SAVVY SEARCHING
Deflated, inflated and phantom citation counts

Peter Jacso
University of Hawaii, Hawaii, USA

Abstract

Purpose – The purpose of this paper is to clarify some issues regarding citation indexing, analysis and searching.

Design/methodology/approach – The paper begins with a discussion on an article in the *D-Lib Magazine* and then focuses on deflated citation counts and inflated and phantom citation counts.

Findings – The combination of the inflated citation count values dispensed by Google Scholar (GS) with the ignorance and shallowness of some GS enthusiasts can be a real mix for real scholars.

Originality/value – The paper offers insight into deflated, inflated and phantom citation counts.

Keywords Information searches, Information retrieval

Paper type Research paper

Introduction

Citation indexing, analysis and searching have lately received increased attention in the press. Until recently the topic has been in the realm of scholarly (and law) journals. Scientists and social scientists, librarians and information professionals observed, measured and analyzed the citation behaviour and practice in different disciplines, and argued about the value of citation indexes. As Google wandered into the territory by launching Google Scholar (GS) at the end of 2004, the topic is expected to appear in the ultra-light morning television chat shows run by ultra-light TV personalities who are meant to light up our mornings. Journalists of reputable scientific magazines also jump on the bandwagon and report about developments. As one day they report about the public health aspects of the avian flu, then the next one summarize the geological background of yet another earthquake and/or tsunami, and finish off the week with an account of the half-read and even less understood studies about the impressively high citation counts

---

The opinions expressed in this paper are the author's; and are not necessarily shared by the editor or publisher.

Editor's note: Serendipitously, *Encyclopaedia Britannica Online* www.eb.com./ has recently reinforced Professor Jacso's comments in this issue's Savvy Searching column regarding the accuracy of *Nature's* reporters. Dated March 2006 and titled "Fatally flawed: refuting the recent study on encyclopedic accuracy by the journal *Nature*" the Britannica rebuttal addresses claims in the 15 December 2005 article in *Nature* that claimed to compare the accuracy of *Encyclopaedia Britannica Online* with *Wikipedia*. Here is an extract that gives the flavour of the *Britannica* findings: "almost everything about the journal's investigation, from the criteria for identifying inaccuracies to the discrepancy between the article text and its headline, was wrong and misleading" – strong stuff indeed, but I for one am convinced by the evidence marshalled in this rebuttal of the *Nature* findings, and believe that this confirms Professor Jacso's suspicions about the accuracy of *Nature* reporting. Read the *Britannica* piece yourself, and see what you think . . .

dispensed by GS, the shallowness and hollowness of those easy pieces is not surprising. Some scholars (who are possibly experts in their own field but novices in using the citation databases and interpreting their results) are also tempted to jump on the bandwagon, and run with some numbers as Forrest Gump ran with the imaginary ball. It seems to be an appropriate time to clarify some issues, as searchers who make a living by being information savvy have to look behind those numbers and put them in proper perspective.

The *D-Lib Magazine* article

D-Lib Magazine, one of the premier, open access web-born journals in library and information science published an article about a comparison of citation counts reported by Web of Science (WoS), Scopus and GS for the *Journal of the American Society for Information Science and Technology* (*JASIST*) and its predecessor (*JASIS*). The authors analyzed data for two years: 1985 and 2000. For 1985 the mean citation counts were 14.1 for WoS, 6.8 for GS and 3.7 for Scopus. (To the credit of Scopus it makes it clear that its abstracting/indexing records are enhanced with cited references only from 1996.) For the year 2000 the mean citation counts were 7.6 for both WoS and Scopus, and 12.1 for GS. Most of the blog entries and the print commentaries paid attention only to the statistics of 2000.

The test was based on a sample of 105 items. The researchers rightly warned about the implications of the very limited test, but on the "what bleeds, leads" principle of media coverage, journalists quickly produced catchy titles, added their two cents' worth and omitted to mention the results for the 1985 comparisons. Some researchers then joined in and shared their findings in blogs, magazine and journal articles.

The counts published in *D-Lib Magazine* seem to be correct. I know simply because I happen to have contemporaneous test results. About the same time that the authors of the *D-Lib Magazine* article did their test in May 2005, I did my own hit and citation counts comparisons in WoS, Scopus and GS for a number of topics and a number of journals (Jacso, 2005a, b), including *JASIS* and *JASIST*. My total hit counts and citation counts for the two years are not the same as the counts in the *D-Lib Magazine* article because I did not filter out the short communications, as some of these are also cited.

More important, my analysis covered the entire run of *JASIS/JASIST* (not including the predecessor title, *American Documentation*) since 1970. The differences of hit counts and citation counts provide a more comprehensive picture for the dynamics of the coverage of the journal than the two sample years measured by the authors of the *D-Lib Magazine* article. It also shows that 2002 and 2003 were far the best years of coverage for these journals in GS (Figure 1).

It is to be noted that searching by source journal name is fairly swift in WoS, which normalizes its source titles and offer browsable indexes both for the source and cited journals and items. Scopus has a browsable source journal name index, but its normalization is less than perfect. GS has no browsable indexes, so in GS searching by journal name is a Sisyphean task because of the many variants of the abbreviations, punctuation and plain typos in the journal names in the sources from which GS extracts them.

If the source document contains them correctly, GS takes care of generating large volume of errors in every data element (including also the author names, and the publication years) in the scanning and parsing process. A browsable journal name index would reveal the errors and inconsistencies, but GS does not even provide a list of sources covered, let alone a browsable index or normalization.
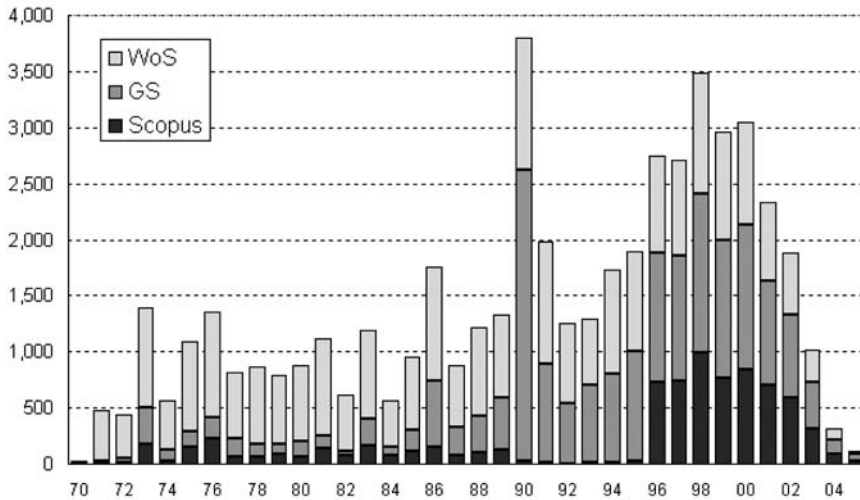
Academics and practitioners who are competent in the field of citation indexing cannot ignore these or take them lightly. You do not need to be an expert in bibliometrics to doubt the data provided by GS that the most productive author is "I INTRODUCTION" who published 40,100 documents covered by GS. To be fair, about 5 per cent may be by another author with additional and/or different middle initial, such as "II" "III" and "IV" – depending on which section is the introductory one in the book or journal article parsed by GS as author name.

Before you dismiss this as a small thing, you must realize that the primary and single authors of these items become unsearchable because their names do not appear in the GS record as author. Although it is not known to be a leading cause for myocardial infarction, Flavio Fento and Alain Karma, the authors of the most cited paper about the effect of vortex dynamics on the myocardium, may be prone to a heart attack when they look up their names in GS and do not see their paper published in *Chaos*. Ditto for the tens of thousands of other authors whose name has been replaced by "I INTRODUCTION" in GS.

Deflated citation counts

Some information professionals have mastered the knowledge of searching the citation index databases of ISI using the advanced search operators hosted by Dialog and STN. ISI introduced a few months ago in WoS the automatic reporting of citations received. This was a good idea, but it has a negative side effect on the interpretation of the citation counts. The automatically reported numbers reflect only the citations which match exactly the last name, first initial, middle initial, journal name, publication year, and chronological-numerical designation of the cited item as it appears in the master records. Most of the authors are rather sloppy with citations and very often misspell many of these data elements. The hundreds of citation styles used by journals makes cited references incredibly inconsistent and erroneous, and this chaos backfires significantly in citation indexing and citation searching.

ISI also offers the cited reference search mode to create a browsable index list of cited references. Savvy searchers can combine up to three citation elements to see the variations of the references: cited author name, cited journal name and cited year. This is the result, for example for a savvy search about Tom Wilson's article in *Information Research*. He used his journal to check the citation count reported for this article by WoS (3) and Google (34), as well as 16 articles published in *Information Research* (Figure 2).

When reporting the citation count, ISI plays it safe and shows only the exact match. It practically deflates the citation count and allows the searchers to visually scan and check the imperfect but close matches, which miss the article identification and/or the volume number. The syntax specified by the searcher in the cited reference search determines how many items will be shown. To keep this example simple, the search left wriggle room only for the author's middle initial by truncating after "WILSON T" and for abbreviation variations in the journal name, using the query string $< INF^*RES^* >$. All of the references indeed cite the article about "The nonsense of knowledge management" so for the savvy searcher the reward and the citation count is not three articles but 13 (Figure 3).

Inflated and phantom citation counts

The infatuation with GS is common because for certain disciplines (e.g. physics, computer science, and information science and technology) GS often reports higher citation counts than WoS – at least for current publications. This is because GS covers conference papers, books, book chapters and dissertations, while WoS practically limits itself to journal articles. The picture is different for the social sciences and humanities, where self-archiving and other types of web postings are not as prevalent as in the disciplines mentioned above, with the possible exception of economics, which has an excellent, citations-enhanced document archive, EconoPapers. It helps make GS look bigger that journals and topics in science, technology and medicine dominate all the comparative evaluations that I have seen. However, the hit counts and citation counts are often inflated.

The inflated counts reported by GS are partly due to the inclusion of non-scholarly sources, like promotional pages, table of contents pages, course reading lists, and to a much larger extent to the inability of GS to recognize real matches between cited and citing items, claiming a match where there is not even minimal "chemistry". The underlying problem is that GS inflates the citation counts as much as it inflates the hit counts, illustrated in the previous issue of this column (Jacso, 2006).

If it were not enough that GS obviously cannot handle Boolean operators, reporting 556,000 hits for the word "scholar" 939,000 for the word "scholars" but only 611,000 for the query "scholar or scholars" its range operation is also dysfunctional, as shown

Title: The nonsense of 'knowledge management'
Author(s): Wilson TD
Source: INFORMATION RESEARCH-AN INTERNATIONAL ELECTRONIC JOURNAL 8 (1): Art. No. 144 OCT 2002
Document Type: Article
Language: English
Cited References: 52   Times Cited: 3   FIND RELATED RECORDS ⓘ

Figure 2.
Automatically reported citation count in WoS

Cited Reference Search

<< Back to query

**Your search has found the following references.
Select only those cited references you want to include,
then click FINISH SEARCH.**
(Hint: Look for variants. Papers are sometimes cited incorrectly.)

FINISH SEARCH >> View the articles that cite the selected references.
The completed search will be added to the search history.

(Limit by language and document type)

CITED REFERENCE INDEX	Go to Page: 1 of 1 GO
References 1 -- 6	| ◀◀ ◀ [1] ▶ ▶▶ ▶|

SELECT PAGE SELECT ALL* CLEAR ALL or select specific references from the list.
When desired references have been selected from all pages, click FII
SEARCH to complete your search.

Select	Times Cited**	Cited Author	Cited Work [SHOW EXPANDED TITLES]	Year	Volume	Page	Article ID
☐	1	WILSON T	INFORM RES	2002	8		⬭
☐	3	WILSON TD	INFORM RES ◀	2002	8		Art. No. 144!
☐	1	WILSON TD	INFORM RES	2002	8		
☐	1	WILSON TD	INFORMATION RES	2002	⬭		⬭
☐	6	WILSON TD	INFORMATION RES	2002	8		
☐	1	WILSON TD	INFORMATION RES	2002	8	1	

* "Select All" adds the first 500 matches to your cited reference search, not all matches.
** Times Cited counts are for all databases and all years, not just for your current database and

Figure 3.
The list of citing
references for a test article

in Figure 4. How it can report half as many hits for the same word for 2005 as for
2005-2006 is more than enigmatic.

The less than savvy searchers happily accept the reported hit counts, just as naïve
tourists accept the first price in a bazaar. Unfortunately, they also present the hit
counts and the citation counts as facts – in journal articles and conference papers,
spreading the myth. GS knows very well that reporting any hit counts above 1,000 is
safe, as from then on GS does not show any more hits. That domain remains the world
of 1,001 + nights, with equally loose number of promised virgins and harem dancers,
which was never corroborated.

No wonder that Anurag Achariya, one of the developers of GS, flashes a big
smile from a small or large portrait from the pages of *Nature* as if he were going
to the bank when reading the news items in *Nature*. These tout the superiority of
GS, rehashing the promotional materials for the press, and/or summarizing and
commenting on the findings of some comparison tests by casual and/or very
novice users of WoS and GS, who probably never used citation databases before,
let alone understood or bothered to corroborate their hit counts and citation
counts.

Undoubtedly, GS made finding citation counts so easy that even the absolute
outsiders may feel that they "got it". Researchers often find it irresistible to take the hit
counts and the citation counts from the result list as is, plug them in a spreadsheet, let a
statistical package produce a graph or two (preferably with logarithmic scales), throw
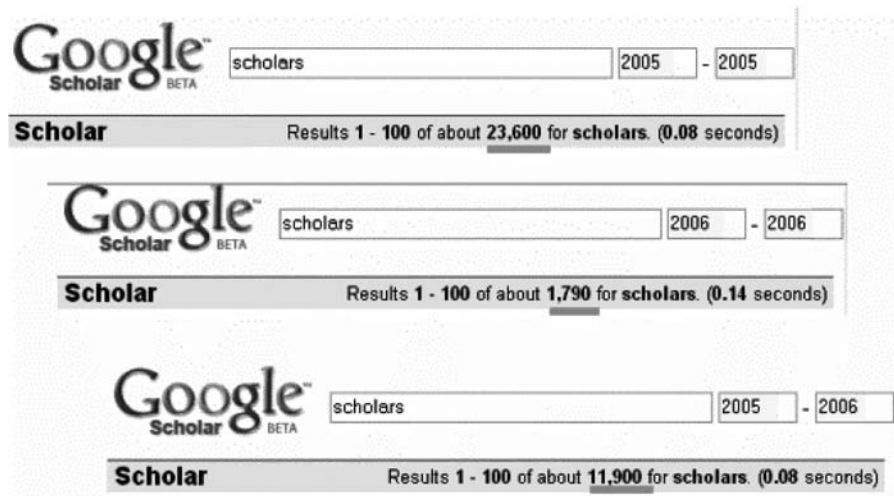in some statistical measures, add some comments and publish a paper.

Figure 4.
Strange range operation in GS

At first sight anyone would be impressed by the results retrieved for the search below. After all, having collected by mid-March of 2006 information about 1,790 items published in 2006 which are about scholars or at least have the word scholars is a big deal. Having citation counts for such documents for four of the top five items is even bigger deal. Having an item that accrued 15 citations in two months is an even bigger catch. Or is it? (Figure 5)

The fact that the first item is a digital pamphlet rather than a scholarly paper is a bit of a downer, but the second one seems to be a fantastic catch (even if day after day the GS link to the source document – or at least to its abstract – did not link, just timed out. Going straight to the archive, bypassing GS, would give you with the abstract.) This article must be hot, as it has already garnered 15 citations in two months (Figure 6).

The GS search links on the article page for Cited by Articles just returns the current article's URL, not a good sign. So we go back to GS to see the 15 items, which cite this paper. Interestingly, the articles citing this philosophical, sociological document are from physics journals. Maybe it signals a new synergy between the two cultures? But wait, how can articles from 1972, 1975, 1976 and 1998 cite a 2006 publication? Maybe they do not, but in GS we trust. If it brought up these records, they must cite the January 2006 Burke article which deal with Regge poles and Regge calculus, named after the theoretical physicist, Tulio Regge.

And what this has to do with the *European Review* article? Its abstract mentions reggae, the music. By the very lax matching criteria of GS, it is close enough, matching five of the six letters. I wonder if the designer of this matching algorithm would be surprised when landing in Haiti instead of Tahiti for a honeymoon courtesy of the matching intelligence of his/her travel agent. After all, if not much else, five letters are common in the names of the two countries, and reggae is popular in Haiti. On a side note let me mention that the items below have 50, 34, and 40 citations in WoS. Yes, they cite an article by PG Burke, not P Burke, and not from *European Review* but from *Computer Physics Communication*, titled "Program for calculating Regge trajectories" not Introduction, published in 1969, not in 2006. This is vintage fool's gold from GS (Figure 7).

Figure 5.
Result list for search for
documents about scholars
published in 2006

As an important side note, I must admit that one of the beauties of GS is that it shows
readily its nonsense for a searcher who is not utterly naïve. A few months ago a
reporter from *The Scientist* called to find out what my problem was with GS. The first
search I did was to locate articles from *The Scientist*. The first item GS regaled us with
was purportedly an article about molecular biology in *The Scientist*, cited 7,390 times.
It looks impressive for the "ignoranti" but suspiciously high for searchers with
common sense. The citations were received for a substantial article in the *Journal of
Crystallography*, not the quarter-page news item in *The Scientist* as GS claimed – see
the story at www2.hawaii.edu/ ~ jacso/extra/gs/. This impromptu search made my
message clear to *The Scientist* reporter. I was not so successful with the reporter of
Nature, who just did not seem to understand my illustrative examples when he
questioned my criticism of GS, even if my examples were as much tailor-made for
Nature as bespoke suits by Savile Row tailors for the ultra rich. It may be my inability,
but Stevan Harnad (www.ecs.soton.ac.uk/ ~ harnad/), who knows a great deal about
cognition as Canada Research Chair in Cognitive Science and Founder and Editor of
Behavioral and Brain Sciences, voiced a similarly frustrating experience in the
American Scientist Forum (www.ecs.soton.ac.uk/ ~ harnad/Hypermail/Amsci/0883.
html) about open access policy, being unable to get the message through to the senior
reporter of *Nature* (Figure 8).

Figure 6.
The home page of an article published in 2006

Back to the fresh citing references from 2006. The third hit, with its four citing references, may seem especially promising, as it is a social science journal, and it discusses what is hard to follow for non-scholars (beyond citation searching, that is). But wait, it is a publication from 1985, not 2006. What about the next one? It is even older, published in 1980. GS just cannot deliver what it promises, 15 citing references for an article published in *European Review* in 2006. It cannot retrieve a single one, let alone 15. Did I say fool's gold? (Figure 9)

These errors of artificial unintelligence in matching cited and citing references one hopes will be noted by the natural intelligence of real scholars and practitioners who are competent in citation indexing. Currently, however, their voices are drowned out by too many authors who discovered the value of citation searching (evangelized by Eugene Garfield for 50 years), and in their neophyte enthusiasm embrace the free GS and pontificate about virtues it does not have – of course, GS does have some virtues which I discussed in this column a year ago (Jacso, 2005a, b).

For genuine scholars GS cannot "substitute for ISI" as Pauly and Stergiou claim after analyzing the citation counts for 114 articles in GS and WoS. They handpicked 99 articles in 11 disciplines from three authors with articles of high, medium and low citation counts. They then added 15 of the highly cited authors. Considering that WoS has about 37.5 million records for articles, and GS may have also have a large (but of indeterminate size) collection of records for all types of materials posted on the web, this is a tiny sample.
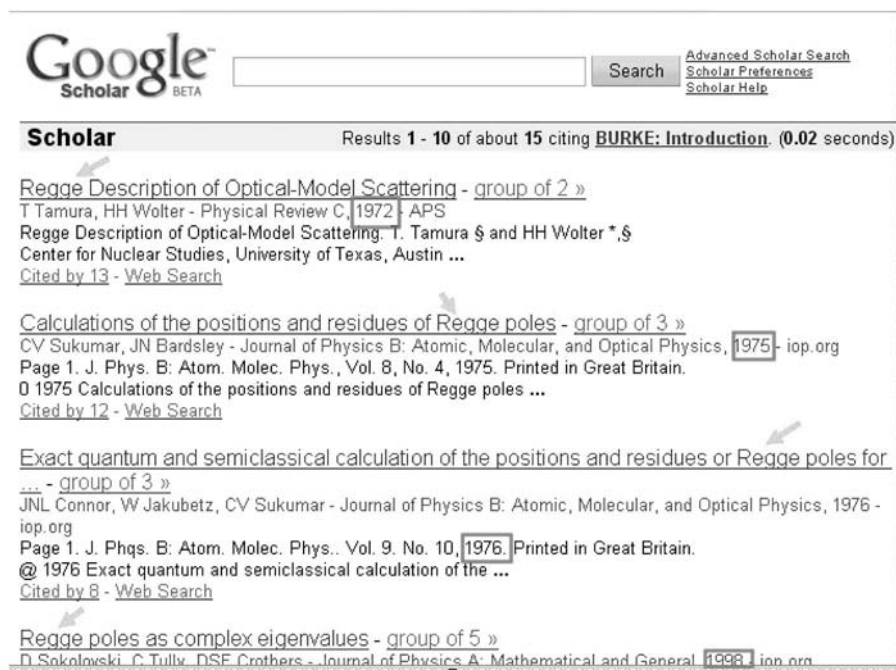
The 114 articles from the 75 unique journals in the sample were published between 1925 and 2004, and garnered from 1 to 300,000 citations. The authors claim this range to be from 1 to 100,000, but they base their highest citation count on a more than 20-year-old essay written by Eugene Garfield in 1984. There is a note in the spreadsheet, but this should have been mentioned in the article.

Apparently the authors are unaware of the fact that the citations to Lowry's landmark article trebled since the 1984 survey. A little search on the web (or contacting a citation indexing specialist) could have revealed a much more current value for the citations received by this paper. Overall, the authors should not have limited their readings about GS and WoS to Richard K. Belew, and Duncan Butler, the senior reporter at *Nature* whom they cite four times in a single paragraph. (I trust that they realize that this still counts only as a single citation for citation indexing purposes even in GS.)

306

The Right to Intervene
Llyod N. Cutler
From *Foreign Affairs*, Fall 1985

Article preview: first 500 of 6,130 words total.

📄 Purchase Article   🖨 Print   ✉ Email to a Colleague

Summary: We face many foreign policy decisions?how to respond to the fighting in Afghanistan, Lebanon, Nicaragua, Salvador, Angola, Kampuchea, the Philippines and soon, perhaps, South Africa?that involve the legality of intervening in a civil war. The international law journals are full of scholarly discussions on this subject. They are hard for non-scholars to follow. They disagree sharply, as scholars are wont to do, in their argumentation and conclusions. For readers who are not scholars of international law, this article tries to explain how the rules have evolved, where they now stand, and how they might be clarified to relieve the rising tension between the principle of nonintervention and the human rights of self-determination and open democratic elections.

Figure 9.
An article which GS
claims to have been
published in 2006

There are a number of articles by information professionals who paint a more realistic picture of the comparative values of the two systems. Some of them are listed in the bibliography for this column.

Pauly and Stergiou also managed to select the second and third most highly cited articles in the history of scientific publishing, one published in 1970 by Laemmli in *Nature*, the other published in 1976 by Bradford in *Analytic Biochemistry*. They registered both of these in their spreadsheet with a 65,535-citation count. For a reality check, my relatively simple test search (limited to the period from 1974 to 2005, and not trying very hard to herd the thousands of inaccurate citations) yielded 196,045 citations for Laemmli, and 118,671 for Bradford. The WoS counter is 16-bit and therefore, stops counting at the 65,536th citation. WoS warns you in all the four cases (out of the 37.5 million records) when this happens, by reporting the "cited by" value as >65,535. Selecting these extreme outliers for this test is akin to choosing the Himalaya for a first climbing expedition by socialites from Southern California. The authors should not have gone for the extreme.

It would have also deserved a note as to why the authors have chosen two articles for the Economics discipline from *Fisheries Research* and *Marine Policy*. To their credit, the only other social science discipline, Psychology, had no articles from fisheries journals.

Most important, it is unfair to compare the automatically reported citation counts of WoS with those of GS without corroboration. Corroborating citation counts is relatively easy in WoS by clicking on the cited references button and checking the one-liner, structured cited references which are sorted alphabetically by author – no matter in what order they appear in the original documents.

On the other hand, corroborating the citation counts reported by GS is very cumbersome for several reasons. One is that a significant portion of the purportedly citing documents is not available digitally in full-text format for many of the researchers who do bibliometric analysis. Even if some of them are available, one constantly needs to switch between databases and jump through hoops to get to the documents, as was the case with the archive of the Cambridge University Press in my tests mentioned earlier.

Often the right version for the researcher is not the first or second among the link options listed by GS for the full text. When testing for the articles citing Tom Wilson's paper, the primary links were often to Ingenta Connect, which no longer carries the journals of the large publisher, Sage, which is now hosted by HighWire Press. Locating the purportedly citing documents once they are accessed can be cumbersome and frustrating. In one case an entire issue of the 122-page *IFLA* journal was in a single pdf file. Repeated efforts to find the article purportedly citing Don Hawkins' article in *Information Research* proved to be fruitless.

The phantom and false references cannot be easily spotted. If you know something about the coverage of disciplines by GS through many hours of testing, you might spot the most suspiciously high citation counts. Rather than taking at face value the large difference between the citation counts for a journal article published in 2001 in *Media Psychology* reported by WoS as seven and by GS as 90 at the time of the "testing" by Pauly and Stergiou, they should have probed at least some of the purportedly citing documents.

They would have found that about 80 per cent of the citations in GS are to a chapter in a book published in 1994, and not the 2001 article they compared. Had they checked some of the purportedly citing references, they could have seen this anomaly, which is certainly not limited to this article in their sample, and which certainly distorts their results. GS shows the snippets only for the first 100 "citing" documents, and from then on it is a question of faith.

GS is good for locating relevant items, leading users some of the time to an open access version of a document, but it is not an appropriate tool for bibliometric studies. It plays fast and loose, (make that too fast and too loose), with its hit counts and citation counts to allow fair comparisons without tiresome verification, as both numbers are too inflated. There is not yet a credibility gap factor for the GS counts as there have been for politicians in adjusting the real value of their campaign statements.

The combination of the inflated citation count values dispensed by GS with the ignorance and shallowness of some GS enthusiasts can be a lethal mix for real scholars who might end up with funny money when their real money budget is cut for many of the scholarly databases, especially the ones enhanced by cited references, such as the ones hosted by CSA of the Cambridge Information Group.

Considering the authors' very limited test, flawed results based partly on phantom citations and inflated citation counts, it is especially disturbing to read their suggestion that "GS can substitute for WoS" and their claims that the two systems "lead to essentially the same results, in spite of their different methods for identifying citing sources". What Pauly and Stergiou did not say is that the total number of citation counts for their small sample is 409,220 in WoS, and 171,519 in GS, even without discounting the GS counts for the inflated and phantom citation counts, and without crediting WoS for retrieving (through the searchers' cooperation) the non-exact but real matches.

GS is unable to its maths, cannot stick to its numbers, and often cannot tell a publication year apart from a PO box or a page number, a chapter title from an author name, or a journal name from a publisher name. GS does not even seem to try to create combined author, journal name and publication year searches for refining its matching process, or applying plausibility tests. Testing for the condition that the citing year must be equal to or larger than the publication year of the cited document could filter out a large number of false matches. But Boolean logic is not the forte of GS.

It is quite alarming that the authors suggest that their perceived equivalence of the two systems "have strong implications for future citation analyses, and the many promotion, tenure and funding decisions based thereon". In academia, referring to citation counts presented by GS may embarrass many scholars if they are asked to show the citing references, and then they realize the phantom citations.

This is the place to come back to Gorman's editorial in this journal (Gorman, 2005) and to his question "How do we count our chickens? Or do citation counts count?" My answer is that, if they come from GS, whose results cannot be corroborated, they do not count. As for WoS, its source base needs to expand to structured, well-organized, metadata-enriched preprint and reprint archives, as well as to many more high-quality open access journals and conference proceedings which do not even exist in print, such as *D-Lib Magazine*, *First Monday* and *Ariadne*, to name a few sources from our fields of specialization which are not covered by WoS. It should also add title and abstract search to the cited reference search options and lead the users there by a click, passing a smartly constructed query for searching the huge index of cited works more swiftly after some adjustment by the searcher.

References

Gorman, G.E. (2005), "How do we count our chickens? Or do citation counts count?", *Online Information Review*, Vol. 29 No. 6, p. 581.

Jacso, P. (2005a), "As we may search: comparison of major features of the Web of Science, Scopus and Google Scholar citation-based and citation-enhanced databases", *Current Science*, Vol. 89 No. 9, p. 1537, available at: www.ias.ac.in/currsci/nov102005/1537.pdf

Jacso, P. (2005b), "Comparison and analysis of the citedness scores in Web of Science and Google Scholar", *Proceedings of Digital Libraries: Implementing Strategies and Sharing Experiences: 8th International Conference on Asian Digital Libraries, ICADL, Bangkok, Thailand, 12-15 December*, p. 360.

Jacso, P. (2006), "Dubious hit counts and cuckoo's eggs", *Online Information Review*, Vol. 30 No. 2, pp. 188-93.

Further reading

Bauer, K. and Bakkalbasi, N. (2205), "An examination of citation counts in a new scholarly communication environment", *D-Lib Magazine*, Vol. 11 No. 9, available at: www.dlib.org/dlib/september05/bauer.html

Butler, D. (2004), "Science searches shift up a gear as Google starts Scholar Engine", *Nature*, Vol. 432, p. 423.

Deis, L.F. and Goodman, D. (2005), "Web of Science (Version 204) and Scopus", *Charleston Advisor*, Vol. 6 No. 3, available at: www.charlestoncocom/compcfm?id = 43

Gardner, S. and Eng, S. (2005), "Gaga over Google? Scholar in the social sciences", *Library Hi Tech News*, Vol. 22 No. 8, p. 42.

Giles, J. (2005), "Science in the web age: start your engines", *Nature*, pp. 438-554, available at: www.nature.com/nature/journal/v438/n7068/full/438554a.html

Jacso, P. (2005c), "Google Scholar: the pros and the cons", *Online Information Review*, Vol. 29 No. 2, p. 208, available at: www.emeraldinsight.com/Insight/viewPDF.jsp?Filename = html/ Output/Published/EmeraldFullTextArticle/Pdf/2640290206.pdf

Kennedy, S. and Price, G. (2004), "Big news: Google Scholar is born", *ResourceShelf*, Vol. 18, 18 November, available at: www.resourceshelf.com/2004/11/wow-its-google-scholar.html

Myhill, M. (2005), "Google Scholar – a review", *The Charleston Advisor*, Vol. 6 No. 4, available at: www.charlestoncocom/reviewcfm?id = 225

Noruzi, A. (2005), "Google Scholar: the new generation of citation indexes", *LIBRI*, Vol. 55 No. 4, available at: www.librijournal.org/2005-4toc.html

Notess, G. (2005), "Scholarly web searching: Google Scholar and Scirus", *Online*, Vol. 29 No. 4, p. 39, available at: www.infotoday.com/online/jul05/OnTheNet.shtml

Pauly, D. and Stergiou, K.I. (2005), "Equivalence of results from two citation analyses: Thomson ISI's Citation Index and Google's Scholar service", *Ethics in Science and Environmental Politics*, p. 33, available at: www.int-res.com/articles/esep/2005/E65.pdf

Vine, R. (2006), "Google Scholar", *Journal of the Medical Library Association*, Vol. 94 No. 1, p. 97, available: www.pubmedcentral.nih.gov/articlerender.fcgi?artid = 1324783

Corresponding author
Peter Jacso can be contacted at: jacso@hawaii.edu