Accuracy and Completeness of Publication and Citation Records in the Web of Science, PsycINFO, and Google Scholar: A Case Study for the Computation of *h* Indices in Psychology

Miguel A. García-Pérez
Departamento de Metodología, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain. E-mail: miguel@psi.ucm.es

Hirsch's *h* index is becoming the standard measure of an individual's research accomplishments. The aggregation of individuals' measures is also the basis for global measures at institutional or national levels. To investigate whether the *h* index can be reliably computed through alternative sources of citation records, the Web of Science (WoS), PsycINFO and Google Scholar (GS) were used to collect citation records for known publications of four Spanish psychologists. Compared with WoS, PsycINFO included a larger percentage of publication records, whereas GS outperformed WoS and PsycINFO in this respect. Compared with WoS, PsycINFO retrieved a larger number of citations in unique areas of psychology, but it retrieved a smaller number of citations in areas that are close to statistics or the neurosciences, whereas GS retrieved the largest numbers of citations in all cases. Incorrect citations were scarce in WoS (0.3%), more prevalent in PsycINFO (1.1%), and overwhelming in GS (16.5%). All platforms retrieved unique citations, the largest set coming from GS. WoS and PsycINFO cover distinct areas of psychology unevenly, thus applying different penalties on the *h* index of researches working in different fields. Obtaining fair and accurate *h* indices required the union of citations retrieved by all three platforms.

Numerous indices of scientific achievement have been developed (Lehmann, Jackson, & Lautrup, 2008), but not all of them are valid for all purposes. For instance, journal impact factors are no longer deemed adequate as measures of an individual's achievements (Cameron, 2005; Campbell, 2008; Colquhoun, 2003; Garfield, 1999; Lawrence, 2007; Moed, 2002; *Nature*, 2005a, 2005b; Rogers, 2002; Seglen, 1997a, 1997b; Simons, 2008; Smith, 1998; Waheed, 2003; Williams, 1998; Wróblewski, 2008). In fact, the validity of journal impact factors for any purpose whatsoever has

recently been questioned because the method by which they are computed is neither transparent nor reproducible (Brumback, 2008a, 2008b; Carrió, 2008; Hernán, 2008; Joseph & Hoey, 1999; Porta & Álvarez-Dardet, 2008; Rogers, 2002; Rossner, van Epps, & Hill, 2007, 2008; Seglen, 1997b; The *PLoS Medicine* Editors, 2006; Wilcox, 2008) and also because of the ease with which journal impact factors can be manipulated (Agrawal, 2005; Brumback, 2009; Cameron, 2005; Della Sala & Brooks, 2008; Falagas & Alexiou, 2007, 2008; Lavie, 2009; Opatrný, 2008; Reedijk & Moed, 2008; Schutte & Švec, 2007; Seglen, 1997b; Sevinc, 2004; Topo Universitario, 2008; van Diest, Holzel, Burnett, & Crocker, 2001; Yu & Wang, 2007).

For measuring an individual's research accomplishments, Hirsch's (2005) *h* index has swiftly gained popularity, perhaps because it directly measures the impact of an author's papers instead of the impact of the journals in which they came out. Research on (and application of) the *h* index has been extensive (see Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2009; Bornmann & Daniel, 2009; for an up-to-date list of publications on the topic, see http://sci2s.ugr.es/hindex/biblio.php). Accurate computation of an individual's *h* requires an exhaustive source of publication records and an also exhaustive and errorless source of citation records.

The number of sources that provide publication and citation records has increased in the past few years (Ballard & Henry, 2006; Neuhaus & Daniel, 2008; Roth, 2005). Some of these are free tools, such as Google Scholar (GS; http://scholar.google.com) or Publish or Perish (http://www.harzing.com/pop.htm), a software tool that analyzes raw citations retrieved from GS and is often complemented with the "clean-up" utility CleanPoP (http://cleanpop.ifris.net). Other platforms are provided by for-profit organizations covering either the broad scientific arena—such as Thomson Reuter's *Web of*

TABLE 1.   Description of the research outcomes of each individual in the study sample.

	Bajo, M.T.	García-Pérez, M.A.	Salgado, J.F.	Vázquez, C.
Academic age[a]	24	22	26	27
Journal articles[b]	65	66	58	82
In Spanish or in Spanish journals	25	4	24	52
In English in international journals	40	62	34	30
Books (as author)[c]	2	0	2	8
Textbooks	1	–	0	4
Reference books	1	–	2	4
Books (as editor)	0	0	0	4
Textbooks	–	–	–	2
Reference books	–	–	–	2
Book chapters[d]	13	6	22	68
In textbooks	4	0	4	39
In reference books	9	6	18	29

[a]Academic age is defined here as the number of years elapsed since publication of the author's third paper.
[b]Eight of these 271 journal articles were in press when this research was carried out.
[c]One of these 12 books was in press when this research was carried out.
[d]Seven of these 109 book chapters were in press when this research was carried out.

Science (WoS; http://thomsonreuters.com) and Elsevier's *Scopus* (http://www.scopus.com)—or specific research fields such as American Psychological Association's *PsycINFO* (http://www.apa.org/psycinfo).

These platforms differ in their characteristics (Falagas, Pitsouni, Malietzis, & Pappas, 2008; Harzing & van der Wal, 2008; Jacsó, 2005a; Schroeder, 2007) and journal coverage (Glavel & Iselid, 2008) and, hence, in their limitations and in the number of publication and citations records that they include (Bakkalbasi, Bauer, Glover, & Wang, 2006; Baneyx, 2008; Bar-Ilan, 2006; Bauer & Bakkalbasi, 2005; Bornmann et al., 2009; Jacsó, 2005b; Kousha & Thelwall, 2008; Kulkarni, Aziz, Shams, & Busse, 2009; Levine-Clark & Gil 2009a, 2009b; Meho & Rogers, 2008; Meho & Yang, 2007; Neuhaus, Neuhaus, Asher, & Wrede, 2006; Norris & Oppenheim, 2007; Pauly & Stergiou, 2005; Shultz, 2007; Vaughan & Shaw, 2008; Vieira & Gomes, 2009; Walters, 2007, 2009; Whitley, 2002). These differences also have an effect on the *h* indices calculated from citation records retrieved through them (Bar-Ilan, 2008; Jacsó, 2008a, 2008b, 2008c, 2008d; Sanderson, 2008).

The studies just mentioned rendered divergent results. Which database fared better with respect to the number of publication or citation records varied with discipline but also with the time in which the study was carried out and the publication year of the papers under study. Perhaps the only conclusion of these studies is that publication and citation records provided by WoS can often be complemented, sometimes substantially. Thus, the work reported here aimed at investigating how best the available databases can be used to collect the largest number of valid citations for known publications.

If the *h* index or some of its variants (see Bornmann, Mutz, & Daniel, 2008) is to accompany other measures of an individual's research impact (Bornmann, Mutz, Neuhaus, & Daniel, 2008; Haeffner-Cavaillon & Graillot-Gak, 2009; Harnad, 2009), then it is foreseeable that researchers will be asked to report it in their CV. For the computation of one's own *h*, the accuracy with which databases include publication records is not critical, provided that the records are to be found there in one way or another (e.g., by article title in case that the author's name was misspelled in the record). Yet, the accuracy and completeness of citation records are crucial because there is no alternative way in which authors can find out how many times each of their papers has been cited. Furthermore, if researchers' own *h* as reported in their CV is expected to be backed up on demand by lists of citing papers, platforms including incorrect citation information are useless. The work described here investigated the concordance, accuracy, and validity of citation records retrieved through three platforms (WoS, PsycINFO, and GS) for each of four well-published Spanish psychologists whose complete list of publications was available.

Method

Sample Individuals

Four Spanish psychologists (including the author) were selected for study. These individuals have a scientific age around 25 years and each has produced more than three publications per year on average (see Table 1). They have an *h* in the vicinity of 10 according to automatic computation using WoS (García-Pérez, 2009a; Salgado & Páez, 2007) and they carry out their research in different areas of psychology. These individuals provided complete lists of publications.

Platforms, Databases, and Dates of Retrieval

Three platforms were used: WoS, PsycINFO, and GS. Of these, GS and WoS are well-known multidisciplinary platforms (Falagas et al., 2008; Jacsó, 2008b, 2008c); PsycINFO, on the other hand, is provided by the American Psychological Association on a subscription basis and covers the psychological literature in several languages since 1890

TABLE 2. Search pattern that renders the most complete list of publications (in Web of Science) for each individual.

Bajo, M.T.	**AU**=(bajo mt or bajo t)
	NOT (SO=international journal of psychology **AND PY**=2000)
	NOT (SO=bulletin of the psychonomic society **AND PY**=1985)
García-Pérez, M.A.	**AU**=(garcia-perez m **OR** garciaperez ma **OR** garcia-perez ma)
	AND OG=(univ complutense **OR** univ complutense madrid **OR** harvard univ **OR** univ pais vasco **OR** virginia polytech inst state univ **OR** univ madrid **OR** univ murcia)
	NOT SO=(molecular* **OR** blood **OR** clinical* **OR** immunology **OR** perception **OR** investigative* **OR** journal of math*)
	NOT (SO=spanish journal of psychology **AND PY**=2007)
Salgado, J.F.	**AU**=(salgado j **OR** salgado jf)
	AND OG=(univ santiago de compostela **OR** univ santiago **OR** univ santiago compostela)
	NOT SO=(physic* **OR** nuclear* **OR** journal of chemical* **OR** journal of physics* **OR** analytical* **OR** journal of thermal* **OR** Brazilian* **OR** engineering* **OR** revista espanola de cardiologia **OR** fluid* **OR** thermo* **OR** geochimica*)
	NOT (AU=ones ds)
	OR (AU=salgado je **AND PY**=(1996 **OR** 2004))
	OR (AU=salgado jdf **AND PY**=2001)
	OR (AU=spector pe **AND PY**=2002 **AND SO**=academy of management journal)
Vázquez, C.	**AU**=(vazquez c **OR** valverde cv)
	AND OG=(univ complutense madrid **OR** univ complutense **OR** ucm **OR** univ madrid)
	NOT AU=(de las heras **OR** errasquin **OR** fontan **OR** huedo **OR** patino)
	NOT SO=(fems* **OR** canadian journal of micro* **OR** journal of the science* **OR** myco* **OR** chemo* **OR** microbio* **OR** international journal of food* **OR** food* **OR** european journal of plant* **OR** journal of applied microbiology **OR** phyto* **OR** letters in applied micro* **OR** journal of food* **OR** journal of microbio* **OR** siam* **OR** systematic* **OR** journal of pediatric* **OR** transactions of the british myco* **OR** archives of bio* **OR** medicina clinica **OR** international journal of psychology)
	OR (AU=vazquez c **AND SO**=psychological bulletin)

Note. For clarity, indentation is used and keywords and operators are shown in bold uppercase, whereas values are shown in lowercase.

(see http://www.apa.org/psycinfo for details). Of the 2,438 journals included as of October 2009, 1,403 are reportedly indexed cover to cover, whereas the rest are examined for psychological content and only the articles considered relevant by PsycINFO staff are indexed.

Our WoS subscription includes Science Citation Index Expanded 1899–present, Social Sciences Citation Index 1956-present, Arts & Humanities Citation Index 1975-present, and the two Conference Proceedings Citation Indexes (Science 1990-present and Social Science & Humanities 1990-present), updated on November 21, 2009. PsycINFO was accessed on its October 2009 update through the CSA Illumina interface (http://www.csa.com). Records were retrieved on November 23–24, 2009.

Search for Publication and Citation Records

Mimicking what individual researchers would do to collect citation records for their papers, our search used whichever reasonable means seemed necessary to retrieve the most complete list of citing papers from each database. This required slightly different strategies across platforms.

In WoS, the "Author Finder" interface was first used as described by García-Pérez (2009a; see also García-Pérez, 2001) to obtain a working list of publications. The CV of the individual was next used for two purposes. One was to remove from this list the papers by other researchers with the same name; the other was to search manually for papers that did not turn up in the working list. The latter was usually accomplished by retrieving the entire list of WoS records for papers published in the applicable journal in the year of

concern. This strategy retrieved records that had errors, but, more generally, it simply revealed that a record for that paper did not exist in WoS. The one-shot query that renders the most complete list of publications for the individuals in our sample is listed in Table 2. When entered into the query box of the "Advance Search" interface in WoS, those queries retrieve the same set of publication records as the three-step strategy just described (as of the date in which this work was carried out). Some of the known papers for which no publication record was found in WoS could nevertheless be found elsewhere in the Web of Knowledge (WoK; of which WoS is just a part). Specifically, the "All Databases" tab next to "Web of Science" in the opening menu of WoK provides a search interface that gains access to publication records (accompanied by citation records) that are unreachable from within WoS.[1] This route retrieved six additional publication records through the queries in Table 3.

In PsycINFO, the author index was first eyeballed to create a query that included all variants in which author names could have possibly been indexed. Publication records were then retrieved using the "Command Search" interface (under the "Search Tools" tab) in PsycINFO with a query that included all these name variants, which rendered an initial working list of publications that was subsequently refined and expanded as described for WoS. The one-shot query that rendered the

---

[1]In our subscription, the set of "All Databases" comprises (besides WoS) *Current Contents Connect* (1998–present), *Derwent Innovations Index* (1980–present), *Biological Abstracts* (1926–present), *Inspec* (1969–present), MEDLINE (1950–present), and *Journal Citation Reports* (1997–2008). We did not investigate in which of these databases were those additional records included.

TABLE 3. Search patterns that retrieve additional publication records under "all databases" in Web of Knowledge for each of the individuals for whom this strategy was useful.

García-Pérez, M.A.	**AU**=(garcia-perez) **AND SO**=(spatial vision) **AND PY**=(1988 **OR** 1992) **OR** **AU**=(garcia-perez) **AND SO**=(span j psychol) **AND PY**=(1999 **OR** 2001) **OR** **AU**=(garcia-perez) **AND SO**=(communications in statistics*) **AND PY**=(2009)
Vázquez, C.	**AU**=(vazquez) **AND SO**=(psiquis) **AND PY**=(1989)

Note. For clarity, keywords and operators are shown in bold uppercase, whereas values are shown in lowercase.

TABLE 4. Search pattern that renders the most complete list of publications (in PsycINFO) for each individual.

Bajo, M.T.	**AU**=(molina teresa bajo **OR** bajo maria teresa **OR** bajo maria-teresa **OR** bajo mariá teresa **OR** bajo mary-teresa **OR** bajo maría t **OR** bajo maría teresa **OR** bajo maría-teresa **OR** bajo teresa **OR** bajo m - teresa **OR** bajo m t **OR** bajo m teresa **OR** bajo ma teresa **OR** bajo maria t) **NOT AN**=(1987-54357-001) **OR (JN**=journal of cross-cultural psychology **AND TI**=Love and power*)
García-Pérez, M.A.	**AU**=(garcía-pérez miguel **OR** garcía-perez m a **OR** garcía-pérez miguel a **OR** garcía-pérez miguel Ángel **OR** garcia-perez miguel a **OR** garcia-perez miguel angel **OR** garcia-perez m a **OR** garcía pérez miguel) **NOT JN**=(a n a e approche neuropsychologique*) **OR (JN**=spatial vision **AND TI**=a comparison of fixed*)
Salgado, J.F.	**AU**=(salgado jesus f **OR** salgado jesus **OR** salgado jesús f **OR** salgado jesús **OR** salgado velo jesús f **OR** salgado velo jesús **OR** salgado-velo j) **OR (JN**=applied* **AND TI**=(an international study* **OR** the pitfalls of poor*))
Vázquez, C.	**AU**=(vazquez carmelo **OR** vázquez-valverde c **OR** vázquez-valverde carmelo **OR** vázquez valverde c **OR** vázquez c **OR** vázquez carmelo **OR** valverde carmelo vázquez) **NOT JN**=(mathematical social sciences **OR** neuropharmacology) **NOT IS**=(1130-9512 **OR** 0214-7599 **OR** 0361-9230 **OR** 0031-9384 **OR** 0091-3057) **OR (JN**=acta psychiatrica scandinavica **AND PY**=1989 **AND TI**=clinical implications*) **OR (JN**=psiquis **AND PY**=1989 **AND TI**=prediccion de la respuesta*)

Note. For clarity, indentation is used and keywords and operators are shown in bold uppercase whereas values are shown in lowercase.

most complete list of publications for the individuals in our sample is given in Table 4.

As for GS, its "Advanced Scholar Search" interface is rather unsophisticated and less than user-friendly. In addition, GS extracts authors' names from digital documents very poorly (Jacsó, 2008a, 2008b, 2008e). Then, publication records were retrieved by searching for major unique word sequences ("exact phrase") "in the title of the article" and by using the author's name only when it proved useful. When a paper that was known to have appeared in journal J in year Y could not be retrieved with this method, all papers published in journal J in year Y were searched for. This strategy usually retrieved a record for the target paper; when it did not (or when the target publication was a book or a book chapter), no further search was attempted.

Lists of citing articles accompanying publication records retrieved through each platform were stored for post hoc analyses. Confirmation that the publication was actually cited (a process that we refer to as "authentication") was sought by inspecting the purportedly citing articles.

Results

Completeness of Publication Records

Records for the vast majority (321/380, or 84.5%) of the target publications could actually be located through GS. Yet, in most cases, the search did not retrieve actual records but mere links to (a) personal Web pages, where documents were posted by their authors, (b) the download pages of publishers or suppliers of full-text articles, or (c) Web pages where records are listed (e.g., the British Library *BL Direct* service at http://direct.bl.uk or a variety of catalogs, repositories, book sellers, or Google Books). Most of the items not found in GS (44/59, or 74.6%) were Spanish textbooks and old articles in either discontinued journals or active journals whose back files have never been posted on the Web. The remaining 15 items are most likely somewhere on the Web too, but finding them seemed a task of formidable dimensions with the crippled "Advance Scholar Search" tool.

In WoS, the queries in Tables 2 and 3 retrieved 60.1% (158/263) of the journal articles in our sample,[2] and two of the 62 chapters in reference books (see Table 1). This low hit rate attests to a known bias against social sciences and non-English sources (Nederhof, 2006; Norris & Oppenheim, 2007). Other target papers could not be retrieved because the international journals in which they appeared started to be covered by WoS only later. For instance, *Spatial Vision* is covered only from Volume 8, 1993 (although it is available in WoK from Volume 1, 1985). Considering the material that WoS actually covers, its hit rate is 98.1% (152/155), and our explicit search for three papers that resisted retrieval revealed

[2]Recall, however, that publication records for six of these 158 journal articles were retrieved through the "All Databases" search in WoK and not through WoS.

TABLE 5.   Overlap of publication records across platforms.

		Records included in WoS					Records not included in WoS		
		Google Scholar					Google Scholar		
		included	not included				included	not included	
PsycINFO	included	130 | 130	1 | 1	131 | 131		PsucINFO — included	58 | 68	5 | 5	63 | 73
	not included	26 | 27	1 | 2	27 | 29		not included	36 | 96	6 | 51	42 | 147
		156 | 157	2 | 3	158 | 160			94 | 164	11 | 56	105 | 220

Note. WoS = Web of Science.

Numerals on the left of each cell or margin indicate counts of journal articles only; numerals on the right indicate counts of all publications.

errors of omission: For instance, nearly 73% (11/15) of the papers published in Volume 38, No. 3, 2009 of *Communications in Statistics – Simulation and Computation* are nowhere to be found in WoS.

Evaluating PsycINFO in this way is difficult because of its discretionality and also because its yearly coverage is not readily available. Even with those uncertain boundaries, PsycINFO coverage statement creates expectations if only because authors can legitimately consider that their papers satisfy whichever requirements for "psychological content" PsycINFO staff may use.

Consider first the case of book chapters. The researchers in our sample have published 109 books (see Table 1), 32 of them in English-language reference books published by major international publishers (a type of book reportedly covered by PsycINFO). Five of these were in press when this research was carried out, and only 10 of the remaining 27 (or about 37%) were actually included in PsycINFO. As for journal articles, our queries retrieved 73.8% (194/263) of them, a figure that is meaningfully larger than the 60.1% reported earlier for WoS. Establishing the denominator for a figure of completeness within the intended coverage of PsycINFO is hampered by lack of precise coverage information. There are, however, some signs that PsycINFO does not include all the items that it should. For instance, none of the seven papers that came out in Volume 13, No. 4, 2000 of *Spatial Vision* are included in PsycINFO; similarly, 25% (5/20) of the papers that came out in Volume 29, No. 3, 1997 of *Behavior Research Methods, Instruments, & Computers* are missing, as are 16% (4/25) of the papers that came out in Volume 8, 2000 of the *International Journal of Selection and Assessment*. These journals are reportedly indexed cover to cover (http://www.apa.org/psycinfo/covfullselect.html). As for journals indexed discretionally, it is striking that all articles lately published in *Psicologica* (http://www.uv.es/psicologica) are indexed with the exception of all the articles that came out in Volume 28, No. 2, 2007. Besides the 194 articles in our sample that were found in PsycINFO, we counted 20 more articles that should have been there according to the coverage statement or because of unquestionable psychological content. Then, the coverage of PsycINFO amounts to 90.7% (194/214).

In sum, WoS includes publication records for almost all (98.1%) of the journal articles that it nominally covers, whereas PsycINFO includes records for slightly less (90.7%) of the journal articles that users can expect it to cover. When all journal articles are considered (whether nominally covered or not), the hit rates for WoS and PsycINFO decrease to 60.1% and 73.8%, respectively, and reflect a larger coverage in PsycINFO. On the other hand, GS retrieves records for a larger percentage (84.5%) of the entire set of publications in our sample, including books and book chapters barely covered by PsycINFO and not covered by WoS.

Overlap of Publication Records Across Platforms

It is interesting that PsycINFO retrieved publication records for a percentage of journal articles (73.8%) that is meaningfully larger than that retrieved by WoS + WoK (60.1%). Nevertheless, the two sets (i.e., the $152 + 6 = 158$ records retrieved by WoS + WoK and the 194 records retrieved by PsycINFO) comprised only 131 common papers (see Table 5). Fifty-three (or about 84%) of the 63 journal articles included in PsycINFO and not in WoS had been published in Spanish journals and as far back as the early 1980s, which effectively adds a significant source of unique publication (and citation) records in psychological research. Yet, these extras are not provided systematically, owing to the discretionality of PsycINFO. On the other hand, most of the 27 journal articles included in WoS and not in PsycINFO represented research published in major international journals which PsycINFO does not index cover to cover or has started to do so only recently (e.g., *NeuroReport*) or which it barely or not at all indexes because of their substantial amount of non-psychological content (e.g., *Journal of Applied Statistics*).

Figure 1a shows the overlap of the three databases as to journal articles. Records for 49.4% of them (130/263) were found in all platforms. Across the two commercial platforms, PsycINFO included 63 unique records (24.0%) and WoS included less than half as many (27, or 10.3%). GS, on its side, virtually engulfed all the publication records included in either WoS or PsycINFO and 85.7% of those included in neither of them. GS misses only 13 of 263 (4.9%) papers,
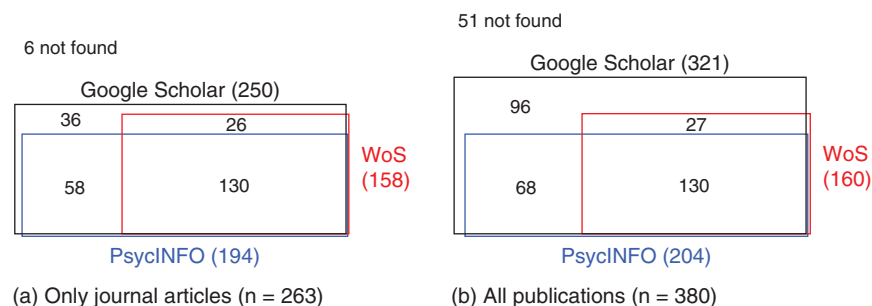
FIG. 1. Overlap of publication records in Web of Science (red), PsycINFO (blue), and Google Scholar (black). Inset numerals indicate the number of common papers in the applicable region. Full data are given in Table 5. (a) Overlap for journal articles. (b) Overlap for all publications, including books and book chapters.

six of which are not to be found in any other database either, whereas five can be found only in PsycINFO, one can be found only in WoS, and one more can be found in both PsycINFO and WoS. If books and book chapters are also considered (Figure 1b), GS continues to gain access to almost all of the publications that can be accessed through WoS or PsycINFO. In sum, GS gains access to a larger overall number of publication records than any of the two other platforms, particularly when book and book chapters are considered.

Accuracy of Publication Records in WoS and PsycINFO

Publication records were generally but not exquisitely accurate regarding article title, journal name, publication year, and volume and page numbers, but they were much less accurate in what regards authors' names. Unfortunately, this latter jeopardizes the success of a search by author name, and our results confirm earlier reports that this problem affects non-English names (Brown, 1999; Kotiaho, 1999; Kotiaho, Tomkins, & Simmons, 1999; Osca-Lluch, Civera Mollá, & Peñaranda Ortega, 2009; Price, 1998). It is also important to note that PsycINFO does not always include in its records names beyond the third or fourth author of a paper and, hence, a search by author name cannot possibly retrieve those records (which is the reason for the final "OR" clauses involving journal name and article title in the query for authors Bajo and Salgado in Table 4).

Validity of Citation Counts and Authentication of Citations

Publication records come with citation counts and lists of citing papers in all platforms. Anecdotal evidence encountered before this research started and reported in other papers (Jacsó, 2005a, 2006, 2008b; Kousha & Thelwall, 2008; Levine-Clark & Gil 2009b; Meho & Yang, 2007) revealed that raw citation counts are sometimes inflated. Across the board, WoS was least affected by this problem and GS was most affected by this problem, as described next.

WoS listed a raw number of 1,335 citing papers for 157 publications[3] across the individuals in our sample, and four

of those purported citations (0.3%) were in error. Of them, three were "phantom citations" (i.e., papers that did not actually cite the target paper; Jacsó, 2008a) and one reflected the presence of a duplicate, rambling record in the database (see Figure 2).

Analogously, PsycINFO listed a raw number of 1,247 citing papers for 201 publications,[4] and 14 of them (1.1%) were in error. Of these errors, eight were duplicate records for articles reprinted in another source (which are counted twice within the database, although the reprints are tagged as such in the records), three were phantom citations, one reflected also a duplicate record (this time for the new edition of a book that was already in the database), one was a mere miscount (the paper was tagged as "cited by 5" but the list of citing papers included only four entries), and the last one was much harder to understand: An announcement of appointed new editors turned up in a list of citing papers when not even the article ending on the page in which the announcement was inserted cited the target paper.

Finally, GS listed a raw number of 3,287 citing papers for 318 publications,[5] of which 541 (16.5%) were in error. Eleven of these errors were mere miscounts (e.g., an item reported to be "cited by 17" and accompanied by a list of only 15 citing papers), and the remaining errors had a very diverse nature as follows: 97 of the 541 errors did not actually supply any links that could help trace down the purported citation (the usual entry with a "[CITATION]" prefix); 57 of the remaining 433 errors were phantom citations, because either the target paper was not to be found in the reference list of the document that the link retrieved or the link did not retrieve any document with citations; 215 represented duplicate links that pointed to the same citing paper, although on a different site and, more often than not, in such a disguise that the duplicity could not have been reasonably anticipated (see Figure 3); 100 errors were also duplicates but in the form of reprints published in different sources (as discussed for PsycINFO) or in different languages (e.g., for journals that publish each paper in two or three languages, or for journals that publish translations

---

[3]Three publications with raw citation counts of 141, 196, and 270 in WoS were excluded.

[4]Three publications with raw citation counts of 145, 170, and 326 in PsycINFO were excluded.

[5]Three publications with raw citation counts of 190, 237, and 575 in GS were excluded.

Citing Articles

Title: **SEMANTIC FACILITATION WITH PICTURES AND WORDS**
Author(s): BAJO, MT
Source: **JOURNAL OF EXPERIMENTAL PSYCHOLOGY-LEARNING MEMORY AND COGNITION** Volume: 14 Issue: 4
Pages: 579-589 Published: OCT 1988
Citation Map

The above article has been cited by the articles listed below.
Note: The Times Cited count is calculated across all *Web of Science* editions. More information.

Results: **64**   Page 3 of 7 (Go)   Sort by: First author

(Print) (E-mail) (Add to Marked List) (Save to EndNote Web) (Save to EndNote, RefMan, ProCite) more options   Analyze Resu

☑ 21. Title: The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading
Author(s): Ganis G, Kutas M, Sereno MI
Source: **JOURNAL OF COGNITIVE NEUROSCIENCE** Volume: **8** Issue: **2** Pages: **89-106** Published: **MAR 1996**
Times Cited: **102**
UCM Otros recursos

☑ 22. Title: PICTURE NAMING
Author(s): GLASER WR
Source: **COGNITION** Special Issue: **Sp. Iss. SI** Pages: **61-105** Published: **1994**
Times Cited: **0**
UCM Otros recursos

☑ 23. Title: PICTURE NAMING
Author(s): GLASER WR
Source: **COGNITION** Volume: **42** Issue: **1-3** Pages: **61-105** Published: **MAR 1992**
Times Cited: **196**
UCM Otros recursos (Full Text)

☑ 24. Title: Understanding corporate logos: Lexical and analogical considerations

FIG. 2.   A duplicate, rambling record in Web of Science. The paper listed as #22 in this cutout of citing papers does not exist in the actual journal in 1994 and seems to imply the 1992 paper listed as #23.

FIG. 3.   Two different-looking links provided by Google Scholar that nevertheless point to the same true document. The link at the top (the implied URL has been copied at its bottom) retrieves a reprint of the journal article from one of the coauthors' Web site; the link at the bottom (the implied URL has been copied at its bottom) retrieves a preprint of the accepted version of the same journal article from the Web site of another coauthor.

of articles published in other journals); 58 errors took the form of links that pointed to course syllabi;[6] and three errors

---

[6]It might be argued that items referenced in course syllabi should not be counted off in an analysis of citations. Resolving this controversial issue is not critical at this point, given the purpose of the present article, but the author's personal position is that textbooks are not part of an individual's *research* output and, hence, that these citations in course syllabi should not be counted.

pointed to the CV posted on the Web page of a coauthor of one of our target authors, where papers are not cited in any realistic sense.

The foregoing data corroborate that citation counts in GS are seriously inflated, but they also reveal that citation counts in WoS and PsycINFO are not error free. Also, it seems that errors in GS are less prevalent today than they were a few years ago (compare with figures given by

Jacsó, 2006). In any event, care should always be exercised in interpreting raw citation counts, but it is understandable that the extra step of authentication is rarely taken: It took about 10 hours to retrieve and store publication records with their accompanying citation counts and lists of citing papers for all authors in our sample, and it took almost 70 hours for authentication.

Comparison of Authenticated Citation Counts Across Platforms

Figure 4 shows scatter plots of authenticated citation counts across databases. Visual inspection reveals that citation counts in PsycINFO are generally lower than in WoS (top panel), whereas citation counts in GS are generally higher than citation counts in either WoS (center panel) or PsycINFO (bottom panel). Some idiosyncracies are also apparent upon separate consideration of data points pertaining to different authors (colors and rotation angles). For instance, in the top panel, where data points are generally below the diagonal, blue crosses (for author Salgado) and gray crosses (for author Vázquez) lie more often at or above the diagonal than below it. This means that the implied publications (mostly journal articles) are cited more often in the PsycINFO database than in the WoS database. This characteristic seems related to the research fields of these authors (work and organizational psychology and clinical psychology), which are more broadly covered by PsycINFO than by WoS. In contrast, green crosses (for author García-Pérez) lie well below the diagonal and red crosses (for author Bajo) lie also generally below it, perhaps reflecting that these authors' publications (in the fields of memory, language, perception, and methodology) are more often cited in neuroscience, methodology, and statistics journals that, by not being of a strict psychological nature, are not well covered by PsycINFO. This distinctive characteristic seems to also be the reason that green and red crosses generally lie around the diagonal in the center panel of Figure 4 (i.e., GS does not seem to find citations for those papers that were not already found within the WoS database) and above the diagonal in the bottom panel (i.e., GS finds citations that were not found in the PsycINFO database).

Interestingly, then, WoS and PsycINFO cover unevenly, and in different ways, the various areas of psychology, as also reported by Frandsen and Nicolaisen (2008). A similar result was reported by Meho and Yang (2007) for different areas in library and information science, although their comparison involved WoS and Scopus rather than WoS and PsycINFO.

Overlap of Authenticated Citation Records Across Platforms

All platforms retrieved unique citations. Figure 5a illustrates by showing the number of unique (authenticated) citations in PsycINFO against the number of citations in WoS for each publication; Figure 5b does the same for unique citations in GS against unique citations in the union of WoS and
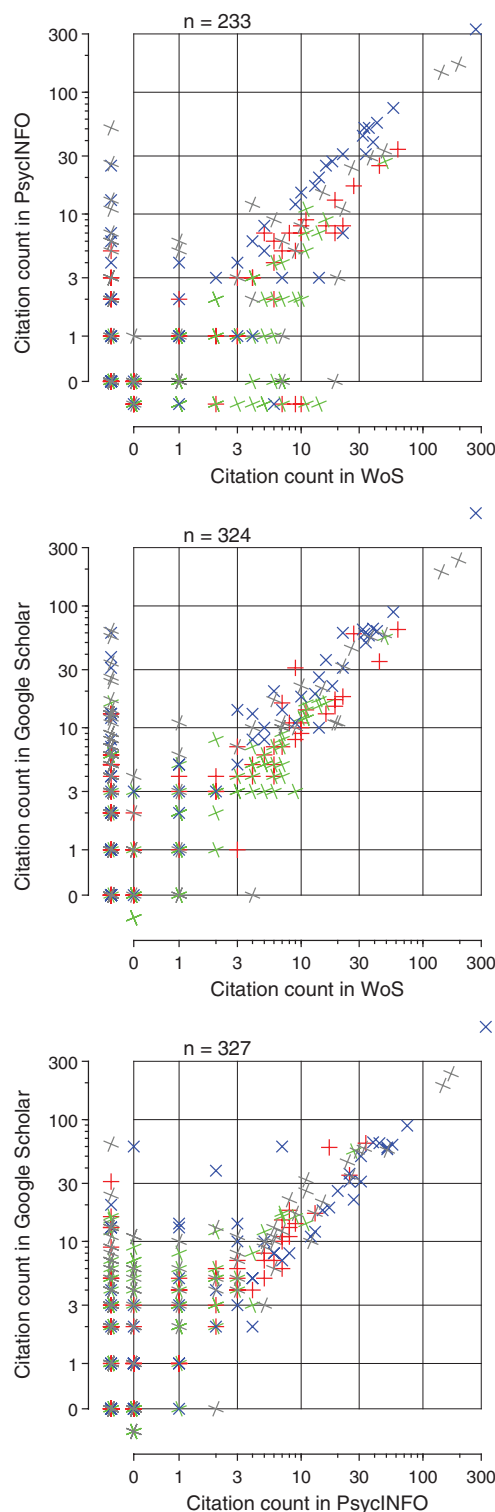

FIG. 4. Scatter plot of authenticated citation counts for all publications across paired databases. Publications from each of the four authors in our study sample are indicated with crosses of different colors and rotation angles so that they remain visible when plotted at the same location in each panel. Data points on the left of the grid area in each panel reflect publications for which a record could not be found through the platform indicated in the horizontal axis, but which had a record in the database indicated in the vertical axis for which the citation count is given by the ordinate of that data point. Data points below the grid area in each panel have a similar interpretation.
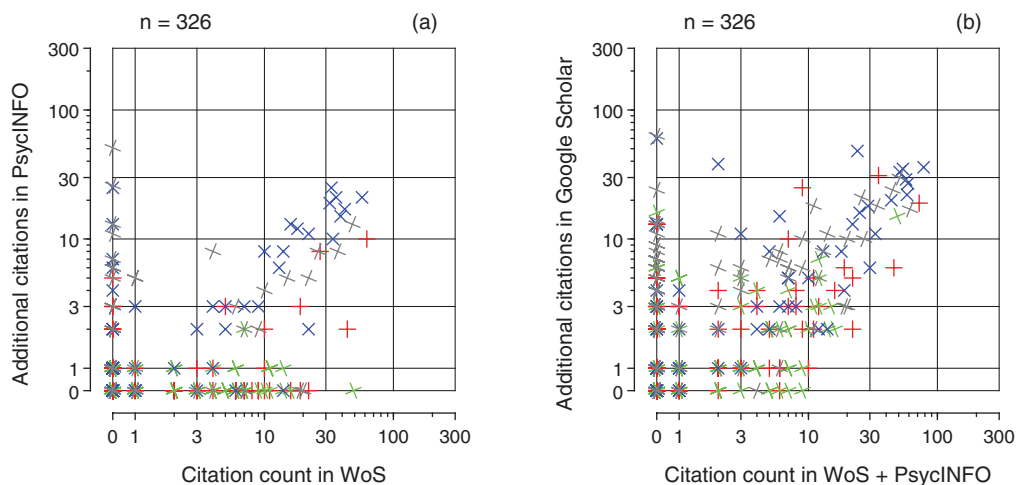
FIG. 5. New and unique authenticated citations supplied by PsycINFO against the number of authenticated citations supplied by Web of Science (WoS) (a), and new and unique authenticated citations supplied by Google Scholar against the number of authenticated citations jointly supplied by WoS and PsycINFO (b). Crosses of different colors and rotation angles pertain to publications by different researchers. The 51 publications with null citation counts in all of the databases are excluded, as are three publications with citation counts in excess of 140 in all platforms.

PsycINFO. This analysis includes all 326 publications for which a record existed in at least one of the platforms[7] and null citation counts reflect that either a record for the publication was not found in that platform or it was found but had a null citation count.

Regarding the *incremental utility* of additional sources, the overall number of authenticated citations across researchers and publications was 1,331 in WoS, 1,879 (i.e., an additional 548 citations, or 41.2% more) in the union of WoS and PsycINFO, and 3,405 (i.e., an additional 1,526 citations, or 81.2% more than WoS + PsycINFO and 155.8% more than WoS) when citations from GS were also included. PsycINFO delivered 1,233 authenticated citations, 648 of which (52.6%) had also been retrieved through WoS. Likewise, GS delivered 2,746 citations, 1,220 of which (44.4%) had also been retrieved through WoS + PsycINFO. On evaluating these figures, it should be recalled that considering additional platforms also brings in further publications for which the preceding set of platforms did not provide citation counts: 157 publications in WoS, 230 (46.5% more) in WoS + PsycINFO, and 326 (41.7% more) in WoS + PsycINFO + GS. Then, consideration of additional platforms not only increased the count of valid citations but also enlarged (doubled) the number of cited publications.

Figure 5 carries the messages that WoS may yield incomplete citation information even for the publication records that it actually includes, that PsycINFO can certainly supplement that information, and that GS completes the picture, although at the extra expense of manual authentication of long lists of purported citations, in which 16.5% of the entries are faulty. The average citation count across the 326 publications in Figure 5 is 4.08 for WoS, 5.76 for WoS + PsycINFO,

and 10.45 for the union of the three platforms. In other words, compared with the number of citation records in WoS, the overall number is more than 2.5 times larger when information from PsycINFO and GS is additionally used.

A substantial number of the new citations found through PsycINFO were book sources, which WoS does not cover, but some also came from scholarly journals not covered by WoS. Similarly, a substantial number of the new citations found through GS came from further books and chapters (often a result of the access to Google Books, although most references to and from Spanish textbooks were actually found by GS in digital copies hosted by their authors on their Web pages; a similar finding was reported by Meho & Yang, 2007). Other sources of unique citations in GS were doctoral dissertations, proceedings, or other documents that are generally available on the Web and are substantially more narrowly or not at all covered by WoS or PsycINFO,[8] and also papers in scholarly journals covered by neither WoS nor PsycINFO. These citations are unreachable within WoS or PsycINFO because they are buried as "orphan references" (i.e., references that cannot be linked to a master record in the database for lack of coverage of the cited source; Jacsó 2008a).

---

[7]Three publications with citation counts in excess of 140 in all platforms were excluded.

[8]On considering the "value" or "quality" of citations retrieved through GS, it should be kept in mind that what is being considered is the impact of the cited author's work. Shadbolt, Brody, Carr, and Harnad (2006) stated that "the *impact* of a piece of research is the degree to which it has been useful to other researchers and users in generating further research and applications: how much the work has been read, used, built-upon, applied and cited in other research as well as in educational, technological, cultural, social and practical applications." From this perspective, the quality of the citing item, whether or not it has been peer reviewed, or the scientific stature of its author (in regard, e.g., to master's theses, technical reports, or research reports) are not pertinent. In any case, citation sources of a controversial nature were scarcely found in GS lists. No precise tallying method was used with them, but the author's feeling is that they represented less than 5% of the material.

Interestingly, PsycINFO listed citing papers that WoS should also have listed; similarly, GS listed citing papers that WoS and/or PsycINFO should also have retrieved. All these judgments are based on the coverage statements of WoS and PsycINFO. It was then mandatory to look into the reason for these unexpected outcomes, because they might unveil further errors in the WoS and PsycINFO databases. The results of this inquiry are described next.

Missing Citations and Stray References

Todd and Ladle (2008) estimated that about 34% of the items in the reference lists of journal articles contain errors. Their estimate comes from five separate studies in the biomedical literature, but the figure is unlikely to be very different in other disciplines. In WoS, some of the missing citations referred to at the end of the preceding section turned out to be "stray references" (i.e., references that cannot be linked to a master record in the database due to errors of referencing by the authors of the citing paper; Jacsó, 2008a, 2008d). Errors in volume or page numbers and authors' names when all remaining data were correct produced stray references; in contrast, even gross errors in the title of the cited article did not render stray references. In other cases, a missing citation in WoS was actually the result of errors committed by indexers or data entry operators (whether human or machine), as illustrated in Figure 6. The hunt that identified these errors also revealed phantom citations.

In PsycINFO, errors of referencing were also found to render stray references, although they were usually accompanied by clickable citation counts. Then, lists of articles in which the cited publication was referenced in the same incorrect way could be retrieved. Also, stray references arising from errors of commission by indexers or data entry operators are frequent in PsycINFO, particularly for articles with unusual page numbers such as those preceded by an "S" to indicate a Supplement issue.

Our search for missing citations in WoS and PsycINFO and the discovery of stray references could not have been possible without the information retrieved through GS. This shows that a nonnegligible number of unique citations encountered through GS come from the scholarly sources represented by refereed and reputable journals, which could not be retrieved otherwise because of referencing or indexing errors. We must nevertheless stress that finding these unique citations was possible because our GS search was based on major word sequences in the known title of the publications for which citation records were sought; it is not at all clear that GS would have actually been able to find these publications (and, hence, citations for them) if the search had been based on authors' names.

Multidimensional h Indices

Large differences in citation counts across platforms must render different h indices. Also, given the disjoint sets of citation records in each database (Figure 5), a realistic h index can only be estimated by aggregating citations across platforms. But there is also the issue that the h index disregards citations to papers not contributing to h, leading to a weakness illustrated in Table 1 of Bornmann and Daniel (2009), in Table 1 of Mingers (2009), and in Table 2 of Moed (2009). For this reason, here we will use a multidimensional extension of the h index that has been shown to solve part of these problems (García-Pérez, 2009b). In this extension, the first component is the conventional h, (i.e., $h_1 = h$), whereas the remaining components (h_2, h_3, and so on) arise by iteratively applying the same logic to the remaining papers. For instance, consider an individual who has published 40 papers, whose ordered citation counts are $C = $ (64, 59, 35, 31, 18, 17, 16, 15, 14, 13, 13, 13, 11, 11, 10, 9, 8, 7, 7, 7, 6, 6, 6, 5, 5, 5, 5, 5, 5, 4, 4, 4, 4, 4, 4, 3, 3, 2, 1, 1). In this case, $h = 12$ and the 5-dimensional index is $\widetilde{H}_5 = $ (12, 7, 5, 5, 4), whose components are the cardinalities of the underlined sets of frequencies in the above vector C: the first set comprises the 12 papers contributing to h; the second set includes the next seven papers, each of which has been cited seven times at least; the third set comprises the next five papers; and so on. This multidimensional extension can be computed to any arbitrary number of components (depth).

Figure 7 plots the five-dimensional h index for each individual in our sample as estimated from authenticated citation counts in each separate database (colored crosses with different rotation angles) from aggregated citation counts in the union of WoS and PsycINFO (gray circles) and from aggregated citation counts in the union of all three platforms (solid circles). A number of features are worth commenting on. First, all the components of $\widetilde{H}_5$ from WoS (red crosses) are at or above their counterparts from PsycINFO (green crosses) for the two individuals on the left, whereas the opposite holds for the two individuals on the right, something that further attests to the already discussed uneven coverage of different areas of psychology in each database. What this reveals is that researchers in various fields of psychology are differentially penalized according to the platform that is used to compute their h indices: One of the authors in our sample has an h of 9 in WoS which goes down to only 5 in PsycINFO, whereas another has an h of 10 in WoS which goes up to 11 in PsycINFO. Second, citations in the union of WoS and PsycINFO render a 5-dimensional index (gray circles) whose components are generally slightly higher than those of the indices separately computed from either platform. Third, GS yields a 5-dimensional index (blue crosses) whose components are generally higher than those obtained from the union of WoS and PsycINFO (gray circles). Finally, the components of $\widetilde{H}_5$ computed from the union of citations across all platforms (black circles), are only occasionally and minimally higher than their counterparts computed from GS alone (blue crosses), a natural outcome of the fact that GS retrieved most of the citations that could also be retrieved through WoS and PsycINFO.

The characteristics just discussed are not to be mistaken for a proof of the adequacy of GS in citation analysis: They are

VÁzquez Carmelo, Muñoz Manuel, Rodriguez Abelardo, 1999, « Spain », *in* Carl O. Helvie
et Wilfried Kuntsmann (ed.) *Homelessness in the United States, Europe and Russia:
a comparative perspective*, Westport, CT: Bergin & Garvey, USA, p. 169-205.
VÁzquez Carmelo, Muñoz Manuel, Sanz Jesus, 1997, « Lifetime and 12-month prevalence
of DSM-III-R mental disorders among the homeless in Madrid: a european study using
the CIDI », *Acta Psychiatrica Scandinavica* 95, p. 523-530.

Cited References

Title: An advantage with limits: the lower risk for women of becoming homeless

Author(s): Maryse, M

Source: POPULATION Volume: 54 Issue: 6 Pages: 885 Published: NOV-DEC 1999

▦ **Citation Map**

References: **96**   ◄◄ ◄ Page [1] of 4 (Go) ► ►►

To find Related Records: Clear the checkbox to the left of an item if you do not want to retrieve articles that cited
when finding Related Records. Then click "Find Related Records."

(Clear All Pages) (Find Related Records)

☑ **1.** SYMPTOME DOMICILE FI : 1995

☑ **23.** CARMELO V
        ACTA PSYCHIAT SCAND 95 : 523 1997

☑ **90.** VAZQUEZ C
        HOMELESSNESS US EURO : 169 1999

Results Year Published=(1997) AND Publication Name=(acta psychiatrica scandinavica) AND Author=(vazquez)
        Refined by: Authors=(VAZQUEZ, C)
        Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH.

Results: **1**   ◄◄ ◄ Page [1] of 1 (Go) ► ►►   Sort by: [Latest Date]

(Print) (E-mail) (Add to Marked List) (Save to EndNote Web) (Save to EndNote, RefMan, ProCite)   ▤ Analyze Results
more options   ▥ Create Citation Report

☑ **1.** Title: Lifetime and 12-month prevalence of DSM-III-R mental disorders among the homeless in Madrid: A European
study using the CIDI
Author(s): Vazquez C, Munoz M, Sanz J
Source: **ACTA PSYCHIATRICA SCANDINAVICA** Volume: **95** Issue: **6** Pages: **523-530** Published: **JUN 1997**
Times Cited: **26**

FIG. 6.   A correct reference in the citing article is introduced with error in the Web of Science (WoS) database. The box at the top shows a cutout of the reference list of the citing article where two publications by the same first author are given. The first of these refers to a book chapter that naturally yields the orphan reference #90 in the WoS record for the citing article (see the box in the center); for the second of these publications, which has a master record within WoS (see the box at the bottom), the first author's given name was introduced as his last name, yielding the stray reference #23 in the WoS record of the citing article (see the box in the center).

determined by the conditions of our search (i.e., by known article title and not by author name, so that all publication records could be found if they existed) and by our analysis (which painstakingly authenticated citations in GS so as to eliminate what turned out to be 16.5% of erroneous entries). For a realistic description of the inadequate and misleading characteristics of GS under typical conditions, see Jacsó (2006, 2008a, 2008b, 2008e).

Results shown in Figure 7 also prove empirically a nonnegligible influence of missing publications and missing citations on *h*, in contrast to what Rousseau (2007) claimed (see also Sanderson, 2008). In particular, the *h* index varies from 8 to 13, 5 to 11, 14 to 20, and 10 to 17, respectively, for each of the four authors in our sample when citation counts come from either the least favorable database (which, in turn, varies across individuals) or the combination of all three databases.
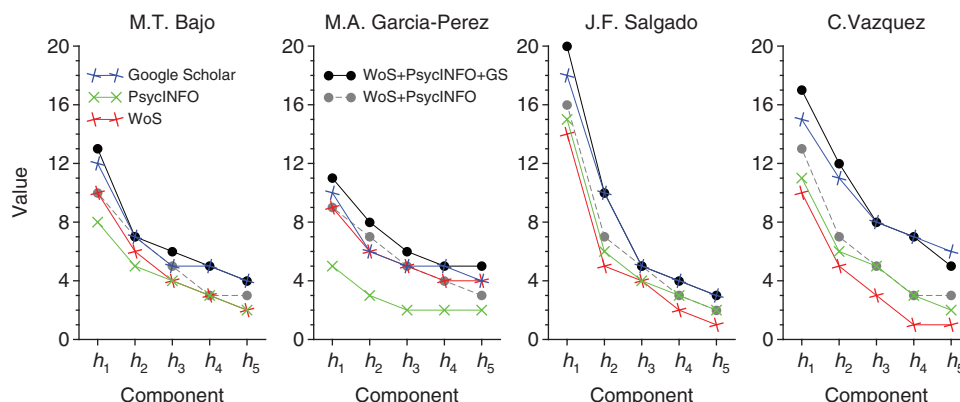
FIG. 7. Five-dimensional *h* index for each author in our study as obtained from authenticated citation counts in each database (colored symbols and lines), and their counterparts when unique citations are aggregated across Web of Science (WoS) and PsycINFO (gray symbols and dashed lines) or across all three platforms (WoS, PsycINFO, and Google Scholar; black symbols and continuous lines).

Jacsó (2008a) reported an also dramatic increase (from 13 to 26) in the *h* index of F.W. Lancaster when citations are willingly searched for beyond what standard search tools provide automatically.

Discussion

Numerous studies have compared coverage and overlap of publication and citation records in various databases at different points in time and in different disciplines, and some of these studies have warned about the limitations of the databases and the need for highly skilled and defensive searching (e.g., Levine-Clark & Gil, 2009b; Meho & Rogers, 2008; Meho & Yang, 2007; Norris & Oppenheim, 2007). The question lurking behind most of those studies was whether dependable data for the computation of scientometric indicators could be obtained from a single platform and, if so, which one that is. The diversity of results had the same cardinality as the set of studies, although studies carried out when GS had been improved and its reach was expanded concurred in reporting that GS usually renders the largest number of publication and citation records. But those studies also showed in one way or another that GS is not dependable.

The present study aimed at searching for citations of known publications in the manner in which authors would seek this information for computation of their *h* index. Thus, we omitted the initial search for unknown publication records by author name, which is the first step in some studies. The complex queries (see Tables 2–4) required to retrieve these authors' publications (and only theirs) from WoS and PsycINFO attest to the difficulties of (and warn about the potential errors incurred by) blind searches by author name with no available disambiguation information. And a search for publication records by author name in GS is well-known to yield such poor results that the endeavor should not even be attempted. Even if GS were endowed with the ability to identify the author of a publication, its unsophisticated "Advance Scholar Search" interface prevents searchers from filtering out publications by other authors with the same name.

Pilot tests for our study revealed that GS is still a long way from being competitive for the accrual of citation records by author name. Yet, it is evident from our results that GS can, in its present form, play a valuable role in the retrieval of citation records for known publications, if only because of the deficiencies of its profiteering counterparts in this respect. Citations not unveiled by commercial platforms (and, yet, coming from unquestionably scholarly sources) generally represent a significant contribution to the *h* index, and these citations can only be dug out by GS.

Our results also corroborate those of Bakkalbasi et al. (2006) or Meho and Yang (2007) in that GS provides unique material and the provision is often substantial. Yet, at a 16.5% rate of inadequate items in the lists of citing papers, material retrieved by GS requires authentication. Our results support Baneyx's (2008, p. 370) conclusion that "GS can be very useful in demonstrating evidence of a broader intellectual and international impact than is possible with WoS." We also agree with Meho and Yang (2007, p. 2111) who assert that "unless a system is developed that automatically and accurately parses result sets into error-free, meaningful, and usable data, GS will be of limited use for large-scale comparative citation and bibliometric analyses." Nonetheless, GS output is usable at the scale of individual researchers for computation of their own *h* index. Although the likelihood of a generalized, small-scale, individual use of GS for this purpose should not deter Google from developing adequate automatic parsing and cleansing tools, the prospect that this capability is eventually achieved will hopefully spur its profiteering competitors to further improve their products by allowing easy handling and integration of the substantial number of orphan and stray references present in their records.

All things considered, we agree with Jacsó (2008e, p. 103) that "using [GS] for bibliometric and scientometric evaluation, comparison and ranking purposes can produce very unscholarly measures and indicators of scholarly productivity and impact." This caveat notwithstanding, our results show that GS offers invaluable help to *collect citations for known publications*, with a generous share coming from

unquestionably valid but unreachable scholarly sources not covered by subscription databases, but also including a nonnegligible share from journals covered by subscription databases that are sprinkled with stray and orphan references. Having said this, searchers should remember that they will have to face the gruesome process of authenticating the citations listed by GS to harvest the benefits. Our study has also revealed a number of little known characteristics of WoS and PsycINFO that have practical implications in the current atmosphere of impact measures and research evaluation, as discussed next.

Some Reasonable Improvements in WoS and PsycINFO

The prevalence of stray and orphan references in WoS and PsycINFO need to be investigated. Although it is understandable that for-profit organizations that provide subscription database services would not be interested in providing figures for these undesirable features, there are a number of ways in which their products could be improved to better serve their customers.

WoS should consider broadening its coverage, given that WoS records include many orphan references to journals that are not covered for what appears to be no good reason. Titles such as *Spatial Vision* or the *Spanish Journal of Psychology* (among many others), which are covered extensively and with citation counts elsewhere in WoK, are potential candidates. In contrast, and inexplicably, obscure, old, and nonpeer-reviewed journals that only published papers in Spanish (examples intentionally omitted) are covered by WoS at least since 1970.

WoS should also consider covering books, more as sources than as targets of citations. Garfield (2008) argued that "a book that is cited in hundreds of published articles will rarely be cited in more than a dozen or so books on a related topic, so even when Google processes citations to all books in the future, I do not think this will change the measurable impact of books." This being true, the point of considering books goes exactly in the opposite direction, namely, to provide a more accurate measure of the impact of journal articles by additional consideration of the number of times that they are cited in books. On another front, it should not be overlooked that book chapters are also an integral part of the publication culture in some fields (Huang & Chang, 2008), and it has actually been shown that citations in books cannot be predicted from citations in journal articles (Cronin, Snyder, & Atkins, 1997).

Finally, WoS should consider unearthing orphan and stray references by at least providing links to the citing articles in the output of "Cited Reference Search," which will allow searchers to seek authentication of the purported citations.

On its side, PsycINFO should consider abandoning discretional indexing of journals that are definitely psychological in content, and retrospective information should be added to consolidate its position as a competitive source of publication and citation records in psychology. The need for name authority control (Bennett & Williams, 2006) is also urgent in PsycINFO, whose current carelessness in this respect burdens the searcher unnecessarily on consideration that authors' names as written on their papers are substantially and inconsistently altered upon indexing.

A Protocol for the Collection of Citation Records

The coverage of a given platform seems to be guided by the assumption that a "carefully selected subset of journals would produce the majority of important citing literature for any given article" (Bakkalbasi et al., 2006), but this assumption is not empirically realizable when articles in different research fields require their own and unique selected subset of journals (Levine-Clark & Gil, 2009b; Meho & Rogers, 2008; Meho & Yang, 2007; Norris & Oppenheim, 2007). We have shown that PsycINFO, which is specific to psychology, is inadequate for the subfields of psychology that are close to the neurosciences, whereas it seems adequate for the subfields that have few or no connections with other sciences. This poses the question of how could the available platforms be used for a fair computation of measures of individual achievement that are not biased by uneven coverage across disciplines or across the subfields of a discipline. This question, in turn, has ramifications for measuring achievement at the aggregate levels of departments, institutions, or countries (see, e.g., Jacsó, 2009a, 2009b), with the additional difficulty that affiliation data cannot always be searched through all platforms, and, hence, final figures must be obtained by aggregating individuals data.

Psychologists seeking to collect citation records for their papers could overcome biases of differential coverage with the three-step strategy pursued in this article, as itemized next.

1. Use a multidisciplinary platform such as WoS as a start. Citation errors in WoS are rare and authentication is generally easy.
2. Continue with PsycINFO. Errors in PsycINFO are also relatively rare, although somewhat more common than in WoS, and they are sometimes easily identifiable. Authentication of the new purported citations is also easy with PsycINFO.
3. Seek additional citations through GS, searching for records by major word sequences in the title of each publication. This last step will unearth stray and orphan references otherwise unreachable in WoS and PsycINFO, but it will also locate citations in academic sources not covered by WoS and PsycINFO. This last step may yield about as many new citations as the two previous steps together, but authentication is mandatory.

For the accrual of stray and orphan references, Jacsó (2007, 2008a) has demonstrated how to use the "Cited Reference Search" interface in WoS. Yet, this approach has limited utility: Indeed, (a) the searcher must be lucky enough to guess correctly the character string that embodies the stray or orphan reference in the database; (b) disambiguation is impossible and often a major problem when the author has a common name (Bar-Ilan, 2006; García-Pérez, 2009a); (c) the

search is limited to sources covered by the database, which are often a narrow subset of all the relevant sources; and, most important, (d) authentication becomes impossible because there is no link to the citing document. In contrast, the search for known publications in GS in the way illustrated in this article is free of all these problems.

We should emphasize that our results may not be replicated in other disciplines and, hence, that the general applicability of our conclusions may be limited in an unknown way. This disclaimer notwithstanding, veteran users of the traditional abstracting/indexing tools surely remember the old days in which it was crystal clear that to be successful in a comprehensive search, several databases ought to be searched. As of the day of this writing, and well into the computerized online searching era, the situation does not seem to have changed a bit.

Acknowledgments

References

Agrawal, A.A. (2005). Corruption of journal impact factors. Trends in Ecology and Evolution, 20(4), 157.

Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E., & Herrera, F. (2009). h-index: A review focused in its variants, computation and standardization for different scientific fields. Journal of Informetrics, 3(4), 273–289.

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomedical Digital Libraries, 3,7. Retrieved May 19, 2010, from http://www.bio-diglib.com/content/3/1/7

Ballard, S., & Henry, M. (2006). Citation searching: New players, new tools. Searcher, 14(9), 24–33.

Baneyx, A. (2008). "Publish or Perish" as citation metrics used to analyze scientific output in the humanities: International case studies in economics, geography, social sciences, philosopy, and history. Archivum Immunologiae et Therapiae Experimentalis, 56(6), 363–371.

Bar-Ilan, J. (2006). An ego-centric citation analysis of the works of Michael O. Rabin based on multiple citation indexes. Information Processing and Management, 42(6), 1553–1566.

Bar-Ilan, J. (2008). Which h-index?—A comparison of WoS, Scopus and Google Scholar. Scientometrics, 74(2), 257–271.

Bauer, K., & Bakkalbasi, N. (2005). An examination of citation counts in a new scholarly communication environment. D-Lib Magazine, 11(9). Retrieved May 19, 2010, from http://www.dlib.org/dlib/september05/bauer/09bauer.html

Bennett, D.B., & Williams, P. (2006). Name authority challenges for indexing and abstracting databases. Evidence Based Library and Information Practice, 1(1), 37–57. Retrieved May 19, 2010, from http://ejournals.library.ualberta.ca/ index.php/EBLIP/article/view/7/86

Bornmann, L., & Daniel, H.-D. (2009). The state of h index research: Is the h index the ideal way to measure research performance? EMBO Reports, 10(1), 2–6.

Bornmann, L., Marx, W., Schier, H., Rahm, E., Thor, A., & Daniel, H.-D. (2009). Convergent validity of bibliometric Google Scholar data in the field of chemistry—Citation counts for papers that were accepted by Angewandte Chemie International Edition or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. Journal of Informetrics, 3(1), 27–35.

Bornmann, L., Mutz, R., & Daniel, H.-D. (2008). Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h index using data from biomedicine. Journal of the American Society for Information Science and Technology, 59(5), 830–837.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. Ethics in Science and Environmental Politics, 8(1), 93–102.

Brown, N.L. (1999). On the trail of the prolific Dr Path. Nature, 398, 555.

Brumback, R.A. (2008a). Worshiping false idols: The impact factor dilemma. Journal of Child Neurology, 23(4), 365–367.

Brumback, R.A. (2008b). Response to correspondence, "'Worshiping false idols: The impact factor dilemma': Correcting the record." Journal of Child Neurology, 23(9), 1092–1094.

Brumback, R.A. (2009). Impact factor wars: Episode V—The empire strikes back. Journal of Child Neurology, 24(3), 260–262.

Cameron, B.D. (2005). Trends in the usage of ISI bibliometric data: Uses, abuses, and implications. Portal: Libraries and the Academy, 5(1), 105–125.

Campbell, P. (2008). Escape from the impact factor. Ethics in Science and Environmental Politics, 8(1), 5–7.

Carrió, I. (2008). Of impact, metrics and ethics. European Journal of Nuclear Medicine and Molecular Imaging, 35(6), 1049–1050.

Colquhoun, D. (2003). Challenging the tyranny of impact factors. Nature, 423, 479.

Cronin, B., Snyder, H., & Atkins, H. (1997). Comparative citation rankings of authors in monographic and journal literature: A study of sociology. Journal of Documentation, 53(3), 263–273.

Della Sala, S., & Brooks, J. (2008). Multi-authors' self-citation: A further impact factor bias? Cortex, 44(9), 1139–1145.

Falagas, M.E., & Alexiou, V.G. (2007). Editors may inappropriately influence authors' decisions regarding selection of references in scientific articles. International Journal of Impotence Research, 19(5), 443–445.

Falagas, M.E., & Alexiou, V.G. (2008). The top-ten in journal impact factor manipulation. Archivum Immunologiae et Therapiae Experimentalis, 56(4), 223–226.

Falagas, M.E., Pitsouni, E.I., Malietzis, G.A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. FASEB Journal, 22(2), 338–342.

Frandsen, T.F., & Nicolaisen, J. (2008). Intradisciplinary differences in database coverage and the consequences for bibliometric research. Journal of the American Society for Information Science and Technology, 59(10), 1570–1581.

García-Pérez, M.A. (2001). The decade 1989–1998 in Spanish psychology: An analysis of research in statistics, methodology, and psychometric theory. Spanish Journal of Psychology, 4(2), 111–122. Retrieved May 19, 2010, from http://www.ucm.es/sjp

García-Pérez, M.A. (2009a). The Hirsch h index in a non-mainstream area: Methodology of the Behavioral Sciences in Spain. Spanish Journal of Psychology, 12(2), 833–849. Retrieved May 19, 2010, from http://www.ucm.es/sjp

García-Pérez, M.A. (2009b). A multidimensional extension to Hirsch's h index. Scientometrics, 81(3), 779–785.

Garfield, E. (1999). Journal impact factor: A brief review. Canadian Medical Association Journal, 161(8), 979–980.

Garfield, E. (2008). Introduction. Archivum Immunologiae et Therapiae Experimentalis, 56(6), 353.

Glavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. Online Information Review, 32(1), 8–21.

Haeffner-Cavaillon, N., & Graillot-Gak, C. (2009). The use of bibliometric indicators to help peer-review assessment. Archivum Immunologiae et Therapiae Experimentalis, 57(1), 33–38.

Harnad, S. (2008). Validating research performance measures against peer rankings. Ethics in Science and Environmental Politics, 8(1), 103–107.

Harzing, A.-W.K., & van der Wal, R. (2008). Google Scholar as a new source for citation analysis. Ethics in Science and Environmental Politics, 8(1), 61–73.

Hernán, M.A. (2008). Epidemiologists (of all people) should question journal impact factors. Epidemiology, 19(3), 366–368.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the USA, 102(46), 16569–16572.

Huang, M.-H., & Chang, Y.-W. (2008). Characteristics of research output in social sciences and humanities: From a research evaluation perspective. Journal of the American Society for Information Science and Technology, 59(11), 1819–1828.

Jacsó, 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, 89(9), 1537–1547.

Jacsó, P. (2005b). Comparison and analysis of the citedness scores in Web of Science and Google Scholar. In E.A. Fox, E. Neuhold, P. Premsmit, & V. Wuwongse (Eds.), Digital libraries: Implementing strategies and sharing experiences (pp. 360–369). Berlin, Germany: Springer.

Jacsó, P. (2006). Deflated, inflated and phantom citation counts. Online Information Review, 30(3), 297–309.

Jacsó, P. (2007). Software issues related to cited references. Online Information Review, 31(6), 892–905.

Jacsó, P. (2008a). Testing the calculation of a realistic h-index in Google Scholar, Scopus, and Web of Science for F.W. Lancaster. Library Trends, 56(4), 784–815.

Jacsó, P. (2008b). The pros and cons of computing the h-index using Google Scholar. Online Information Review, 32(3), 437–452.

Jacsó, P. (2008c). The pros and cons of computing the h-index using Scopus. Online Information Review, 32(4), 524–535.

Jacsó, P. (2008d). The pros and cons of computing the h-index using Web of Science. Online Information Review, 32(5), 673–688.

Jacsó, P. (2008e). Google Scholar revisited. Online Information Review, 32(1), 102–114.

Jacsó, P. (2009a). Errors of omission and their implications for computing scientometric measures in evaluating the publishing productivity and impact of countries. Online Information Review, 33(2), 376–385.

Jacsó, P. (2009b). The h-index for countries in Web of Science and Scopus. Online Information Review, 33(4), 831–837.

Joseph, K.S., & Hoey, J. (1999). CMAJ's impact factor: Room for recalculation. Canadian Medical Association Journal, 161(8), 977–978.

Kotiaho, J.S. (1999). Papers vanish in mis-citation black hole. Nature, 398, 19.

Kotiaho, J.S., Tomkins, J. L., & Simmons, L.W. (1999). Unfamiliar citations breed mistakes. Nature, 400, 307.

Kousha, K., & Thelwall, M. (2008). Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. Scientometrics, 74(2), 273–294.

Kulkarni, A.V., Aziz, B., Shams, I., & Busse, J.W. (2009). Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. Journal of the American Medical Association, 302(10), 1092–1096.

Lavie, P. (2009). The race for the impact factor. Journal of Sleep Research, 18(3), 283–284.

Lawrence, P.A. (2007). The mismeasurement of science. Current Biology, 17(15), R583–R585.

Lehmann, S., Jackson, A.D., & Lautrup, B.E. (2008). A quantitative analysis of indicators of scientific performance. Scientometrics, 76(2), 369–390.

Levine-Clark, M., & Gil, E.L. (2009a). A comparative citation analysis of Web of Science, Scopus, and Google Scholar. Journal of Business and Finance Librarianship, 14(1), 32–46.

Levine-Clark, M., & Gil, E. (2009b). A comparative analysis of social sciences citation tools. Online Information Review, 33(5), 986–996.

Meho, L.I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science. Journal of the American Society for Information Science and Technology, 59(11), 1711–1726.

Meho, L.I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. Journal of the American Society for Information Science and Technology, 58(13), 2105–2125.

Mingers, J. (2009). Measuring the research contribution of management academics using the Hirsch-index. Journal of the Operational Research Society, 60(9), 1143–1153.

Moed, H.F. (2002). The impact-factors debate: The ISI's uses and limits. Nature, 415, 731–732.

Moed, H.F. (2009). New developments in the use of citation analysis in research evaluation. Archivum Immunologiae et Therapiae Experimentalis, 57(1), 13–18.

Nature. (2005a). Not-so-deep impact [Editorial]. Nature, 435, 1003–1004.

Nature. (2005b). Ratings games [Editorial]. Nature, 436, 889–890.

Nederhof, A.J. (2006). Bibliometric motinoring of research performance in the Social Sciences and Humanities: A review. Scientometrics, 66(1), 81–100.

Neuhaus, C., & Daniel, H.-D. (2008). Data sources for performing citation analysis: An overview. Journal of Documentation, 64(2), 193–210.

Neuhaus, C., Neuhaus, E., Asher, A., & Wrede, C. (2006). The depth and breadth of Google Scholar. Portal: Libraries and the Academy, 6(2), 127–141.

Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. Journal of Informetrics, 1(2), 161–169.

Opatrný, T. (2008). Playing the system to give low-impact journal more clout. Nature, 455, 167.

Osca-Lluch, J., Civera Mollá, C., & Peñaranda Ortega, M. (2009). Consecuencias de los errores en las referencias bibliográficas [Consequences of the error in bibliographical references: The specific case of the journal Psicothema]. El caso de la revista Psicothema. Psicothema, 21(2), 300–303.

Pauly, D., & 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, 5, 33–35.

Porta, M., & Álvare-Dardet, C. (2008). How come scientists uncritically adopt and embody Thomson's Bibliographic Impact Factor? Epidemiology, 19(3), 370–371.

Price, N.C. (1998). What's in a name (or a number or a date)? Nature, 395, 538.

Reedijk, J., & Moed, H.F. (2008). Is the impact of journal impact factors decreasing? Journal of Documentation, 64(2), 183–192.

Rogers, L.F. (2002). Impact factor: The numbers game. American Journal of Roentgenology, 178(3), 541–542.

Rossner, M., van Epps, H., & Hill, E. (2007). Show me the data. Journal of Cell Biology, 179(6), 1091–1092.

Rossner, M., van Epps, H., & Hill, E. (2008). Irreproducible results: A response to Thomson Scientific. Journal of Cell Biology, 180(2), 254–255.

Roth, D.L. (2005). The emergence of competitors to the Science Citation Index and the Web of Science. Current Science, 89(9), 1531–1536.

Rousseau, R. (2007). The influence of missing publications on the Hirsch index. Journal of Informetrics, 1(1), 2–7.

Salgado, J.F., & Páez, D. (2007). La productividad científica y el índice h de Hirsch de la psicología social española: Convergencia entre indicadores de productividad y comparación con otras áreas [Scientific productivity and Hirsch's h Index of Spanish social psychology: Convergence between productivity indexes and comparison with other areas]. Psicothema, 19(2), 179–189.

Sanderson, M. (2008). Revisiting h measured on UK LIS and IR academics. Journal of the American Society for Information Science and Technology, 59(7), 1184–1190.

Schroeder, R. (2007). Pointing users toward citation searching: Using Google Scholar and Web of Science. Portal: Libraries and the Academy, 7(2), 243–248.

Schutte, H.K., & Švec, J.G. (2007). Reaction of Folia Phoniatrica et Logopaedica on the current trend of impact factor measures. Folia Phoniatrica et Logopaedica, 59(6), 281–285.

Seglen, P.O. (1997a). Why the impact factor of journals should not be used for evaluating research. British Medical Journal, 314(7079), 498–502.

Seglen, P.O. (1997b). Citations and journal impact factors: Questionable indicators of research quality. Allergy, 52(11), 1050–1056.

Sevinc, A. (2004). Manipulating impact factor: An unethical issue or an Editor's choice? Swiss Medical Weekly, 134(27–28), 410.

Shadbolt, N., Brody, T., Carr, L., & Harnad, S. (2006). The open research Web. In N. Jacobs (Ed.), Open access: Key strategic, technical and economic aspects (pp. 195–206). Oxford: Chandos.

Shultz, M. (2007). Comparing test searches in PubMed and Google Scholar. Journal of the Medical Library Association, 95(4), 442–445.

Simons, K. (2008). The misused impact factor. Science, 322 (no. 5899), 165.

Smith, R. (1998). Unscientific practice flourishes in science: Impact factors of journals should not be used in research assessment. British Medical Journal, 316(7137), 1036.

The PLoS Medicine Editors. (2006). The impact factor game. PLoS Medicine, 3(6): e291.

Todd, P.A., & Ladle, R.J. (2008). Hidden dangers of a 'citation culture.' Ethics in Science and Environmental Politics, 8(1), 13–16.

Topo Universitario. (2008). Factor de impacto del International Journal of Clinical and Health Psychology [Impact factor of the International Journal of Clinical and Health Psychology]. Retrieved May 19, 2010, from http://unnombrealazar. blogspot.com/2008/09/factor-de-impacto-del-international.html

van Diest, P.J., Holzel, H., Burnett, D., & Crocker, J. (2001). Impactitis: New cures for an old disease. Journal of Clinical Pathology, 54(11), 817–819.

Vaughan, L., & Shaw, D. (2008). A new look at evidence of scholarly citation in citation indexes and from university. Scientometrics, 81(2), 587–600.

Vieira, E.S., & Gomes, J.A.N.F. (2009). A comparison of Scopus and Web of Science for a typical university. Scientometrics, (4), 1121–1132.

Waheed, A.A. (2003). Citation rate unrelated to journals' impact factors. Nature, 426, 49.

Walters, W.H. (2007). Google Scholar coverage of a multidisciplinary field. Information Processing and Management, 43Web sources. Scientometrics, 74(2), 317–330.

Walters, W.H. (2009). Google Scholar search performance: Comparative recall and precision. Portal: Libraries and the Academia, 9(1), 5–24.

Whitley, K.M. (2002). Analysis of SciFinder Scholar and Web of Science citation searches. Journal of the American Society for Information Science and Technology, 53(14), 1210–1215.

Wilcox, A.J. (2008). Rise and fall of the Thomson impact factor. Epidemiology, 19(3), 373–374.

Williams, G. (1998). Misleading, unscientific, and unjust: The United Kingdom's research assessment exercise. British Medical Journal, 316(7137), 1079–1082.

Wróblewski, A.K. (2008). A commentary on misuses of the impact factor. Archivum Immunologiae et Therapiae Experimentalis, 56(6), 355–356.

Yu, G., & Wang, L. (2007). The self-cited rate of scientific journals and the manipulation of their impact factors. Scientometrics, 73(3), 321–330.