \equiv E\xi$ is necessary but not sufficient for independence. Because the hypothesis of independence was already rejected we have conducted a regression analysis for the two variables ζ and $E(\xi|\zeta)$. Figure 3 presents the plot of foreign citations vs. mean self-citation rate for a three-year citation window, particularly 1992-1994, for all fields combined. The strong correlation between the two variables $E(\xi|\zeta)$

and ζ led to the hypotheses that there exists an appropriate real function f, so that $E(\xi|\zeta) = f(\zeta)$. The regression analysis resulted in the approximation $f(k) \approx (k + ¼)^{½}$ for the stationary case and for all fields combined. This means that the expected self-citation rate increases with growing number of foreign citations but at a sub-linear level, roughly following a square root law (cf. Figure 3).



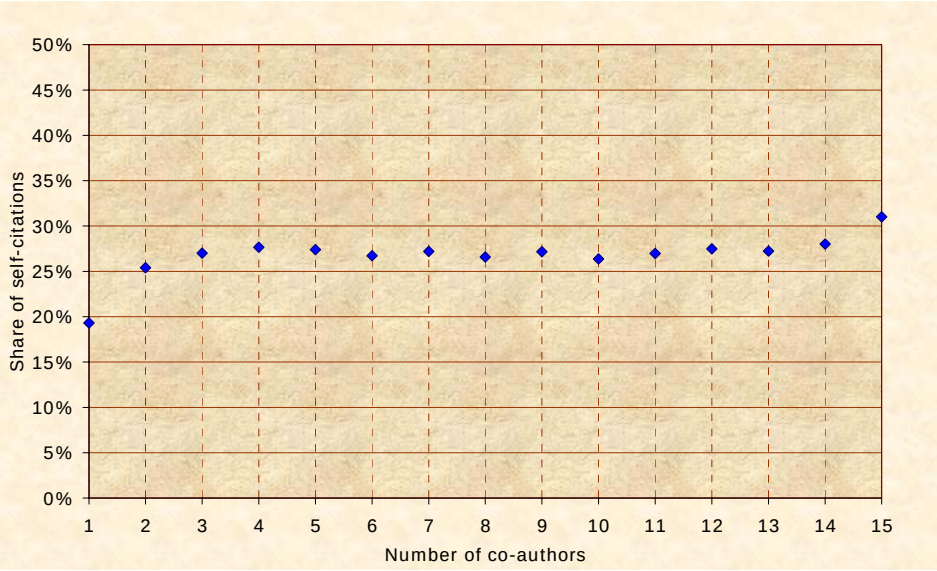
Source: Glänzel et al. (2004)

Figure 3 The plot of foreign citations vs. mean self-citation rate for three selected citation windows with at most 100 foreign citations for all fields combined
(publication year: 1992, citation window: 1992-1994, all science fields combined)

In a further study by *Glänzel* and *Thijs* (2004a) these methods have been applied to the analysis of national citation indicators in 15 subject fields in the sciences, social sciences and humanities. Although self-citation indicators are clearly influenced by subject-specific factors and national peculiarities, the correlation between relative citation indicators including and excluding self-citations is always strong. This led to the conclusion that – at the macro level – there is no need for excluding self-citations in evaluative bibliometrics. The results of the institutional analysis, conducted by *Thijs* and *Glänzel* (2004), partially deviate from both their field standard and the corresponding national standard. Research at the meso level is more characterised by specific profiles than the national level is. This is in line with the recommendation by *Aksnes* (2003) who argued that at lower levels of aggregations, such as the meso-level, self-citations might represent a serious problem. The outcomes of both studies confirm that self-citation indicators should be used as supplementary indicators in evaluative bibliometrics at the meso level.

According to a widespread view, increasing self-citations are a by-effect of intensifying scientific collaboration. From the "combinatorial" viewpoint this seems to be quite plausible. One can thus give a simple estimate of the probability of getting cited in function of the number of co-authors (cf. *Rousseau*, 1992). However, practice looks always a bit different. A recent study by *Glänzel* and *Thijs* (2004b) dealt with the question of in how far co-authorship influences self-citation patterns. The study confirmed that multi-authorship goes with increasing self-citation rates, but the increase is much weaker than that of foreign citations. Thus, multi-authorship increases above all the probability to be cited by others. However, the

share of self-citations of single-authored papers is pronouncedly lower than that of multi-authored papers. The actual number of co-authors in the set of multi-authored papers has no essential influence on the share of self-citations in all citations a paper receives. Figure 4 shows the plot of the conditional mean share of self-citations vs. number of co-authors of individual papers in a three-year citation window. The two effects mentioned are clearly visible in the chart.



Source: Glänzel and Thijs (2004)

Figure 4 Plot of the conditional mean share of self-citations vs. number of co-authors (publication year: 1992, citation window: 1992-1994, all science fields combined)

The method proposed by *Snyder* and *Bonzi* (1998) does not take into account the weight of self-citing authors among co-authors of both the cited and citing papers. It is quite obvious a self-citation to a single-authored paper given in a single-authored paper is much "stronger" than a self-citation to a multi-authored paper given in a multi-authored paper but caused by one joint co-author only. In a recent paper, *Glänzel* et al. (2005) made the attempt to quantify the weight of self-citations with respect to co-authorship through replacing binary "0–1" formula for the 'foreign citation' – 'self-citation' relation by a continuous measure reflecting a fuzzy situation. The method was based on the *Jaccard Index*. The most striking feature of such fractional self-citation count is the extremely fast ageing. Author self-citations become practically negligible after period of ten year, but also three years after publication, the effect of author self-citations is quite low if compared with binary counts. Fractionation of self-citations might also help to bridge the gap between the two approaches to self-citations for individual authors and for document sets, respectively, as described in the outset.

4. Conclusions

Citing of scientific literature has to be considered part of social processes in the science system. The results substantiate on the one hand that "pure arithmetic" models such as the increase of self-citations as a function of the number of co-authors and on the other hand, the "pure psychological" approach, namely, that the self-citation is subject to the will and arbitrariness of the authors alone, are in practice not applicable. Of course, individual citation behaviour might sometimes extremely deviate from statistical patterns, but no alarming trends

could be found so far. The macro studies show that there is no reason for condemning self-citations in general or for removing them from citation statistics. On the other hand, supplementary indicators based on self-citations are useful to understand communication patterns. Above all, in meso and micro studies such indicators might help to clarify if the observed citation impact really reflects the reception of the research results by the scientific community.

***

Figures 2–4 are reproduced with permission of the publisher.

References

[ANONYMOUS] (2003), Citing self, *Science*, 5616 (4), 47.

[ANONYMOUS] (2004), *Eigen lof stinkt, Knack*, 21 April 2004.

Aksnes DW (2003), A macro-study of self-citations. *Scientometrics*, 56 (2), 235-246.

Berghmans T, Meert A P, Mascaux C, Paesmans M, Lafitte JJ, Sculier JP (2003), Citation indexes do not reflect methodological quality in lung cancer randomised trials, *Annals of Oncology*, 14, 715-721.

Butler L (2004), What happens when funding is linked to publication counts? In: H.F.M. Moed, W. Glänzel, U. Schmoch (Eds), *Handbook of Quantitative science and Technology Research. The use of Publication and patent statistics in studies on S&T Systems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 389-405.

Braun T, Glänzel W, Schubert A (1985), *Scientometric Indicators. A 32 Country Comparison of Publication Productivity and Citation Impact*. World Scientific Publishing Co. Pte. Ltd., Singapore * Philadelphia.

Cozzens SE (1989), What do citations count? The rhetoric-first model. *Scientometrics*, 15, 437-447.

Cronin B (1981), The need for a theory of citation, *Journal of Documentation* 37, 16-24.

Garfield E, Sher IH (1963), New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, 18, 195-201.

Garfield E (1979), *Citation Indexing: Its Theory and Applications in Science, Technology and Humanities*. New York: Wiley.

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, Schoepflin U (1995), A Bibliometric Study of Reference Literature in the Sciences and Social Sciences, *Information Processing and Management*, 35, 31-44.

Glänzel W, Debackere K (2003), On the opportunities and limitations in using bibliometric indicators in a policy relevant context, In: R. Ball (Ed.), *Bibliometric Analysis in Science and Research: Applications, Benefits and Limitations*, Forschungszentrum Jülich (Duitsland), p. 225-236.

Glänzel W, Thijs B, Schlemmer B (2004), A bibliometric approach to the role of author self-citations in scientific communication, *Scientometrics*, 59 (1), 63-77.

Glänzel W, Thijs B (2004a), The influence of author self-citations on bibliometric macro indicators, *Scientometrics*, 59 (3), 281-310.

Glänzel W, Thijs B (2004b), Does co-authorship inflate the share of self-citations? *Scientometrics*, 61, (3), 395-404.

Glänzel W, Debackere K (2005), *De rol van auteur-zelfcitaties zoals beschouwd door informatiewetenschappen, bibliometrie en wetenschapsbeleid*, In: Debackere K, Veugelers R (Eds.), Vlaams Indicatorenboek 2005, in press.
Gross PLK, Gross EM (1927), College Libraries and Chemical Education, *Science*, 66, 385-389.

Glänzel W, Thijs B, Schubert A (2005), *The weight of author self-citations*, paper to be presented at the 10th International Conference of the International Society for Scientometrics and Informetrics, Stockholm (Sweden), 24-28 July, 2005.

Holmes A, Oppenheim C (2001), Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management, *Information Research*, 6 (2).

Lawani SM (1982), On the heterogeneity and classification of author self-citations, *JASIS*, 33 (5), 281-284.

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

MacRoberts MH, MacRoberts BR (1989), Problems of citation analysis: A critical review. *JASIS*, 40 (5), 342-349

Nederhof AJ, Meijer RF, Moed HF, van Raan AFJ (1993), Research Performance Indicators for University Departments - A Study of an Agricultural University, *Scientometrics*, 27 (2), 157–178.

Pichappan P, Sarasvady S (2002), The other side of the coin: The intricacies of author self-citations, *Scientometrics*, 54 (2), 285-290.

Pritchard A (1969), Statistical bibliography or bibliometrics? *Journal of Documentation* 24, 1969, 348-349.

Rousseau R (1992), Why am I not cited, or why are multi-authored papers more cited than others? *Journal of Documentation*, 48 (1), 79.80.

Smith LC (1981), Citation Analysis. *Library Trends*, 30 (1), 83-106.

Snyder H, Bonzi S (1998), Patterns of self-citation across disciplines. *Journal of Information Science*, 24, 431-435.

Thijs B, Glänzel W (2004), *The influence of author self-citations on bibliometric meso-indicators. The case of European universities.* Paper presented at the 8th S&T Indicators Conference, Leiden (Netherlands), 24 September 2004.

van Raan AFJ (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.

Warner J (2000), A critical review of the application of citation studies to the Research Assessment Exercises, *Journal of Information Science*, 26, 453-459.
Weinstock N (1971), Citation indexes, In Kent A. (Ed.). *Encyclopedia of Library and Information Science*, New York: Marcel Dekker, 5, 16-41.

Westney LCH (1998), Historical rankings of science and technology: A citationist perspective, *The Journal of the Association for History and Computing*, 1 (1), June 1998.

Wouters P (1999), *The citation culture*, PhD Thesis, Private edition.