Percentile Ranks and the Integrated Impact Indicator (I3)

Dear Sir:

In a brief communication, Rousseau (in press) proposed to redefine percentile classes by counting as follows: $x_i \leq x$; $i = 1,\ldots, n$; instead of using $x_i < x$; $i = 1,\ldots, n$. According to him "this completely solves the issue raised by Leydesdorff & Bornmann (in press) that in the case of small numbers (e.g., reviews), papers would for arithmetic reasons have lower percentile values." Leydesdorff & Bornmann (2011, at p. 2137) proposed to add 0.9 to the count arguing as follows: "For example, if a journal with many articles publishes only 10 reviews each year, the highest possible percentile of reviews within this set would be the 90^{th} (i.e., 9 of 10) whereas this could be the 99^{th} (i.e., 9.9 of 10) when 0.9 is added to the count."

Since the counting rule employed for computing percentile values is not uniquely determined (Hyndman & Fan, 1996), we accepted Rousseau's suggestion as a further improvement, implemented it into the program for computing *I3* in Web-of-Science data (at http://www.leydesdorff.net/software/i3), and began to use it in a recent study (Bornmann & Leydesdorff, in press). However, Zhou *et al.* (in preparation) noted that in the case a set of nine uncited papers and one with citation, the uncited papers would all be placed in the 90^{th} percentile rank. A lowly-cited document set would thus be advantaged when compared with a highly-cited one. Rousseau (*personal communication*, Dec. 23, 2011) thereupon suggested disregarding the zero-counts. We followed this suggestion and re-analyzed the data studied by Leydesdorff &

Bornmann (2011). In Table 1, we use the values provided in their Table 4 (at p. 2139) for the comparison.

Table 1: Rankings of 15 LIS journals with highest values on *I3* (expressed as percentages of the sum) compared with *IF*s, total citations, and % *I3*(*6PR*) with different calculation rules for the percentiles.

Journal	% I3 (L&B)* (a)	%I3 (Rousseau)** (b)	% I3 (quantiles) (c)
J Am Soc Inf Sci Technol	9.72 [1]	7.32 [2]	9.73 [1]
Scientometrics	7.23 [2]	5.20 [4]	7.24 [2]
J Amer Med Inform Assoc	6.80 [3]	4.53 [5]	6.80 [3]
Inform Process Manage	6.14 [4]	4.41 [6]	6.14 [4]
Inform Management	4.01 [5]	2.63 [7[	4.01 [5]
Int J Geogr Inf Sci	3.14 [6]	2.32 [8]	3.14 [6]
MIS Quart	2.61 [7]	1.61 [21]	2.61 [7]
J Manage Inform Syst	2.60 [8]	1.76 [15]	2.60 [8]
J Health Commun	2.52 [9]	1.80 [14]	2.51 [9]
J Acad Libr	2.50 [10]	2.15 [9]	2.51 [10]
J Inform Sci	2.43 [11]	1.88 [12]	2.43 [11]
J Comput-Mediat Commun	2.37 [12]	1.89 [11]	2.37 [12]
J Informetr	2.28 [13]	1.49 [24]	2.28 [13]
J Med Libr Assoc	2.21 [14]	1.97 [10]	2.21 [14]
Telecommun Policy	2.15 [15]	1.74 [18]	2.15 [15]

* Source: Leydesdorff & Bornmann (2011); ** cf. Rousseau (in press).

The differences between the quantiles and our correction with +0.9 are only in the second decimal of the percentages and negligible ($r = 1.00$; $p < 0.01$; $N = 65$). However, the differences with the values provided in column *b* based on the normalization suggested by Rousseau are considerable. For example, *JASIST* would lose its first position in this ranking to *The Scientist*. Indeed, *The Scientist* contained 392 citable items in 2008 and 2009, of which 352 (98.1%) were not cited at the time of this download (February 2011). Using Rousseau's counting rule, the

journal obtains a %*I3* of 7.50. Using quantiles or the correction of +0.9, *The Scientist* is rated 33rd with a %*I3* of 1.00.

In summary, we regret with hindsight our suggestion to deviate from quantiles (however computed) as a basis for the ranking because Rousseau's contribution makes clear that we may have opened a parameter space of other possibilities. The mathematical discussion about other possibilities easily obscures our message that one is not allowed nor does one have to use central tendency statistics for analyzing citation distributions (Seglen, 1992).

I3 is an *impact* indicator which can be used as an alternative to parametric statistics such as the ratio of citations over publications (*c/p*) or the Impact Factor. More recently, both the *SCImago Institutions Rankings* and the *Leiden Ranking 2011/2012* used the top-10% most-highly cited papers as a non-parametric *excellence* indicator (cf. Bornmann & Leydesdorff, 2011). As Rousseau (in press) mentions these excellence indicators can be considered as special cases of *I3*: only two percentile rank classes are distinguished for the evaluation. Both excellence and impact indicators can be tested statistically using the *z*-test for independent proportions (Bornmann *et al.*, 2011).

Loet Leydesdorff [1] and Lutz Bornmann [2]

[1] University of Amsterdam, Amsterdam School of Communication Research (ASCoR), Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands; loet@leydesdorff.net.
[2] Max Planck Society, Administrative Headquarters, Hofgartenstrasse 8, D-80539 Munich, Germany; bornmann@gv.mpg.de.

References

Bornmann, L., de Moya-Anegón, F., & Leydesdorff, L. (in press). The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *Journal of Informetrics*. Retrieved on Dec. 25, 2011 from http://arxiv.org/abs/1110.2305.

Bornmann, L., & Leydesdorff, L. (2011). Which cities produce excellent papers worldwide more than can be expected? A new mapping approach—using Google Maps—based on statistical significance testing. *Journal of the American Society for Information Science and Technology, 62*(10), 1954-1962.

Bornmann, L., & Leydesdorff, L. (in press). Which cities' paper output and citation impact are above expectation in information science? Some improvements of our previous mapping approaches. *Journal of Informetrics*.

Hyndman, R. J., & Fan, Y. (1996). Sample quantiles in statistical packages. *American Statistician*, 361-365.

Leydesdorff, L., & Bornmann, L. (2011). Integrated Impact Indicators (I3) compared with Impact Factors (IFs): An alternative design with policy implications. *Journal of the American Society for Information Science and Technology, 62*(11), 2133-2146.

Rousseau, R. (in press). Basic properties of both percentile rank scores and the I3 indicator. *Journal of the American Society for Information Science and Technology*, doi: 10.1002/asi.21684.

Seglen, P. O. (1992). The Skewness of Science. *Journal of the American Society for Information Science, 43*(9), 628-638.

Zhou, P., Ye, F. Y., & Zhong, Y. (in preparation). Percentage Citation Impact and the Combined Impact Indicator: Another perspective for measuring academic impact.