
National bias: a comparison of citation practices by health professionals

By Frank M. Campbell, M.A.L.S.

Virginia Commonwealth University
Medical College of Virginia
Tompkins-McCaw Library
Box 582 MCV Station
Richmond, Virginia 23298-0582

It is hypothesized that health professionals in the United States and the United Kingdom are nationally biased in their citation practices. Articles published in the *New England Journal of Medicine* and *Lancet* were used to study citation practices of U.S. and U.K. authors. Percentages of cited references to material published in a specific country were calculated for both the *New England Journal of Medicine* and *Lancet*. Using a variation of a citation publication ratio based on Frame and Narin's original ratio, an attempt was made to quantify author bias. To calculate these ratios, values from SERLINE* and the British Library Lending Division were employed to find world journal counts. The results suggest that U.S. authors publishing in the *New England Journal of Medicine* and U.K. authors publishing in *Lancet* tend to cite material produced in their own countries more than would be warranted by the amount of material produced by these countries. In addition, these authors cited material produced in non-U.S. and non-U.K. countries far less than the amount of material produced by these countries would indicate.

INTRODUCTION

Citation practice analysis has been a growing area of study in library and information science. Much information can be gathered from citation patterns that appear in professional literature. This article studies the possibility of a national bias in citation practices on the part of health professionals in the United States and the United Kingdom.

It is hypothesized that health professionals in the United States and the United Kingdom chiefly cite material produced in their own country and rarely material produced in other countries. Articles in the *New England Journal of Medicine (NEJM)* and *Lancet* have been chosen to test this hypothesis. An attempt is made to quantify any bias found.

If bias actually does occur among health professionals, it could be causing problems in medical literature. Authors who use only a portion of the total amount of available information run the risk of overlooking important data. Research published in coun-

tries that are routinely ignored is lost to the world's health community or repeated needlessly in other countries.

Authors who use only a portion of the total amount of available information run the risk of overlooking important data. Research published in countries that are routinely ignored is lost to the world's health community or repeated needlessly in other countries.

LITERATURE REVIEW

Before discussing studies related to the research conducted in this article, a look should be taken at authors' motivational factors. An author can be motivated to cite a reference for a number of reasons. Using earlier studies, Brooks synthesized seven reasons why authors cite specific works. Ten authors picked from a number of scientific fields were asked

* SERLINE is a registered trademark of the National Library of Medicine.

to select the motivational factor that best defined their reasons for citing information. Timeliness, "a prestige factor driving authors to show how up-to-date they are by referring only to the latest output of their contemporaries," and reader alert, used "to alert the reader to new, different or obscure sources," ranked highest [1]. A smaller portion of the authors chose persuasiveness and positive credit as motivational factors. Brooks stated,

This exploratory study has shown that not all authors are motivated by the same galaxy of motivations. . . . No longer can we naively assume that authors cite only noteworthy pieces in a positive manner. Authors are revealed to be advocates of their own points of view who utilize previous literature in a calculated attempt to self-justify [2].

Kaplan, another individual working in the field of citation motivation, stated,

My own very preliminary studies lead me to suspect that the citation practices of scientists today are in large part a social device for coping with problems of property right and priority claims. Only incidentally do these citations serve as a careful and accurate reconstruction of the scholarly precursors of one's own contribution [3].

By examining these two studies, one can assume that authors do not always cite material logically. The decision to use some references might be due to emotional reasoning. A variety of questionable motivational factors that impact on an author could contribute to national bias.

Although there were few studies found dealing with possible health professional bias against citing foreign publications, there were studies in other specific fields and in the general field of science. These studies help put the research reported in this article into perspective.

One study that touches on possible national bias in citing material was reported by Cronin. Psychology journal editors and editorial advisory board members were polled using a questionnaire that was created to discover the nature and possible functions of citations. One statement on the form, "National bias in references is inevitable," was particularly important to this research [4]. A majority of those polled (65%) "agreed that national bias was unavoidable," 24% disagreed, and 12% were undecided (percentages do not equal 100 due to a rounding up of figures) [5]. It is apparent that those polled believed that "national bias is in certain circumstances to be expected (and even tolerated)" [6].

Another study by Inhaber and Alvo measured information created by eighteen of the largest scientific nations. These nations produced over 95% of the major scientific literature. Using an input-output concept originally created to measure industry interaction be-

Table 1

Citation in single-author papers in the *British Journal of Educational Psychology* (*BJEP*) and the *Journal of Educational Psychology* (*JEP*) (1978)*

	JEP	BJEP
References to North American literature	97%	47%
References to United Kingdom literature	2%	51%
References to other literature	1%	2%
Self-citation (author citing self)	18%	22%
References to sister journal (references between <i>JEP</i> and <i>BJEP</i>)	0%	4%

* Data taken from Cronin, Transatlantic citation pattern (reference 8).

tween nations, Inhaber and Alvo measured the flow of information among nations. Results showed that the United States ranked highest in the category of self-citation, that is, authors from a country citing material from that same country. The study also indicated that authors from other countries cite U.S. authors much more often than they are cited by their counterparts in the United States [7].

Results showed that the United States ranked highest in the category of self-citation, that is, authors from a country citing material from that same country.

A study that is very similar to the research undertaken in this article was conducted by Cronin in the field of psychology. Two journals—the *British Journal of Educational Psychology* (*BJEP*), published in the United Kingdom, and the *Journal of Educational Psychology* (*JEP*), published in the United States—were studied using citation analysis. Communication patterns were examined by dividing cited material by country of origin. *BJEP* articles were used if written by a U.K.-based author, and *JEP* articles if written by a North American (Cronin did not define the term "North American"). A final tabulation of percentages measured whether the individual citations were references to U.K. literature, North American literature, or other literature. Table 1 lists values for 1978 [8].

The values in Table 1 demonstrate the high number of authors publishing in North America citing other authors publishing in North America. Authors publishing in the U.S. journal cited material published in North America 97% of the time. The British journal, on the other hand, cited material from its own country 51% of the time. While the British journal cited material originating from North America a great deal

Table 2
World journal counts of the British Library Lending Division (BLLD) 1973: percentage of journals in each field from each country*

	Clinical medicine	Biomedical research
United States	21.0%	22.2%
United Kingdom	7.4%	14.1%
Other	71.6%	63.7%
Total	100.0%	100.0%

* Data taken from Carpenter, The subject composition (reference 10).

(47%), the U.S. journal rarely cited U.K. material (2%). Cronin stated

The difference in both volume and direction of the traffic is quite striking, not least in view of the fact that there are no language or political barriers to information transfer between Britain and North America [9].

To accurately measure national bias, one must take into account from which countries authors are publishing and how much they are publishing. One would assume that a large portion of the world's references would be to publications from those countries where authors publish the most. Publications from countries that produce a small amount of material would be cited less frequently. Therefore, it is important to consider the percentage of material published in a given country when testing national bias.

Carpenter and Narin have computed the percentage of world journal production in a variety of scientific fields. They have used a number of sources to create these percentages. Using a magnetic tape of all serial publications received by the British Library Lending Division (BLLD) in 1973, they were able to calculate the percentage of journals produced in four-teen countries and six geopolitical regions. Journals were divided into nine basic fields of science based on the serials title. Results of calculations relevant to this study are shown in Table 2 [10]. Barr explained that BLLD attempts to acquire all science and technical journals published in the world. Its collection is considered to be comprehensive in the field of medicine. Its policy is to obtain all periodicals that might be of interest to scientists and technologists [11]. BLLD is constantly updated by the removal of discontinued journal titles from its files. Therefore, the percentages from Carpenter and Narin listed in Table 2 are for world journals publishing in 1973. BLLD was receiving over 43,000 journals in 1973 [12].

In their later study, Carpenter and Narin measured the adequacy of the *Science Citation Index (SCI)* as an indicator of international scientific activity. As with BLLD, journals indexed in *SCI* were divided by coun-

Table 3
World journal counts of the *Science Citation Index (SCI)* 1973: percentage of journals in each field from each country

	Clinical medicine	Biomedical research
United States	36.4%	36.1%
United Kingdom	11.8%	13.5%
Other	51.8%	50.4%
Total	100.0%	100.0%

try of publication and by basic field of science. Percentages shown in Table 3 were calculated using Carpenter and Narin's journal counts [13].

Using the percentage of journal output by country and the percentage of citations going to a country can help to measure possible bias in citation practices. Frame and Narin constructed a citation-to-publication ratio to measure such citation practices by country. This ratio takes into account not only the number of citations going to a country, but also the amount of material published in that country. Since a larger producer of material could be expected to receive a larger percentage of references, total amount of material published by country is required:

This suggests that if one wishes to make meaningful intercountry comparisons of citations received, one should compensate for the size of the different countries' research efforts [14].

The citation-to-publication ratio is expressed as

$$\frac{\text{Percentage of World Citation Going to a Country}}{\text{Percentage of World Papers Produced by That Same Country}} = \text{Citation-to-Publication Ratio [15]}$$

If a specific country receives 40% of the world's citations in a given area and produces 30% of the world's published material in that area, then its citation-to-publication ratio would be $40/30 = 1.33$. Thus, this country receives 33% more references than its output would appear to warrant. Using this equation, Frame and Narin found that only the United States and the United Kingdom were referenced more than would be indicated by the amount of material they produce.

METHODOLOGY

To test the hypothesis that health professionals publishing in both the United States and the United Kingdom were nationally biased in their citation practices, it was necessary to examine references from specific articles—articles published in the United States and in the United Kingdom. Only authors who

reside in the same country in which the article is published could be chosen, so that articles produced by foreign authors would not obscure the results.

The two source journals, the *NEJM* and *Lancet*, were chosen using the stated criteria and because of their high impact factor in *SCI* [16]. An impact factor is defined as the measure of frequency in which an average article in a specific journal has been cited in a given period of time. Garfield explained that the impact factor is calculated "by dividing the number of times a journal has been cited by the number of articles it has published" [17]. An impact factor was important because of its possible indication of journal significance. Journals that have been cited more than others might be considered more relevant in the field that they encompass. *NEJM* was given an impact factor of 17.75 and *Lancet* a factor of 12.86 in 1986 [18].

A section of *Journal Citation Reports* from *SCI* ranked journals first by category and then by impact factor. In 1986, *NEJM* and *Lancet* ranked higher than any other journal in the category of "Medicine, General and Internal" [19]. Since both journals were so highly cited by health professionals, the citation practices reflected in these journals appeared worthy of investigation.

For this study, the MEDLINE† database was searched using the title abbreviation field to retrieve all articles in *NEJM* and *Lancet* that had been entered into the database between January 1984 and August 1986. Since only original articles were desired, the retrieved group was limited by rejecting all articles with "letter" or "editorial" appearing in the text word field. Citations within a retrieved set pulled from MEDLARS‡ are automatically numbered by the online system. Using randomly generated numbers and corresponding citation numbers, citations were retrieved from the MEDLINE database.

Limiting the retrieved set to original articles brought the total number of possibly acceptable articles to 1,228 for *NEJM* and 1,459 for *Lancet*. A total of 4% of these articles was selected; thus, 49 articles from *NEJM* and 58 articles from *Lancet* were chosen.

Final acceptance of a source article depended upon the existence of citations and the residency of the primary author. If citations did not exist in the source article, the article was dropped from consideration. Since this study involved the testing of possible national bias of U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet*, it was important to accept only those source articles written by authors residing in the same country as the source journal was published. Author residency was presumed to correspond to place of employment. Location of employment was established by checking author infor-

mation in the source journal listing. Articles from *NEJM* were used only if the author was a resident of the United States, and articles from *Lancet* were used only if the author was a resident of the United Kingdom.

After the source articles had been selected, citations were examined. A number of reference tools were used to verify the country of publication for each cited journal article. *Health Sciences Serials* (April 1986 edition) was the primary verification tool [20]. Percentages of cited references in *NEJM* to journal articles published in the United States, the United Kingdom, and "other" countries were calculated. The same was done for *Lancet*.

To measure actual national bias, one must take into account the amount of material published in a given country. The amount of research output for a country would affect the number of references given to that country. Using the citation-to-publication ratio, one can negate the effects of countries that produce a very large amount of research and those that produce a very small amount of research.

To measure actual national bias, one must take into account the amount of material published in a given country. The amount of research output for a country would affect the number of references from that country.

Frame and Narin's original citation-to-publication ratio consisted of the percentage of the world's citations going to a country divided by the percentage of the world's papers produced by that same country [21]. This research did not measure world citation output by country; instead it measured citation output by select authors in a select journal. These authors were U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet*. Because of these differences, the original ratio was altered to conform with these new components. The revised ratio used in this study became

$$\frac{\text{Percentage of (NEJM or Lancet) Citations to Journals Published in a Specific Country}}{\text{Percentage of Medical Journals Produced by the Same Country}} = \text{Citation-to-Publication Ratio}$$

To use the ratio, the total number of medical journals produced by a country was needed. SERLINE and BLLD provided these values. SERLINE was searched in November 1986 using the country of publication field. The SERLINE database listed both currently received serials and discontinued serials. It also

†‡ MEDLINE and MEDLARS are registered trademarks of the National Library of Medicine.

Table 4

Comparison of articles pulled from the MEDLINE database, January 1984–August 1986

Articles	NEJM		Lancet	
	Number	(%)	Number	(%)
Usable articles	49	(84.5%)	58	(43.9%)
Unusable articles				
foreign residency	6	(10.3%)	54	(41.0%)
lack of citations	3	(5.2%)	20	(15.1%)
Total	58	(100.0%)	132	(100.0%)

covered serials peripheral to the field of medicine, such as dental and hospital management journals. To retrieve the number of serials concurrently being published and in a narrow definition of medicine, the values pulled from SERLINE were limited to those journals indexed in *Index Medicus*. Carpenter and Narin's article provided values from BLLD as of 1973 [22]. Citation-to-publication ratios created using SERLINE and those created using BLLD were then compared.

RESULTS AND DISCUSSION

The process of collecting data to test the hypothesis yielded a few secondary discoveries. These discoveries dealt specifically with the two source journals used in this study, *NEJM* and *Lancet*. While these data might reflect on the two journals, they cannot be expected to stand for all medical journals.

The two source journals showed distinct differences when the specified requirements used in this research were applied to source articles. A breakdown of retrieved articles from the MEDLINE database is shown in Table 4.

NEJM published very few articles by authors residing outside the United States. On the other hand, Lancet published more equal numbers of articles by authors residing in the United Kingdom and authors residing in other countries.

As previously mentioned, 49 articles were required from *NEJM* and 58 articles were required from *Lancet* to achieve 4% of the total number of possibly usable articles listed in MEDLINE. For *NEJM*, 58 articles were pulled from the database to achieve 49 usable articles. Thus, nine articles pulled were rejected because of residency of the author or lack of citations. For *Lancet*, 132 articles were required from the database to achieve 58 usable articles. Seventy-four articles were rejected because of residency or lack of citations. Thus, a great many more *Lancet* articles were required from the database to fulfill the number of articles needed.

After viewing Table 4, it appears that *NEJM* pub-

Table 5

Percentage of cited references to material published in a specific country for the *NEJM* and *Lancet*

Country	NEJM	Lancet
United States	80.65%	40.80%
United Kingdom	11.29%	51.90%
Other countries	8.06%	7.30%
Total	100.00%	100.00%

lished very few articles by authors residing outside the United States. On the other hand, *Lancet* published more equal numbers of articles by authors residing in the United Kingdom and authors residing in other countries.

A high percentage of *Lancet* articles were dropped from this study because of lack of citations (15.1%), as compared to the percentage dropped for *NEJM* (5.2%). From those articles originally chosen for this study, *Lancet* had more articles lacking citations.

Because citations were counted for each source article used in this study, the average and median number of citations for articles appearing in the source journals could be calculated. *NEJM* articles had a mean value of 34.02 citations and a median value of 27.0. *Lancet* articles had a mean value of 19.31 citations and a median value of 17.5. These values illustrate that for source articles used in this research, those originating from *NEJM*, on the average, had more citations than those originating from *Lancet*. The difference between the number of citations in both journals would be even greater if articles having no citations had not been rejected. These values do not necessarily mean that articles authored by U.K. health professionals tend to have fewer citations; the numbers might reflect only the publishing policies of the two journals being studied.

Percentages of cited references to journal articles published in a specific country were calculated for each source article. Total percentages could then be computed for all articles chosen from *NEJM* and *Lancet* (Table 5). The values demonstrate that authors publishing in the two source journals cited journal articles published in their own country more than any other country. While authors in both *NEJM* and *Lancet* cite their own country, *NEJM* did it at a higher proportion than did *Lancet*. Finally, *Lancet* authors residing in the United Kingdom tended to cite U.S. journal publications much more than *NEJM* authors residing in the United States cited U.K. publications. Inhaber and Alvo reported similar results in their study using an input-output system. Both U.K. authors publishing in *Lancet* and U.S. authors publishing in *NEJM* rarely cited other countries (non-U.S. or non-U.K.) [23].

Cronin's study of two educational psychology journals, one originating from the United States and one

Table 6
World journal count in the field of medicine comparing SERLINE and BLLD

	SERLINE (1986)		BLLD (1973)	
	Number	(%)	Number	(%)
Journals published in the United States	1,320	(34.19%)	1,216	(21.24%)
Journals published in the United Kingdom	468	(12.12%)	504	(8.80%)
Journals published in other countries	2,073	(53.69%)	4,006	(69.96%)
Total	3,861	(100%)	5,726	(100%)

originating from the United Kingdom [24], resulted in percentages much like those found in Table 5. Although percentages are not the same, the values follow the same pattern. Cronin's study, like the one undertaken in this article, demonstrated that authors writing in select journals published in the United States and the United Kingdom cited material published in their own country more than any other country. These authors rarely cited non-U.S. or non-U.K. published materials. Finally, both studies demonstrated that U.K. authors cited U.S. material much more than the U.S. authors cited U.K. material. Health-professional and educational-psychologist publishing in the studied journals appear to have similar citation practices.

Both studies demonstrated that U.K. authors cited U.S. material much more than the U.S. authors cited U.K. material.

As discussed previously, one must employ a citation-to-publication ratio to measure national bias. The percentage of journals produced by a country is one of the required values in this ratio. World journal counts based on information derived from Carpenter and Narin [25] and numbers retrieved from SERLINE are shown in Table 6. Carpenter and Narin divided medical journals covered by BLLD into two distinct groups, clinical medicine and biomedical research. For the purposes of this study, these two groups were merged into one.

The calculated percentages for BLLD and SERLINE were not very different. However, values calculated using SERLINE suggested that a higher percentage of material was published in the United States and the United Kingdom when compared with BLLD values. SERLINE was developed by the National Library of Medicine (NLM) in the United States. This might explain the higher percentage of material from the United States. Because it was developed in and for the health community in the United States, NLM's holdings might reflect a form of national bias. One would assume that this could also be true for BLLD. Since BLLD was created in the United Kingdom, its

collection might reflect a bias toward that country. The values found in the BLLD study, however, did not support this theory. Many more world journal counts using other sources must be computed to test that type of hypothesis.

Journals indexed in *SCI* were used to analyze world journal counts in the 1981 study reported by Carpenter and Narin [26]. The percentages reported in their study are remarkably similar to those found in SERLINE. (Percentages from *SCI* can be found in Table 3 of this report.)

With the percentages of journals published in the United States, the United Kingdom, and other countries, it was possible to calculate citation-to-publication ratios. As mentioned earlier, a citation-to-publication ratio is the percentage of world citations going to a specific country divided by the percentage of world papers produced by that same country. Instead of world citations going to a country, this research used the percentage of citations from a population of articles published in *NEJM* and *Lancet*. Using percentages derived from *NEJM* and *Lancet* and dividing them by the percentages of world journal counts (represented by values from BLLD and SERLINE) yielded citation-to-publication ratios (Table 7).

A value of 1 would be the optimum. At this value, the percentage of citations to a country and the percentage of journals produced in that country would be equal. In other words, if the percentage of authors publishing in *NEJM* cite U.K. authors 20% of the time and the United Kingdom produces 20% of the medical journals published in the world, then 20 divided by 20 equals 1. The value 1 indicates that citation practices are not biased in any way. It is clear from Table 7 that both U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet* cite non-U.S. and non-U.K. journals much less than the amount of material produced by these other countries would indicate.

The values for cross-citations between the United States and the United Kingdom are much closer to 1. Values calculated using SERLINE are closer to 1 than those calculated using BLLD. But finally, whether using SERLINE or BLLD, the values produced are close to what would be expected for nonbiased cross-citation practices between the two countries.

The most striking results of this study were the

Table 7
Citation-to-publication ratios

From	To	Using SERLINE	Using BLLD
NEJM	United States	2.36	3.80
Lancet	United Kingdom	4.28	5.90
NEJM	United Kingdom	0.93	1.28
Lancet	United States	1.19	1.92
NEJM	Other	0.15	0.12
Lancet	Other	0.14	0.10

values for self-citation among nations. Both U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet* cited journal articles published in their own country much more than would be expected. U.K. authors publishing in *Lancet*, citing other U.K. material, have higher citation-to-publication ratios than any other group. This is true whether using journal counts from SERLINE or journal counts from BLLD.

U.S. authors publishing in NEJM and U.K. authors publishing in Lancet cited journal articles published in their own country much more than would be expected.

CONCLUSIONS

It is clear that in the articles used in this study, U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet* tended to cite journals from their own country more than those of any other countries. In addition, they rarely cited journals published in

Authors cited material produced in non-U.S. and non-U.K. countries far less than the amount of material produced by these countries would indicate.

countries other than the United States and the United Kingdom. When a citation-to-publication ratio was employed, these observations did not change. U.S. authors publishing in *NEJM* and U.K. authors publishing in *Lancet* still appeared to cite journals from their own country more than would be warranted by the amount of material produced in these countries. Again, these authors cited material produced in non-U.S. and non-U.K. countries far less than the amount of material produced by these countries would indicate.

NEJM and *Lancet* are only two medical journals out of hundreds published in the United States and the United Kingdom, but they are so highly cited and so highly respected that citation practices of U.S. and U.K. authors appearing in these two journals can be considered significant. Whether these practices are representative of other authors publishing in other journals is another matter. Many more studies are needed to properly gauge possible national bias as a whole in the health community.

REFERENCES

1. BROOKS TA. Private acts and public objectives: an investigation of citer motivations. *J Am Soc Inf Sci* 1985 Jul;36(4):226.
2. *IBID.*, 228.
3. KAPLAN N. The norms of citation behavior: prolegomena to the footnote. *Am Doc* 1965 Jul;16(3):181.
4. CRONIN B. Norms and functions in citation: the view of journal editors and referees in psychology. *Soc Sci Inf Stud* 1982;2:71.
5. *IBID.*
6. *IBID.*, 74-75.
7. INHABER H, ALVO M. World science as an input-output system. *Scientometrics* 1978 Sep;1(1):43-64.
8. CRONIN B. Transatlantic citation pattern in educational psychology. *Educ Libr Bull* 1981;24(2):48-51.
9. *IBID.*, 49.
10. CARPENTER MP, NARIN F. The subject composition of the world's scientific journals. *Scientometrics* 1980 Jan;2(1):53-63.
11. BARR KP. Estimates of the number of currently available scientific and technical periodicals. *J Docum* 1967 Jun;23(2):110-6.
12. CARPENTER MP, NARIN F. The adequacy of the *Science Citation Index (SCI)* as an indicator of international scientific activity. *J Am Soc Inf Sci* 1981 Nov;32(6):430-9.
13. *IBID.*
14. FRAME JD, NARIN F. The international distribution of biomedical publications. *Fed Proc* 1977 May;36(6):1792.
15. *IBID.*
16. Journal citation reports, *Science Citation Index*. Philadelphia, PA: Institute for Scientific Information, 1986.
17. GARFIELD E. Citation analysis as a tool in journal evaluation. *Science* 1972 Nov;178(4060):476.
18. Journal citation reports, op. cit., 83.
19. *IBID.*, 266.
20. NATIONAL LIBRARY OF MEDICINE. Health sciences serials. Bethesda, MD: The Library, 1986 Apr.
21. FRAME, op. cit., 1790-5.
22. CARPENTER, The subject.
23. INHABER, op. cit.
24. CRONIN, Transatlantic.
25. CARPENTER, The adequacy.
26. *IBID.*

Received December 1989; accepted April 1990