Open Access Citation Rates and Developing Countries

Michael Norris[1]; Charles Oppenheim[1]; Fytton Rowland[2]

[1]Department of Information Science, Loughborough University,
Loughborough, Leicestershire LE11 3TU, UK
e-mail: M.Norris2@lboro.ac.uk; C.Oppenheim@lboro.ac.uk
[2]73 Dudley Street, Bedford MK40 3TA, UK
e-mail: fytton@googlemail.com

Abstract

Academics, having written their peer reviewed articles, may at some stage in the make their work Open Access (OA). They can do this by self-archiving an electronic version of their article to a personal or departmental web page or to an institutional or subject repository, such that the article then becomes freely available to anyone with Internet access to read and cite. Those authors who do not wish to do this may leave their article solely in the hands of a toll access (TA) journal publisher who charges for access, consigning their article to remain behind a subscription barrier. Lawrence (2003), in a short study, noted that conference articles in computer science that were freely available on the World Wide Web were more highly cited that those that were not. Following this, there have been a number of studies which have tried to establish whether peer-reviewed articles from a range of disciplines which are freely available on the World Wide Web, and hence are OA, accrue more citations than those articles which remain behind subscription barriers (Antelman 2004, Davis and Fromerth 2007, Eysenbach 2006, Harnad and Brody 2004, Kurtz and Henneken 2007, Moed 2007). These authors generally agree that there is a citation advantage to those articles that have been made OA, but are either uncertain about, or find that they cannot agree on, the cause of this advantage. The causes of this citation advantage could simply be that OA articles are available well in advance of formal publication, and so have a longer period in which to accrue citations, or simply that more authors, because they are freely available, can read and cite them. As part of this debate, Smith (2007) asked whether authors from developing countries might contribute to higher citation counts by accessing OA articles and citing them more readily than TA articles. As part of a larger study of the citation advantage of OA articles (Norris, Oppenheim and Rowland 2008), research was undertaken to see whether a higher proportion of citations to OA articles came from authors based in countries where funds for the purchase of journals are very limited. Mathematics was chosen as the field to be studied, because no special programme for access in developing countries, such as HINARI (2007), covers this subject. The results show that the majority of citations were given by Americans to Americans, but the admittedly small number of citations from authors in developing countries do seem to show a higher proportion of citations given to OA articles than is the case for citations from developed countries. Some of the evidence for this conclusion is, however, mixed, with some of the data pointing toward a more complex picture of citation behaviour.

Keywords: Open Access; Citation advantage; Developing countries

1. Introduction

One of the basic arguments for OA is that those who cannot afford access to peer-reviewed journal articles could access them if the authors of these articles self-archived their work somewhere on the World Wide Web. It should follow that a higher percentage of those who cite these OA articles ought to come from countries where access to expensive journals is limited. A number of schemes, such a HINARI

(2007) and AGORA (2007), exist to provide access to scholarly information inexpensively to users in developing countries, but not all disciplines are covered by these. In the overall larger study (Norris, Oppenheim and Rowland 2008), four subjects (sociology, economics, ecology and mathematics) were selected, and a large number of papers were investigated to discover whether they were available on an OA basis anywhere. Citation data on all these papers were collected and subjected to statistical analyses of various kinds to establish whether or not OA availability of itself correlates with a greater number of citations to an article. As part of the larger study, mathematics – which is not covered by any of the assistance schemes – was chosen for an investigation of citation of articles by authors based in developing countries, the hypothesis being that these authors would be unlikely to have access to expensive toll access (TA) journals.

2. Methods

In the main project, articles were selected from high-impact journals and their OA status was sought by using various search tools (OAIster 2007, OpenDOAR 2007, Google Scholar, and finally Google), and their availability or non-availability with OA was noted. Citations to them were then retrieved by using the ISI Web of Science databases.

The country of origin of cited and citing articles was decided by the first author's affiliation. Countries were classified by their per capita income using the World Bank's (2007) categories (see Table 6), and were also grouped by their geographical location into twelve groups (see Table 1). Citation ratios for the TA group and the OA group of articles were calculated separately.

			OA		
			Toll Access	Open Access	Total
	Spain	Count	12	14	26
		% within Region	46.2%	53.8%	100.0%
	Japan	Count	16	9	25
		% within Region	64.0%	36.0%	100.0%
	Italy	Count	20	22	42
		% within Region	47.6%	52.4%	100.0%
	Germany	Count	18	47	65
		% within Region	27.7%	72.3%	100.0%
	France	Count	26	39	65
		% within Region	40.0%	60.0%	100.0%
	Canada	Count	11	16	27
		% within Region	40.7%	59.3%	100.0%
	Pacific Rim	Count	9	22	31
		% within Region	29.0%	71.0%	100.0%
	China	Count	15	11	26
		% within Region	57.7%	42.3%	100.0%
	Rest of World	Count	20	18	38
		% within Region	52.6%	47.4%	100.0%
	UK	Count	18	28	46
		% within Region	39.1%	60.9%	100.0%
	Europe	Count	31	65	96
		% within Region	32.3%	67.7%	100.0%
	USA	Count	102	204	306
		% within Region	33.3%	66.7%	100.0%
Total		Count	298	495	793
		% within Region	37.6%	62.4%	100.0%

Table 1. Cited articles by region and OA status

3. Data

In the overall sample, 1158 mathematics journal articles were taken from 16 high impact journals. Only citation links from other-author citations were counted; all other types of author and journal self-citations were discarded. After this, 365 of the articles were then uncited, leaving 793 articles cited by other authors. Table 1 shows how these 793 were distributed amongst the twelve regions.

All of the citation links to the remaining 793 articles, which totalled 3032, were then analysed. These 3032 citations were from 2680 citing articles; clearly, in some cases, there were multiple citations from some of the citing articles. Table 2 shows how the 3032 citations from the 2680 citing articles that cited the original 793 were broken down. For example, there were 2413 citing articles (80%) that cited just one of the 793 articles, whereas there were three articles which cited six of the original articles each. The first-author affiliations of the original 793 cited articles covered 47 countries. The first-author affiliation of the 2680 citing articles, citing the 793 articles, were drawn from 70 countries; 23 of these were necessarily in addition to the initial 47 countries. The cited and citing countries were classified by their per capita income using The World Bank's (2007) system of classification. China, for example, is designated as being in the lower middle-income group of countries and India in the low-income bracket, whereas most of Western Europe and North America are in the high-income group of countries. To further aid analysis and comparison, the original 47 countries and the 70 citing-author countries were classified by location into USA, Canada, France, Germany, Italy, Japan, Spain, UK, rest of Continental Europe, China, Pacific Rim, and the Rest of World.

Frequency of Citation	Citing Articles	Overall Citations
1	2413	2413
2	208	416
3	43	129
4	9	36
5	4	20
6	3	18
Totals	2680	3032

Table 2. Frequency of Citation

In Table 3 the 3032 citations are shown by their cited article OA or TA status, the region from which they were cited, and whether the cited article was matched by a citation from the same region. By way of illustration, 231 TA citations came from the USA, but only 115 of these were from articles that were originally authored by first-author affiliation from that territory, hence the other 116 were from other regions.

Table 4 takes the data from Table 3 a step further and shows the citing country and the income group of the related cited articles. The 231 citations from the USA to the TA cited articles are shown by the World Bank per-capita income group from which they came, by first-author affiliation. It is evident that 115 of them were from the USA, as shown in Table 3, but overall only 20 of the 231 were from regions outside the high per-capita income bracket. It is noticeable at this stage that a greater percentage of the TA cited articles (13.50%) are being cited outside of the high per-capita income bracket than OA articles, which account for only 4.7%.

The cluster bar chart in Figure 1 further extends the data from Table 4 by giving a comparative percentage

Count

OA			Region to region match		Total
			no match	match	
Toll Access	Region	USA	116	115	231
		Rest of Europe	104	16	120
		UK	48	12	60
		Rest of World	50	5	55
		China	70	20	90
		Pacific Rim	44	7	51
		Canada	26	1	27
		France	45	14	59
		Germany	59	11	70
		Italy	65	16	81
		Japan	28	24	52
		Spain	37	2	39
	Total		692	243	935
Open Access	Region	USA	251	359	610
		Rest of Europe	266	55	321
		UK	97	17	114
		Rest of World	99	5	104
		China	133	7	140
		Pacific Rim	114	10	124
		Canada	55	5	60
		France	108	29	137
		Germany	168	46	214
		Italy	93	27	120
		Japan	84	6	90
		Spain	60	3	63
	Total		1528	569	2097

Table 3. Citations by author country

of the whole count for each category by the OA status of the citations from each region. The USA, for example, receives 24.71% of all the citations to given to TA articles (231/935) and 29.09% of all those to given to OA articles (610/2097). Of the twelve regions, only five receive a greater overall percentage of the OA than of the TA citations, even though in every case each region had a greater number of OA than TA citations. It is noticeable that China receives 9.6% (90/935) of all citations given to TA articles but only 6.7% (140/2097) of all citations to OA articles. The Rest of World shows a similar but narrower disparity: 5.9% (55/935) of all citations to TA articles but 5.0% (104/2097) of all those to OA articles.

Table 5 shows the distribution of citations to cited articles by their OA status and per-capita income group. What is evident is that there is a greater percentage of citations to the TA articles (20.00%, 187/935) from the low to upper middle income groups than is the case for OA articles, where the comparable group is 15.40% (323/2097) of all the citations to the OA articles.

Table 6 shows the distribution of the original 793 cited articles and the distribution of the 3032 citations by the country of their first-author affiliation. Countries are again classified by their World Bank per-capita income grouping. The number of articles appearing in each category has been given by its occurrence and the ratio is the division of the citing articles by the cited articles. The numbers in brackets indicate the number of article records in each category.

OA				Cited Income				Total
				Low	Lower middle	Upper middle	High	
Toll Access	Citing Country	USA	Count	0	17	3	211	231
			% within Region	.0%	7.4%	1.3%	91.3%	100.0%
		Rest of Europe	Count	0	7	10	103	120
			% within Region	.0%	5.8%	8.3%	85.8%	100.0%
		UK	Count	0	2	0	58	60
			% within Region	.0%	3.3%	.0%	96.7%	100.0%
		Rest of World	Count	0	3	9	43	55
			% within Region	.0%	5.5%	16.4%	78.2%	100.0%
		China	Count	3	21	6	60	90
			% within Region	3.3%	23.3%	6.7%	66.7%	100.0%
		Pacific Rim	Count	0	5	2	44	51
			% within Region	.0%	9.8%	3.9%	86.3%	100.0%
		Canada	Count	0	5	1	21	27
			% within Region	.0%	18.5%	3.7%	77.8%	100.0%
		France	Count	0	2	2	55	59
			% within Region	.0%	3.4%	3.4%	93.2%	100.0%
		Germany	Count	0	5	1	64	70
			% within Region	.0%	7.1%	1.4%	91.4%	100.0%
		Italy	Count	2	6	3	70	81
			% within Region	2.5%	7.4%	3.7%	86.4%	100.0%
		Japan	Count	0	3	1	48	52
			% within Region	.0%	5.8%	1.9%	92.3%	100.0%
		Spain	Count	0	3	4	32	39
			% within Region	.0%	7.7%	10.3%	82.1%	100.0%
	Total		Count	5	79	42	809	935
			% within Region	.5%	8.4%	4.5%	86.5%	100.0%
Open Access	Citing Country	USA	Count	0	14	12	584	610
			% within Region	.0%	2.3%	2.0%	95.7%	100.0%
		Rest of Europe	Count	0	11	7	303	321
			% within Region	.0%	3.4%	2.2%	94.4%	100.0%
		UK	Count	0	3	2	109	114
			% within Region	.0%	2.6%	1.8%	95.6%	100.0%
		Rest of World	Count	1	5	3	95	104
			% within Region	1.0%	4.8%	2.9%	91.3%	100.0%
		China	Count	0	7	4	129	140
			% within Region	.0%	5.0%	2.9%	92.1%	100.0%
		Pacific Rim	Count	0	12	0	112	124
			% within Region	.0%	9.7%	.0%	90.3%	100.0%
		Canada	Count	0	0	0	60	60
			% within Region	.0%	.0%	.0%	100.0%	100.0%
		France	Count	0	1	1	135	137
			% within Region	.0%	.7%	.7%	98.5%	100.0%
		Germany	Count	0	3	5	206	214
			% within Region	.0%	1.4%	2.3%	96.3%	100.0%
		Italy	Count	0	1	3	116	120
			% within Region	.0%	.8%	2.5%	96.7%	100.0%
		Japan	Count	1	1	1	87	90
			% within Region	1.1%	1.1%	1.1%	96.7%	100.0%
		Spain	Count	0	0	0	63	63
			% within Region	.0%	.0%	.0%	100.0%	100.0%
	Total		Count	2	58	38	1999	2097
			% within Region	.1%	2.8%	1.8%	95.3%	100.0%

Table 4 Citing country to cited income group

4. Results and Discussion

Overall there is a tendency for authors to cite work from their own country preferentially. Table 3 shows all the citations, analysed by whether there was a match or not between the nationality of the authors of the citing work and of the cited work. Generally, for every one citation that can be paired by country to the

article it is citing, there are three that do not match. This applies to both OA and TA articles. Clearly, however, these data are skewed by the predominance of the region-to-region match for the USA. Given that just over 25% of the citations come from this territory alone, it is not surprising that of all citations almost 42% are to USA-affiliated first authors (data not shown). Perhaps this result is unremarkable, given that a large proportion (38.6%) of the cited articles originate from the USA, and that of all the citations from each region, the largest number are given to USA-affiliated authors. The 230 citations made by Chinese authors accounted for about 10% of all the citations to TA articles and about 7% of all the citations to OA articles. For the Pacific Rim and the Rest of the World territories, the overall TA/OA citation percentages were barely different, at around 5% each. This result appears to confirm the findings from the analysis by per capita income, that is, that there is little evidence to suggest that authors who live in countries that may have difficulties accessing TA journals are citing OA articles in greater numbers and hence boosting the citation count. Figure 1 shows that seven out of the twelve regions have more citations to TA than to OA articles. The seven regions include the UK, Italy, Japan, Spain, China and the Rest of the World, the latter two helping to support the premise that low-income does not generate exceptional OA citations. It is noticeable also, as demonstrated by Table 3, that the regional link between cited and citing article by first author affiliation is generally weak, once the USA has been excluded. For citations of either access status, OA or TA, the overall regional match is about a quarter, but noticeably in the case of China, this is heavily skewed in favour of not citing other Chinese-affiliated authors.

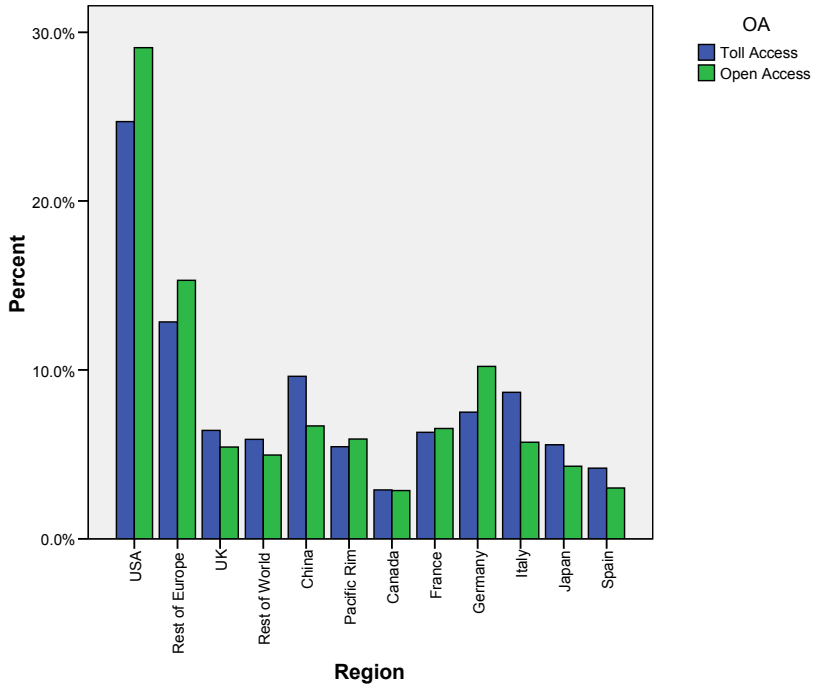

Figure 1 Percentage of citations to cited articles by OA status

Whilst there is unmistakable evidence from the data collected here that there is an overall citation advantage to those articles that are made available as OA (20%), the actual causes of this advantage, here and in other studies, are not always clear. Given that one of the primary arguments in favour of OA is that those who cannot afford access to peer-reviewed journal articles could use them if these articles were self-archived on the World Wide Web, where they could be readily accessed and cited by those with limited incomes. Hence, it could be reasoned that if this were clearly so, that a demonstrable cause of any citation advantage could be shown. Table 6 shows for the TA articles, the highest ratio of citing to cited articles occurs for citing authors in those countries in the lower middle income bracket, regardless of the nationality

of origin of the cited articles. If all but the high income level countries are taken together, then the citation ratio is 4.45 for the TA articles and 9.79 for the OA articles. However, the overwhelming majority of articles are both authored and cited from the high-income countries. Table 6 gives the full data, divided into TA and OA articles, and by the four income categories of the first author's country of residence. Although this appears to be a convincing advantage, the percentage of the 187 lower-income citations to TA articles of all 935 citations to TA articles the result is 20%, and this is greater than the 15.40% figure from the comparable calculation for citations to OA articles. So there is a greater percentage of lower income authors among those citing TA articles than among those citing OA articles.

OA				Cited Country Income				Total
				Low	Lower middle	Upper middle	High	
Toll Access	Citing Country Income	Low	Count	0	1	0	5	6
			% within Citing Country Income	.0%	16.7%	.0%	83.3%	100.0%
		Lower middle	Count	3	22	8	73	106
			% within Citing Country Income	2.8%	20.8%	7.5%	68.9%	100.0%
		Upper middle	Count	0	6	15	54	75
			% within Citing Country Income	.0%	8.0%	20.0%	72.0%	100.0%
		High	Count	2	50	19	677	748
			% within Citing Country Income	.3%	6.7%	2.5%	90.5%	100.0%
	Total		Count	5	79	42	809	935
			% within Citing Country Income	.5%	8.4%	4.5%	86.5%	100.0%
Open Access	Citing Country Income	Low	Count	0	1	0	16	17
			% within Citing Country Income	.0%	5.9%	.0%	94.1%	100.0%
		Lower middle	Count	0	10	5	165	180
			% within Citing Country Income	.0%	5.6%	2.8%	91.7%	100.0%
		Upper middle	Count	1	3	6	116	126
			% within Citing Country Income	.8%	2.4%	4.8%	92.1%	100.0%
		High	Count	1	44	27	1702	1774
			% within Citing Country Income	.1%	2.5%	1.5%	95.9%	100.0%
	Total		Count	2	58	38	1999	2097
			% within Citing Country Income	.1%	2.8%	1.8%	95.3%	100.0%

Table 5. Cited to citing articles by income group

Access Status	**World Bank classification by per capita income**			
TA Articles	Low	Lower middle	Upper middle	High
TA cited articles (298)	2	21	19	256
TA citing articles (935)	6	106	75	748
Ratio of citing to cited articles	3	5.05	3.95	2.92
OA Articles	Low	Lower middle	Upper Middle	High
OA cited articles (495)	1	14	18	462
OA citing articles (2097)	17	180	126	1774
Ratio of citing to cited articles	17	12.85	7	3.84

Table 6. Ratio of citing to cited articles by national income groups

Given that self-citations have been eliminated from the original portion of the data, it is authors citing the work of others to support their work, who are doing the citing. As can be seen in Table 5, however, most of those who are doing the citing to the lower income groups are from the high-income countries. It is clear that a greater percentage of authors from the low and lower income countries cite more TA articles than OA articles, despite there being a higher ratio of citing to cited articles for those of OA status. In fact, 95.33% of all OA citations and 86.5% of TA citations are from high income regions.

5. Conclusions

Taken overall the results give a mixed picture as to whether those in the lower per capita income bracket countries are citing OA articles more frequently than TA ones.

The USA cites itself more than anyone else, which is not surprising given its level of authorship. The other developed countries, except for Japan, are all at about the same level in terms of within-nation citation. Table 6 suggests that while there is a modest difference between the citation ratios of OA and TA articles for citations given by authors in the developed world (3.84 versus 2.92), the difference becomes much greater when citations given by authors from the developing world are studied. The sample from the lowest income countries is very small. It may be that the lack of reliable telecommunications networks in these low income countries could hinder access to OA articles. In this case, scholars in these countries may rely on a limited number of printed journals for which they have subscriptions. The results from the larger sample in the lower middle income group of countries, however, are striking: a citation ratio of 12.85 for OA articles versus 5.05 for TA articles.

6. Notes and References

HINARI, 2007. <http://www.who.int/hinari/en/>, [accessed 16.02.07].
AGORA, 2007. <http://www.aginternetwork.org/en/>, [accessed 16.2.07].
OAIster, 2007. < http://www.oaister.org/>, [accessed 07.03.07].
OpenDOAR. 2007. <http://www.OpenDOAR.org/search.php>, [accessed 07.02.07].
The World Bank, 2007. <http://go.worldbank.org/K2CKM78CC0>, [accessed 08.08.07].

[1] Lawrence, S., 2001. Online or invisible. *Nature*, **411**(6837), 521.
[2] Antelman, K., 2004. Do open-access articles have a greater research impact? *College and Research Libraries,* **65**(5), 372-382.
[3] Eysenbach G., 2006. Citation advantage of open access articles. *PLoS Biology*, **4**(5), 692-698.
[4] Harnad, S. and Brody, T., 2004. Comparing the impact of OA (OA) vs. non-OA articles in the same journals. *D-Lib Magazine* [online], **10**(6), 1-5. <http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/june04/harnad/06harnad.html>, [accessed 22.12.05].
[5] Kurtz, M. & Henneken, E., 2007. *Open access does not increase citations for research articles from The Astrophysical Journal.* <http://arxiv.org/ftp/arxiv/papers/0709/0709.0896.pdf>, [accessed 07.09.07].
[6] Moed, H., 2007. The effect of 'Open Access' upon citation impact: An analysis of ArXiv's Condensed Matter Section. *Journal of the American Society for Information Science and Technology.* **58**(13), 2047-2054.
[7] Smith, J., 2007. Re: The apparent OA citation advantage. To multiple recipients of list *JISC Repositories*, 20 May, 19:19:35 BST.
[8] Norris, M., Oppenheim C. & Rowland, F. (in press). The citation advantage of open access articles. *Journal of the American Society for Information Science and Technology*.