# Gender Differences in Publication and Citation Counts in Librarianship and Information Science Research

Celia Sánchez Peñas and Peter Willett[1]
Department of Information Studies, University of Sheffield, 211 Portobello, Sheffield S1 4DP, UK

**Abstract**:

## INTRODUCTION

Gender is one of the key components in understanding individuals and their working lives. An important component of academic working life is the publication of research findings, and the subsequent citation of those publications, and there have hence been many investigations of the publication and citation behaviour of men and women in a range of subject fields [1-8]. With the notable exception of women's studies [9], the general conclusion from these investigations is that women publish less than men and are less likely to be cited than are men. Many reasons have been suggested for this, most notably the under-representation of women in academe at the higher levels [2, 4, 8, 10-12]. Librarianship and information science (hereafter LIS) is a discipline that attracts both men and women, both as practitioners and as academics, and it is hence of interest to ascertain whether the differences that have been observed in other subjects are equally applicable here. However, with the exception of the study by Korytnyk [13], there do not appear to have been any detailed studies of the extent to which gender affects the publication and citation performance of LIS academics.

In principle, it is very easy to study the effect of gender by identifying a body of male and female LIS academics and then counting the publications they have authored and the citations that those publications have attracted, using public databases such as *Google Scholar*, *Library and Information Science Abstracts*, *Library Literature and Information Science Full Text*, or the *Web of Knowledge*. Bibliometric studies of LIS research using such databases are commonplace but there is a complicating factor that needs to be taken into account when considering a multidisciplinary subject such as LIS, *viz* the fact that publication and citation behaviour is very different in different subject domains [14-16]. This implies that different parts of LIS research are likely to provide very different publication and citation counts [17, 18] unless care is taken to standardise the data in some way. In this paper we attempt to provide such a standardisation of available bibliometric data to enable meaningful conclusions to be drawn as to the publications by, and citations to, the research of male and female LIS academics.

[1] To whom all correspondence should be addressed. Email p.willett@sheffield.ac.uk

## METHODS

Our basic approach involved identifying the staff, both male and female, in five of the top LIS departments worldwide, and searching for all of their publications, and citations to those publications, in the *Web of Knowledge*, an online database (at URL http://wok.mimas.ac.uk/) that provides access for UK higher education institutions to the citation files produced by Thompson ISI. The departments chosen are listed in Table 1, together with the numbers of male and female staff listed on their Web pages as being members of the faculty. It will be seen that the five departments are comparable in size and that they all have approximately equal numbers of male and female staff.

The next step was to identify the subject areas in which these 105 individuals carried out research, this again being achieved by recourse to the departmental web sites to identify words and phrases that described the individuals' research interests; it was not possible at this stage further to process those faculty members from the Royal Danish School of Librarianship whose entries were only in Danish. The intention was then to classify all of the staff into some number of subject areas so that comparisons could be made between the male and female staff in each of these areas. Over 300 different subjects were identified, covering a huge range of topics; for example, limiting attention to just the 34 subjects commencing with the letter "C" we found not only classical LIS subjects such as Cataloguing, Citation Studies, Classification and Comparative Librarianship, but also Cold-War Censorship, Complex Adaptive Systems, Computer-Mediated Discourse Analysis and Cultural Change. This diversity of topics inevitably meant that there were often only a very few researchers in each area, which would limit the possibilities for quantitative analysis. It was accordingly decided to group the various subjects that had been identified into a much smaller number of broad subject categories.

Biggs and Biggs [19] have suggested a seven-part sub-division of LIS research, but this could not be used: some of their categories (such as school librarianship and library history) are poorly represented in our sample, and many of the topics identified here are not represented in their categories. After a fair amount of trial and error, eight categories were identified that covered the bulk of the subject areas identified from the Web sites and, most importantly, that had sufficient male and female academics working in each area 4to enable a sensible comparison of publications and citations to be made. The resulting eight categories are listed in Table 2, where it will seen that we have had to carry out a large amount of rather *ad hoc* grouping of related topics in some cases to ensure that sufficient male and female academics are present. In other cases, it was not possible to do this, i.e., there was a marked preponderance of either male or female academics, and there are thus LIS topics that were studied by some of the academics from Table 1 but that are not represented in Table 2. Examples of such subject areas include health information, legal issues, and chemical and biological information systems. The low counts also meant that it was not possible, as had been intended at the start of the study, to carry out a quantitative comparison of the publications by, and citations to, the sets of US and European male and female academics.

The *Web of Knowledge* was used to retrieve all of the publications that were associated with each of the authors in the chosen subject areas, and the citations to

each of the resulting publications. This enabled counts to be generated for the numbers of publications and citations by male and female academics for each of the eight categories in Table 2. Each row in this table contains a broad subject description and then the numbers of academics, publications and citations to those publications, sub-divided by gender. For example, our analysis identified 11 academics (five males and six females) working on Cataloguing and related subjects; these academics published a total of 38 papers on this subject (28 male papers and 10 female papers) and these 38 papers attracted a total of 382 citations (251 citations to the male papers and 131 to the female papers).

Three comments should be made about the counts in Table 2. First, a paper was regarded as a male (female) paper if a male (female) faculty member from one of the five chosen departments was an author, even if there were other authors with the opposite gender. Thus, a given paper might be considered as both a male paper and a female paper if it had multiple authors from one or more of the chosen institutions with both genders. Second, the decision was taken to omit from the analysis three male academics – Cronin, Ingwersen and Willett – who we considered to represent outlier points in terms of publication and citation counts, when compared with the other academics considered here. In all, these three academics were associated with no less than 312 publications and 5516 citations, numbers that would have swamped the counts associated with their peers. Third, we have used just a single database (*Web of Knowledge*) to obtain the numeric data, and Meho and Spurgin [20] have shown that multiple databases are required for a fully comprehensive survey of the research productivity of LIS faculty; the results presented in the next section can hence be regarded only as indicative.

## RESULTS AND DISCUSSION

The numeric data is presented in Table 2, which details the numbers of publications by, and citations to, male and female authors in each of eight broad subject categories. An inspection of the bottom row of this table reveals a greater number of male and female authors than the total numbers of academics in Table 1: this is because an individual author might publish in more than one of the eight broad subject categories.

Inspection of the publication and citation data in Table 2 suggests strongly that there are substantial differences in the numbers of papers published by, and the numbers of citations to, male and female LIS authors. These impressions are assessed quantitatively in Table 3, where the publication counts in each broad subject area have been normalised by the numbers of authors for each of the two genders, and where the numbers of citations in each broad subject area have been normalised by the numbers of papers for each of the two genders. Thus, the grouping in Table 2 has been carried out to ensure that there is sufficient data to enable a meaningful comparison to be made for a particular broad subject area, and the mean values have been computed in Table 3 for each such area to ensure that the data is appropriately normalised. It will be seen that precisely one-half of the figures in the main body of Table 3 have been starred: this is to denote which gender has the higher mean value in each subject category.

The mean values in Table 3 have been analysed using the Sign Test [21]; the more common approach to the comparison of mean values, the *Z* test, cannot be used here since there is insufficient data for the individual subject areas to enable the test to be used. Specifically, the null hypothesis $H_O$ was tested that there was no difference between the mean male and mean female publication figures and between the mean male and mean female citation figures. In the case of the publications, there are eight occurrences where the male mean figure is higher and none where the female mean figure is higher. In the Sign Test, the numbers of non-equal cases are used in a calculation based on the Binomial Distribution, and $H_O$ could be rejected with $p \leq 0.008$ (two-tailed test) for this set of eight subject categories. In the case of the citations, there are six occurrences where the male mean figure is higher and two where the female mean figure is higher; in this case, $H_O$ could not be rejected ($p \leq 0.289$, i.e., no significant difference). The data was also analysed using the more powerful Wilcoxon Signed Ranks Test [21], with the same conclusions: in the case of the publication data, $H_O$ could be rejected ($p \leq 0.008$), whilst for the citation data $H_O$ could not be rejected ($p \leq 0.250$).

The figures in Table 3 suggest that there are significant differences in the numbers of publications produced by mean and women. This observation might be affected by one or both of the following factors: the men in the sample had worked for a longer time, and hence had more opportunity to publish; the men in the sample had reached a higher level in the profession, and hence had a greater visibility that would facilitate publication and citation. An analysis was hence carried out to ascertain whether the male authors had worked for longer, as determined by the date of their first publication in the sample considered here. There is a slight difference in the mean year of first publication – 1991.21 for the men and 1993.47 for the women - but the difference is not significant ($p \leq 0.136$) in a *Z*-test with the null hypothesis that the mean year of first publication is the same for the two samples of authors. An analysis of the websites for the five departments showed that there were 15 male professors and 10 female professors in our sample (the former figure excludes the three male outliers noted previously); thus while there is a greater percentage of male faculty who are professors, the difference is by no means overwhelming. We hence believe that the data in Table 3 provides a true guide to the publication differences between men and women in LIS.

## CONCLUSIONS

In this paper we have reported a comparison of the publications by, and the citations to, 57 male and 48 female academics in five leading LIS departments of librarianship and information science. The raw bibliometric data that was obtained from the *Web of Knowledge* database was carefully analysed to identify eight broad subject areas for which there was sufficient data to enable meaningful comparisons to be made between the research activities of the male and female faculty. Statistical analysis of the resulting data shows that the male LIS academics do publish significantly more papers on average than do female LIS academics: this situation mirrors that observed in other disciplines, despite the high proportion of women working in the area as professionals and academics. There are, however, no significant differences in the numbers of citations to published papers by male and by female LIS academics.

## REFERENCES


1. S. Acker, Women, the other academics, *British Journal of Sociology of Education* 1 (1980) 81-91.
2. A. Brooks, *Academic Women* (Open University Press, Buckingham, 1997).
3. M.A. Ferber, Citations and networking, *Gender and Society* 2 (1988) 82-89.
4. S.H. Kaplan, L.M. Sullivan, K.A. Dukes, C.F. Phillips, R.P. Kelch and J.G. Schaller, Sex differences in academic advancement — results of a national study of paediatricians, *The New England Journal of Medicine* 335 (1996) 1282-1290.
5. J.S. Long (ed.) *From Scarcity to Visibility: Gender Differences in the Careers of Doctoral Scientists and Engineers* (National Academies Press, Washington DC, 2001).
6. A.L. Mathews and K. Andersen, A gender gap in publishing? Women's representation in edited political science books, *PS: Political Science and Politics* 34 (2001) 143-147.
7. National Science Foundation, Division of Science Resources Statistics, *Gender Differences in the Careers of Academic Scientists and Engineers: a Literature Review, NSF 03-322*, (National Science Foundation, Arlington VA, 2003).
8. K. Prpiû, Gender and productivity differentials in science, *Scientometrics* 55 (2002) 27–58.
9. J.M. Cullars, Citation characteristics of English-language, *Monographs in Philosophy, Library & Information Science Research* 20 (1998) 41-68.
10. S. Luzzadder-Beach and A. Macfarlane, The environment of gender and science: status and perspectives of women and men in physical geography, *Environment of Gender and Science* 52 (2000) 407-424.
11. B. McElhinny, M. Hols, J. Holtzkener, S. Unger and C. Hicks, Gender, publication and citation in sociolinguistics and linguistic anthropology: the construction of a scholarly canon, *Language in Society* 32 (2003) 299-328.
12. L. Morley and V. Walsh, *Feminist Academics: Creative Agents for Change* (Taylor and Francis, London, 1995)
13. C.A. Korytnyk, Comparison of the publishing patterns between men and women Ph.D.s in librarianship, *Library Quarterly* 58 (1988) 52-65.
14. H. Dundar and D.R. Lewis, Determinants of research productivity in higher education, *Research in Higher Education* 39 (1998) 607 –631.
15. A.F.J. Van Raan, The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschatzung – Theorie und Praxis*, 12 (2003) 20-29.
16. L. Vaughan and D. Shaw, Web citation data for impact assessment: a comparison of four science disciplines, *Journal of the American Society for Information Science and Technology*, 56 (2005) 1075-1087.
17. A. Holmes and C. Oppenheim, Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management, *Information Research*, 6(2) (2001) at http://InformationR.net/ir/6-2/paper103.html [accessed 16 January 2006]
18. M. Santos, P. Willett and F.E. Wood, Research degrees in librarianship and information science: a survey of master's and doctoral students from the Department of Information Studies, University of Sheffield, *Journal of Librarianship and Information Science* 30 (1998) 49-56.
19. M. Biggs and V. Biggs, Library and information science faculty: their lives as scholars, *Library Quarterly* 63 (1993) 282-317.
20. L.I. Meho and K.M. Spurgin, Ranking the research productivity of library and information science faculty and schools: an evaluation of data sources and research methods, *Journal of the American Society for Information Science and Technology* 56 (2005) 1314-1331.
21. S. Siegal and N.J. Castellan, *Nonparametric Statistics for the Behavioural Sciences* (McGraw Hill, New York, 1988).

| University and Department Name | Male Faculty | Female Faculty |
|---|---|---|
| University of Loughborough Department of Information Science | 11 | 9 |
| University of Sheffield, Department of Information Studies | 11 | 9 |
| Royal School of Library and Information Science, Department of Information Science | 11 | 8 |
| University of Illinois at Urbana-Champaign, Graduate School of Library and Information Science | 14 | 11 |
| University of Indiana at Bloomington, School of Library and Information Science | 10 | 11 |
| Total | 57 | 48 |

Table 1. Numbers of male and female faculty for the five chosen departments

| Broad Subject Area | Authors | | Publications | | Citations | |
|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female |
| Human and social aspects of information handling; Organisational behaviour; User studies | 7 | 8 | 20 | 20 | 134 | 115 |
| Digital libraries; E-books; E-publishing | 15 | 8 | 50 | 21 | 182 | 38 |
| Information retrieval | 15 | 9 | 49 | 28 | 438 | 223 |
| Books; Collection, records and library management; Literature; Preservation; Printing; Publishing | 12 | 13 | 102 | 39 | 148 | 94 |
| Automation; Database systems; Systems management; Technical issues | 7 | 6 | 38 | 18 | 336 | 95 |
| Cataloguing; Classification; Indexing; Knowledge organisation; Taxonomies; Thesaurus construction | 5 | 6 | 28 | 10 | 251 | 131 |
| Bibliometrics; Citation studies; Informetrics; Webometrics | 9 | 3 | 50 | 7 | 585 | 37 |
| Information literacy; Teaching and learning; | 11 | 9 | 43 | 27 | 311 | 86 |
| Total | 81 | 62 | 380 | 170 | 2385 | 819 |

Table 2. Publications by, and citations to, male and female academics in eight broad subject areas.

| Broad Subject Area | Mean Publications per Author | | Mean Citations per Publication | |
|---|---|---|---|---|
| | Male | Female | Male | Female |
| Human and social aspects of information handling; Organisational behaviour; User studies | 2.86* | 2.50 | 6.70* | 5.75 |
| Digital libraries; E-books; E-publishing | 3.33* | 2.63 | 3.64* | 1.81 |
| Information retrieval | 3.27* | 3.11 | 8.94* | 7.96 |
| Books; Collection, records and library management; Literature; Preservation; Printing; Publishing | 8.50* | 3.00 | 1.45 | 2.41* |
| Automation; Database systems; Systems management; Technical issues | 5.43* | 3.00 | 8.84* | 5.28 |
| Cataloguing; Classification; Indexing; Knowledge organisation; Taxonomies; Thesaurus construction | 5.60* | 1.67 | 8.96 | 13.10* |
| Bibliometrics; Citation studies; Informetrics; Webometrics | 5.56* | 2.33 | 11.70* | 5.29 |
| Information literacy; Teaching and learning; | 3.91* | 3.00 | 7.23* | 3.19 |
| Mean | 4.69* | 2.74 | 6.28* | 4.82 |

Table 3. Mean numbers of publications by, and citations to, male and female academics in eight broad subject areas. The larger figure in each male-female pair is starred. The bottom row contains the mean values computed from the totals in the bottom row of Table 2.