

Citation analysis in research on differential item functioning

Georgina Guilera · Juana Gómez-Benito ·
M. Dolores Hidalgo

Published online: 10 September 2009
© Springer Science+Business Media B.V. 2009

Abstract This paper offers a first approach to the study of citing behaviour in the literature on differential item functioning (DIF) and provides new information about the main characteristics that define that behaviour. With a focus on articles listed in the Web of Science for the period 1975–2000, the present research considers the number of self-citations included in each paper with respect to other bibliometric indicators such as year of publication, the number of signatories to a paper, author's country of origin and the journal in which it was published. In general terms it can be concluded that in studies of DIF the mean percentage of self-citations is around 11%, this figure remaining stable over time; there is a positive correlation between the number of signatories to an article and the number of self-citations. This paper has offered a first approach to the study of citing behaviour in the literature on DIF.

Keywords Bibliometrics · Differential item functioning · Citation analysis · Self-citation

1 Introduction

In recent decades, numerous studies have sought to analyse various facets of citing behaviour, the aims being, among others, to understand why some studies are cited and others not (Case and Higgins 2000), to explore similarities and differences between areas of knowledge (Glänzel and Schoepflin 1999), to determine relationships between bibliometric indicators (Shadish et al. 1995) and to establish predictive models (Basulto Santos and Ortega Irizo 2005). One of the aspects to have aroused most interest in the bibliometric literature concerns

G. Guilera (✉) · J. Gómez-Benito
Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035 Barcelona, Spain
e-mail: gguilera@ub.edu

M. D. Hidalgo
Departamento de Psicología Básica y Metodología, Facultad de Psicología, Universidad de Murcia, Campus Universitario de Espinardo, 30100 Murcia, Spain

interpretation and the role played by self-citations in the process of scientific communication (Glänzel et al. 2004). Citing oneself is considered to be both natural and acceptable (Phelan 1999), and the absence of self-citations would even be surprising given the accumulative nature of research. Some researchers (Bonzi and Snyder 1991; Lawani 1982; Pichappan and Sarasvady 2002) have sought to understand why authors make self-citations and basically conclude that it is usually done to give continuity to their own research lines, to establish scientific authority, to raise the profile of their work and to increase the proportion of citations received. In this context, it is important to determine the characteristics of self-citations in different areas and to explore the relationships with other indicators, the results then being compared with references to other studies. From this bibliometric perspective it would be useful to analyse the subject areas of greatest interest within a given discipline; in this regard, one of the key issues with respect to the measurement of psychological and social phenomena concerns the invariance of measurement instruments, in other words, the degree to which the instruments used measure all subjects equally.

Since the appearance in the USA of the first tests designed to help select personnel for companies, students for admission to university, or soldiers for the army, not to mention their use in many other public and private sectors, standardized tests or measurement instruments have become an indispensable tool in the social and health sciences, especially in the fields of psychology and education.

Indeed, in these two fields a wide range of measurement instruments is used to obtain information about the traits, skills or knowledge of individuals and groups. In this regard it is essential that the professionals who use these tests ensure that the latter function in the same way in all groups with respect to socio-demographic variables such as sex, culture or ethnicity, since all such instruments must be objective in their measurements, that is, they must be able to provide identical results in subjects who have the same level on a measured attribute regardless of the group to which they belong. When one of the test items functions differently in favour (or against) subjects with a similar ability the measurement instrument is said to show differential item functioning (DIF).

Given the social, ethical and legal implications of DIF it continues to be an issue that is subject to debate and analysis. From the bibliometric perspective a study has already been conducted which sought to offer a clear overview of scientific productivity with respect to published studies of DIF (Gómez Benito et al. 2005); however, an analysis of bibliographic references has yet to be carried out in this field.

2 Purpose

Following this approach the aim of the present research was to conduct a bibliometric study of end-of article references in the DIF literature. This was done by analysing the references listed in published studies and by distinguishing between self-citations and those concerning the work of other authors. The relations with other variables are explored, such as the number of authors per paper, the journal in which the papers were published, and the total number of references cited per paper.

3 Methods

The strategy and search criteria are the same as those used in the first bibliometric study of DIF (Gómez Benito et al. 2005). Briefly, this previous study included articles referring to DIF

and published in scientific journals between 1975 and the year 2000, these being gathered by applying the search criteria “(differential item functioning) or DIF or (item bias)” to the databases SCI-EXPANDED and SSCI. This approach yielded a total of 271 articles.

Thus, given the aim of the present study the variables coded for each of the articles were: year of publication, number of signatories to the article, the name of the journal where it was published, the impact factor of this journal, the number of references, the number of times the authors cite their own work (self-citations) and that of others (foreign citations), and, finally, the number of citations made by a journal to studies published in the same journal.

Self-citation was defined as when any one of the signatories to an article also appeared as one of the authors of a study listed in the bibliography. Therefore, a ‘foreign citation’ was when a reference did not meet this criterion. As regards self-citations, it should be clarified that they were counted in a synchronous way and by taking into account any of the authors of the source study, regardless of his or her position in the list of signatories.

The impact factor was taken from Journal Citation Reports, which has been available online since 1998 and thus only values since this date were considered. In coding this variable, each article was assigned the impact factor of the corresponding journal for the year in which it was published.

Various analyses based on frequencies, percentages, means and correlations were conducted using SPSS and Excel.

4 Results

4.1 General descriptive results

The overall mean number of references for all the studies analysed was 28.47 (SD = 18.50), with a range of 0–159; the mean for self-citations was 2.85 (SD = 2.61; range 0–13), while the figure for foreign citations was 25.63 (SD = 17.78; range 0–157). Glänzel and Schoepfelin (1999) reported a mean number of references per study in the field of *Psychology and Psychiatry* of 31.00, a value very similar to that found here.

Similarly, the proportion of self-citations with respect to the total number of references was 10.91%, with a range of 0–75%; in a bibliometric study of the field *Social Sciences and Humanities* Bonzi and Snyder (1990) report a proportion of around 11%, very similar to the figure obtained here. However, according to the paper by Hyland (2003) the proportion of self-citations in the study of DIF is closer to the values found in the hard sciences (12.50%) than in the soft sciences (4.30%).

4.2 Evolution over time

Figure 1 shows the evolution over time of the mean number of references (both self-citations and those of other authors), and it can be seen that the mean number of self-citations per study is always lower than that of foreign citations.

It can also be observed that the mean number of self-citations remains more or less constant over time, whereas the total number of references, and therefore the number of foreign citations, follows a clearly heterogeneous pattern.

Although it has not been directly explored in the present study these results suggest that the mean number of self-citations per study is not related to the accumulated scientific output in a given area by an author or group of authors; rather, the temporal homogeneity would seem to be due either to the tendency among authors only to cite more recent research

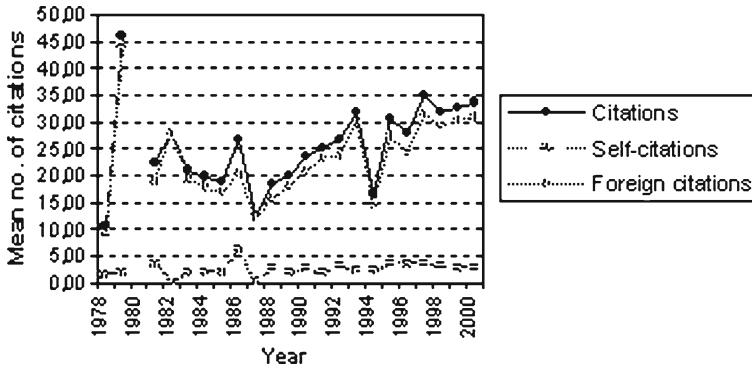


Fig. 1 Evolution over time of the mean number of citations

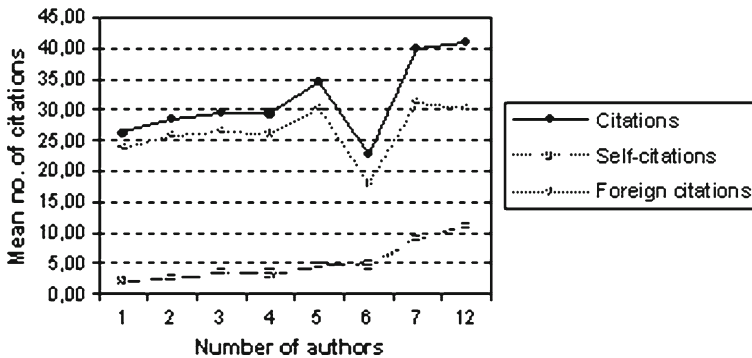


Fig. 2 Mean number of citations according to the number of authors

(Aksnes 2003) or to the supposed existence of a maximum number of self-citations, which when exceeded would be viewed by the scientific community as excessive or unnecessary.

4.3 Number of authors

The correlations between the number of references and the number of foreign citations with the number of signatories to the article were $r = 0.081$ ($p = 0.181$) and $r = 0.038$ ($p = 0.538$), respectively, and it can thus be stated that the relationship between these variables is practically null.

However, when analysing the number of self-citations with respect to the number of authors one would expect to find a general trend toward a greater number of the former as the number of signatories increases (Dimitroff and Arlitsch 1995; Glänzel and Thijs 2004), as generally speaking a greater number of authors yields a greater number of previous studies that might potentially be self-cited. In this regard, the correlation between the number of authors and the number of self-citations was $r = 0.320$ ($p < 0.01$), as shown graphically in Fig. 2.

In a study that explored this same correlation Snyder and Bonzi (1998) report similar results in areas as diverse as *Chemistry* ($r = 0.301$) and *Geology* ($r = 0.361$), although they also found small negative correlations in other areas such as *Economics* and *Art History*,

Table 1 Mean number of citations, self-citations and foreign citations according to country

Country	Citations	Self-citations (%)	Foreign citations (%)
USA	26.94	2.65 (9.84)	24.30 (90.20)
Spain	26.70	1.48 (5.54)	25.22 (94.46)
Holland	34.50	4.17 (12.08)	30.33 (87.92)
Canada	30.30	3.80 (12.54)	26.50 (87.46)
Australia	34.38	4.00 (11.64)	30.38 (88.37)
Denmark	31.38	4.63 (14.74)	26.75 (85.26)

Table 2 Mean number of citations, self-citations and foreign citations according to journal

Journal	Citations	Self-citations (%)	Foreign citations (%)
J Educ Meas	22.11	2.70 (12.22)	19.43 (87.87)
Appl Psych Meas	28.61	4.61 (16.12)	24.00 (83.88)
Educ Psychol Meas	16.87	0.57 (3.38)	16.30 (96.65)
Psicothema	27.33	1.19 (4.35)	26.14 (95.64)
Appl Meas Educ	31.53	1.07 (3.39)	30.47 (96.63)
Psychometrika	30.50	4.63 (15.16)	25.88 (84.84)

as [Tagliacozzo \(1977\)](#) who concluded that self-citation did not appear to be related to the number of co-authors in the articles of *Plant Physiology* and *Neurobiology* areas.

4.4 Countries

Table 1 shows the characteristics of citations according to the country of origin of the authors of the source article. As can be seen, the country with the lowest mean number of references per article is Spain (26.70), while that with the highest is the Netherlands (34.50). Spain is also the country with the lowest percentage of self-citations (5.54%), while in this case it is Denmark which has the greatest proportion of this type of citation (14.74%).

4.5 Journals

Here self-citation refers to the fact that the reference list of an article published in a given journal contains another study published in the same journal, regardless of the names of the authors. The analysis includes those journals which have published ten or more studies on DIF.

As can be seen in Table 2 the journal with the smallest mean number of references per article is *Educational and Psychological Measurement* (EPM), while that with the highest is *Applied Measurement in Education*; EPM also has the lowest percentage of self-citations (3.38%), while in this case *Applied Psychological Measurement* (APM) is the journal with the highest percentage of references to studies published in the same journal (16.12%).

When exploring the relationship between the proportion of self-citations and the impact factor of journals (in this case, journals with less than ten published studies were also included), only those studies published between 1998 and 2000 were analysed. Although one might expect there to be a positive relationship between the two variables the correlation coefficient obtained was $r = 0.164$ ($p = 0.154$), thus indicating the lack of a statistically significant relationship between them.

5 Conclusions

This paper has offered a first approach to the study of citing behaviour in the literature on DIF, providing new information about the main characteristics that define it and describing the relationships between different variables. However, it should be borne in mind that the accuracy of the results obtained here may be affected by both the presence of homonyms and confusion as regards the names of authors.

In general terms it can be concluded that in studies of DIF the mean percentage of self-citations is around 11%, this figure remaining stable over time. In addition, there is a positive correlation between the number of signatories to an article and the number of self-citations, this not being observed in the case of the number of foreign citations. In terms of countries of origin, the lowest percentage of self-citations corresponds to Spain, while Denmark has the highest proportion in this regard. Finally, the journal with the lowest percentage of self-citations is EPM, while the highest corresponds to APM.

It would be interesting to extend this kind of study to the analysis of citations from a diachronic perspective with the aim of providing new evidence about the process of scientific communication in the area of DIF (in terms of the age of references or patterns in the evolution of the literature) and exploring the relationships between the different bibliometric indicators.

Acknowledgements This study was partially supported by grants 2007FIC 00736 and 2005SGR00365 from the “Departament d’Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya”, and SEJ2005-09144-C02-02 from Spain’s “Ministerio de Ciencia y Tecnología” under European Regional Development Found (ERDF).

References

- Aksnes, D.W.: A macro study of self-citation. *Scientometrics* **56**, 235–246 (2003)
- Basulto Santos, J., Ortega Irizo, F.J.: Modelling citation age data with right censoring. *Scientometrics* **62**(3), 329–342 (2005)
- Bonzi, S., Snyder, H.W.: Patterns of self citations across fields of inquiry. *Proceedings of the 53rd Annual Meeting of the American Society for Information and Science* **27**, 204–207 (1990)
- Bonzi, S., Snyder, H.W.: Motivations for citation: a comparison of self citation and citation to others. *Scientometrics* **21**(2), 245–254 (1991)
- Case, D.O., Higgins, G.M.: How can we investigate citation behavior? A study of reasons for citing literature in communication. *J. Am. Soc. Inf. Sci.* **51**(7), 635–645 (2000)
- Dimitroff, A., Arlitsch, K.: Self-citations in the library and information science literature. *J. Doc.* **51**(1), 44–56 (1995)
- Glänzel, W., Schoepflin, U.: A bibliometric study of reference literature in the sciences and social sciences. *Inf. Process. Manag.* **35**, 31–44 (1999)
- Glänzel, W., Thijs, B.: Does co-authorship inflate the share of self-citations?. *Scientometrics* **61**(3), 395–404 (2004)
- Glänzel, W., Thijs, B., Schlemmer, B.: A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics* **59**(1), 63–77 (2004)
- Gómez Benito, J., Hidalgo Montesinos, M.D., Guilera Ferré, G., Moreno Torrente, M.: A bibliometric study of differential item functioning. *Scientometrics* **64**(1), 3–16 (2005)
- Hyland, K.: Self-citation and self-reference: credibility and promotion in academic publication. *J. Am. Soc. Inf. Sci. Technol.* **54**(3), 251–259 (2003)
- Lawani, S.M.: On the heterogeneity and classification of author self-citations. *J. Am. Soc. Inf. Sci.* **33**(5), 281–284 (1982)
- Phelan, T.J.: A compendium of issues for citation analysis. *Scientometrics* **45**(1), 117–136 (1999)

- Pichappan, P., Sarasvady, S.: The other side of the coin: the intricacies of author self-citations. *Scientometrics* **54**(2), 285–290 (2002)
- Shadish, W., Tolliver, D., Gray, M., Gupta, S.: Author judgments about works they cite: three studies from psychology journals. *Soc. Stud. Sci.* **25**(3), 477–498 (1995)
- Snyder, H., Bonzi, S.: Patterns of self-citation across disciplines (1980–1989). *J. Inf. Sci.* **24**(6), 431–435 (1998)
- Tagliacozzo, R.: Self-citations in scientific literature. *J. Doc.* **33**(4), 251–265 (1977)