Accepted for publication in the
Journal of the Association for Information Science and Technology

An Empirical Look at the Nature Index

Lutz Bornmann* & Robin Haunschild**

* Corresponding author:

Division for Science and Innovation Studies

Administrative Headquarters of the Max Planck Society

Hofgartenstr. 8,

80539 Munich, Germany.

E-mail: bornmann@gv.mpg.de

** Max Planck Institute for Solid State Research

Heisenbergstr. 1,

70569 Stuttgart, Germany.

E-mail: R.Haunschild@fkf.mpg.de

Abstract

In November 2014, the Nature Index (NI) was introduced (see http://www.natureindex.com) by the Nature Publishing Group (NPG). The NI is comprised of the primary research articles published in the past 12 months in a selection of reputable journals. Starting from two short comments on the NI (Haunschild & Bornmann, 2015a, 2015b), we undertake an empirical analysis of the NI using comprehensive country data. We investigate whether the huge efforts of computing the NI are justified and whether the size-dependent NI indicators should be complemented by size-independent variants. The analysis uses data from the Max Planck Digital Library in-house database (which is based on Web of Science data) and from the NPG. In a first step of analysis, we correlate the NI with other metrics which are simpler to generate than the NI. The resulting very large correlation coefficients point out that the NI produces very similar results as simpler solutions. In a second step of analysis, relative and size-independent variants of the NI are generated which should be additionally presented by the NPG. The size-dependent NI indicators favor large countries (or institutions) and the top-performing small countries (or institutions) do not come into the picture.

1 Introduction

Currently, there exist five major international university rankings worldwide: (1) Academic Ranking of World University (ARWU), (2) Times Higher Education World University Rankings (THE Rankings), (3) QS World University Rankings, (4) Leiden Ranking by the Centre for Science and Technology Studies (CWTS), and (5) SCImago Institutions Ranking. Whereas the first three rankings use very different indicators to rank universities, the last two use a set of bibliometric indicators only. An overview over these (and other university rankings) can be found in several publications: Safón (2013), Dill and Soo (2005), Buela-Casal, Gutiérrez-Martínez, Bermúdez-Sánchez, and Vadillo-Muñoz (2007), Aguillo, Bar-Ilan, Levene, and Ortega (2010), Rauhvargers (2011). These overviews describe – among other things – the different indicators used and compare the ranking results based on the different ranking methods.

In November 2014, the Nature Index (NI) was introduced (see http://www.natureindex.com) (Campbell & Grayson, 2014). According to Campbell and Grayson (2015) the Nature Publishing Group (NPG) "does not intend the Nature Index to be a ranking and have quite deliberately not referred to it as such anywhere" (p. 1831). However, the NI allows exactly this: a ranking of worldwide institutions and countries based on their publication output in selected journals (see http://www.natureindex.com/country-outputs and http://www.natureindex.com/institution-outputs). Thus, the NI should be discussed against other possibilities of measuring the performance of institutions or countries. The aim of this paper is to empirically study the NI as a new ranking method in more detail and to provide possible improvements. In a first step of analysis, we correlate the NI with other metrics which are simpler to generate than the NI. In a second step, relative and size-independent variants of the NI are generated which should be additionally presented by the NPG.

2 The Nature Index

The NI is comprised of the *primary research articles* published in the past 12 months in a selection of reputable journals. The list of 68 journals is the result of 68 panelists being asked to name the journals (maximal n=10) in which they would want to publish their best research articles. Two panel chairs signed off on the final list. A confirmation attempt was made by sending out an online questionnaire to 100,000 scientists in the life, physical, and medical sciences. They also were asked to name their 10 most preferred journals. The panel chairs recorded a response rate of 2.8%. Overall, Campbell and Grayson (2014) reported "a high degree of convergence between the panel and survey outputs for the most popular journals" (p. S52). The NI contains three quantities: the raw article count (AC), the fractional count (FC), and the weighted fractional count (WFC). The AC is obtained from counting all *primary research articles* published in the past 12 months in the NI journals. The fractional count weights the individual *primary research article* according to the number of co-authors (e.g.: if three scientists from the USA and two scientists from Japan published one paper in a NI journal, this paper is counted as 3/5 for the USA and 2/5 for Japan). The WFC is supposed to account for the fact that papers from journals in the field astronomy and astrophysics are approximately five times as numerous as papers from other fields in the NI. Therefore, papers from the field astronomy and astrophysics are weighted with a coefficient of 0.2.

Recently, we started the discussion about the NI with two comments (Haunschild & Bornmann, 2015a, 2015b). One of our discussion starters has sparked a reply (Campbell & Grayson, 2015).

3 Methods

On February 10, 2015, we saved the country tables from the NI for the publication period 01 January 2014 – 31 December 2014. In June 2015, the NPG published the NI supplement 2015 including a table with the top 100 countries

(http://www.nature.com/nature/journal/v522/n7556_supp/full/522S34a.html). The comparison of both tables shows small deviations. For example, in February the USA had an AC=26,631 and in June an AC=26,638. The NPG seems to continuously update their older data. The deviations between both points in time are so small for all countries that we decided to use the data from February. A further advantage of this data is that the FC is included (in the NI supplement table it is not).

Reference data for the full publication output were taken from the in-house database of the Max Planck Society (MPG) based on the WoS and administered by the Max Planck Digital Library (MPDL). Since we do not have reliable data in the database on the institutional, but on the country level, the study focuses on countries. Also, the period of analysis is 01 January 2014 – 31 December 2014.

In this study only countries are considered with at least n=1000 papers and an AC of at least 30. Although the study focusses on one year only (2014), we would like to produce results which can be generalized for neighboring years. In case of small publication sets for a country, large annual variations of indicator values can be expected (Levy & Lemeshow, 2008). Thus, one needs larger publications sets to obtain results which might be also valid for neighboring years.

4 Results

In section 4.1 we compare the NI with other metrics, namely more simple solutions. Costly generated metrics (as the NI) should not correlate very high with more simple solutions. Otherwise the efforts are not justified. In section 4.2 we recommend to complement the size-dependent results of the NI with size-independent results. For example, the Leiden Ranking (www.leidenranking.com) also provides both perspectives. The size-dependent results are mainly caused by the total publication output of a country or institution.

4.1 The comparison of NI with other metrics

The NPG made significant efforts to generate the NI. Two panels have been constituted which have selected the reputable journals. Further, a comprehensive survey has been performed to validate the selection of the panels. These huge efforts are justified only if the NI does not correlate with simpler solutions (metrics). In case of high correlations, one could question these efforts. In this study, we produced three metrics on the country level which are relatively easy to produce.

(1) Total number of papers (Np). Here, a country's number of papers with the WoS document type "article" is counted. For some journals, the NPG selects some articles as primary research articles. This is not reproducible in an automated manner in our database. We consistently obtain too high values for the AC, but we obtain a nearly perfect correlation ($r^2 = 0.9985$) between our AC and the official AC by the NPG.

(2) AvgAC: We generated five different random NIs. From each journal in the WoS which published papers in 2014 (N=12102), 68 journals were randomly selected. This procedure was repeated five times, resulting in five different (random) ACs for each country. This procedure gave rise to five random NIs with five different (random) ACs for each country. We computed the mean value over the five random ACs which yielded our (random) AvgAC.

(3) $Q1_{JIF}$: In the SCImago Institutions Ranking (SIR) an indicator is considered which reflects the reputation of the journals in which an institution has published. Q1 is the ratio of papers that an institution publishes in the most influential scholarly journals of the world. The most influential journals are those which ranked in the first quartile (25%) of their subject areas (journal sets) as ordered by the SCImago Journal Rank (SJR) indicator (Gonzalez-Pereira, Guerrero-Bote, & Moya-Anegon, 2010).

The Q1 indicator is a size-independent indicator. In order to produce a size-dependent indicator for this study, which can be correlated with the other size-dependent

indicators, we identified those papers for a country published in a first quartile ranked journal. Different from the SCImago Institutions Ranking, we used the Journal Impact Factor (JIF) instead of the SJR to select the journals belonging to the first quartile of their subject areas (Pudovkin & Garfield, 2004). Thus, we name the indicator $Q1_{JIF}$.

Table 1 presents the AC, FC, and WFC values of the NI for the year 2014. Additionally, the Np, $Q1_{JIF}$, and (random) AvgAC are included. For every indicator the corresponding rank numbers were generated. As the results in the table show the indicators lead to the same or similar ranking positions for several countries. For example, the USA and China are at the top positions independent of the used indicators. UK takes up the third or fourth position. However, it is also visible that the ranking positions of many countries differ to a larger extent. For example, Switzerland has the 18^{th} position if the countries are ranked according to the number of papers or random AvgAC. However, the country (significantly) improved its positions if the official NI indicators or the $Q1_{JIF}$ indicator is used.

Since the rank columns in Table 1 do not offer a clear picture of the relationship between the different indicators, we calculated correlation coefficients.

Table 1: Total number of papers (Np), $Q1_{JIF}$*, AvgAC, AC, FC, and WFC for countries in decreasing order of WFC. The rank positions for every country are added based on orderings by the different indicators.

Country	Total number of papers (Np)	Rank Np	$Q1_{JIF}$*	Rank $Q1_{JIF}$	AvgAC	Rank AvgAC	AC	Rank AC	FC	Rank FC	WFC	Rank WFC
United States of America (USA)	365437	1	201555	1	2189	1	26631	1	19765.08	1	17931.85	1
China	215731	2	90233	2	1011.8	2	8632	2	6323.43	2	6032.52	2
Germany	98154	4	53002	4	582	3	8576	3	4561.71	3	4016.03	3
United Kingdom (UK)	102766	3	56942	3	559.2	4	7590	4	3802.69	4	3250.06	4
Japan	76638	5	32281	7	347.8	7	4973	6	3428.46	5	3198.73	5
France	68382	6	37602	5	320	9	5241	5	2586.78	6	2220.89	6
Canada	60340	7	32401	6	370.4	6	3224	7	1659.27	7	1488.07	7
Switzerland	25979	18	15729	14	138.2	18	2709	10	1371.36	9	1287.8	8
South Korea	50951	11	19730	12	287.4	11	1966	13	1232.24	11	1167.49	9
Spain	53876	9	28237	9	303	10	2896	9	1349.39	10	1090.7	10
Italy	58707	8	29305	8	372.2	5	3051	8	1403.6	8	1046.92	11
Australia	50495	12	26740	10	327.4	8	2497	11	1145.32	12	951.22	12
India	51751	10	16200	13	253.2	12	1483	14	1029.92	13	921.77	13
Netherlands	35394	14	21666	11	203.8	14	2188	12	902.53	14	756.33	14
Singapore	11287	30	6732	22	56	33	873	21	521.4	17	520.6	15
Sweden	23716	20	13463	15	123.8	21	1407	15	570.81	15	514.55	16
Israel	12491	26	6604	23	72.4	26	1012	19	546.67	16	492.44	17
Taiwan	27586	16	13043	16	126.2	20	887	20	520.18	18	481.36	18
Russia	28981	15	6342	24	215	13	1145	16	451.61	19	370.08	19
Belgium	19704	22	10839	18	96.6	23	1082	17	426.9	20	348.91	20
Denmark	14944	23	8861	19	102.8	22	1036	18	363.1	21	321.26	21

Austria	13384	24	7281	20	80	24	856	22	350.86	22	318.98	22
Brazil	38309	13	11791	17	175.2	15	715	24	315.22	23	237.01	23
Poland	22872	21	6979	21	140.4	17	685	25	275.58	24	209.69	24
Finland	11348	29	6207	26	67.6	28	590	26	224.32	25	194.97	25
Norway	11443	28	6164	27	54.2	38	414	29	160.37	28	149.77	26
Czech Republic	10602	31	4211	31	56.8	32	372	32	161.07	27	126.33	27
Portugal	12696	25	6150	28	63	29	416	28	143.68	30	117.45	28
Ireland	7302	40	3889	35	44.6	41	331	33	125.92	32	113.91	29
Argentina	8454	36	3485	36	38.6	43	321	34	142.47	31	98.58	30
New Zealand	8373	37	4197	32	55.6	35	275	37	102.96	35	95.69	31
Chile	6380	43	2724	39	68.6	27	856	23	192.35	26	95.12	32
Greece	10096	33	4494	30	58.2	31	319	35	118.03	34	91.76	33
Mexico	11681	27	4035	34	74.2	25	391	30	146.56	29	84.6	34
South Africa	10301	32	4104	33	54.8	37	417	27	122.16	33	81.55	35
Turkey	26206	17	5518	29	134	19	236	38	97.53	36	79.24	36
Hungary	6312	44	2613	40	50.8	39	276	36	96.2	37	78.48	37
Saudi Arabia	9061	35	3191	37	55.8	34	384	31	80.37	39	76.63	38
Iran	25824	19	6244	25	153.2	16	147	40	81.03	38	67.87	39
Ukraine	4869	46	1255	47	31.8	44	173	39	50.5	40	37.06	40
Slovenia	3888	47	1478	44	22.4	48	106	41	37.89	41	36.36	41
Thailand	6583	41	2389	41	27.4	46	92	43	24.98	43	22.68	42
Estonia	1744	53	799	51	9.2	53	92	44	27.42	42	22.31	43
Romania	8246	38	2109	43	40.4	42	79	46	24.66	44	21.41	44
Croatia	3512	48	1024	49	17.8	49	102	42	23.3	47	17.56	45
Serbia	5105	45	1441	45	26	47	55	52	24.62	46	17	46
Lithuania	1996	52	519	55	8.8	54	58	50	24.62	45	16.85	47
Malaysia	9282	34	2971	38	59.4	30	53	53	15.81	50	15.81	48
Egypt	7932	39	2160	42	55.4	36	87	45	14.69	52	14.32	49
Slovakia	3219	50	979	50	15.8	50	63	49	21	48	14.15	50

Pakistan	6446	42	1411	46	28.4	45	42	55	17.04	49	12.83	52
United Arab Emirates	1456	55	583	54	11.2	52	49	54	11.53	55	9.37	54
Colombia	3422	49	1210	48	15.2	51	56	51	11.72	54	9.36	55
Bulgaria	2302	51	707	52	46.8	40	63	48	14.83	51	8.59	57
Indonesia	1563	54	644	53	6.4	55	30	59	6.72	59	5.92	59

Note. * $Q1_{JIF}$ values are from 2013 and not from 2014. The results of Lopez-Abente and Munoz-Tinoco (2005) and Tsigilis, Grouios, Tsorbatzoudis, and Koidou (2010) show that journals' JIFs remained relatively stable over the years.

Based on the indicator values in Table 1, we calculated Spearman's rank correlation coefficients. This coefficient is a descriptive statistical measure which represents the degree of relationship between two indicators. A positive correlation points to a monotonic increase, that means the increase in the value of one indicator is always accompanied by an increase in the value of the other indicator (Sheskin, 2007). The results of the correlations are shown in the correlation matrix of Table 2. All correlation coefficients are at least at r_s=.88. When we interpret the coefficients against the backdrop of Cohen (1988), we can conclude that the coefficients are much larger than typical and on a very high level (see also Kraemer et al., 2003).

The most important NI indicator is the WFC. This indicator is mainly used by the NPG to rank institutions or countries. WFC shows the highest correlations with the other NI indicators (r_s=.98 for the AC and r_s=.99 for the FC). A similar high correlation is pointed out for the $Q1_{JIF}$ indicator (r_s=.96). The correlations with the total number of papers and the (random) AvgAC are somewhat lower with r_s=.91 and r_s=.88, but still much higher than one would expect.

Table 2: Correlations between total number of papers (Np), $Q1_{JIF}$, AvgAC, AC, FC, and WFC

	Np	$Q1_{JIF}$	AvgAC	AC	FC	WFC
Np	1					
$Q1_{JIF}$	0.97	1				
Avg_AC	0.97	0.95	1			
AC	0.90	0.95	0.90	1		
FC	0.91	0.96	0.89	0.99	1	
WFC	0.91	0.96	0.88	0.98	0.99	1

4.2 Size-independent and size-dependent NI indicators

In the previous section, we have shown that the different NI variants correlate much higher than typical with the mere number of papers on the country level. This result points out that the NI variants are size-dependent indicators. Thus, we recommend to complement the size-dependent NI indicators with size-independent NI indicators. Already in Haunschild & Bornmann (2015), we recommended this additional perspective and gave the following supporting example: The NI ranks the Chinese Academy of Sciences (CAS) before Harvard University with 2,661 papers from CAS and 2,555 papers from Harvard in 2013. Considering the full publication output of both institutes in 2013 (31,428 for CAS and 17,836 for Harvard), we see that in relative terms Harvard (14% of Harvard papers in the NI) ranks higher than the CAS (8% of CAS papers in the NI).

The CWTS justifies its presentation of size-independent indicators (besides size-dependent indicators) in the Leiden ranking as follows: "In the case of size-dependent indicators, universities with a larger publication output tend to perform better than universities with a smaller publication output. Size-independent indicators have been corrected for the size of the publication output of a university. So when size-independent indicators are used, both larger and smaller universities may perform well" (http://www.leidenranking.com/methodology/indicators#sthash.52ZO1Kmm.dpuf). On the country-level, many small countries have no way of publishing the same (or larger) number of papers in NI journals as large countries. For example, the USA published n=26,631 papers in journals considered in the NI (see Table 1). Switzerland has a total publication output of n=25,979. Since the papers published in NI journals cannot exceed the total publication output of a country, Switzerland could never reach the top position in the AC, FC, or WFC ranking.

Table 3: Number of papers (Np), AC, RelAC, $RelQ1_{JIF}$, and RelAvgAC by country. The rank positions for every country are added based on orderings by the different indicators. The table is sorted in decreasing order by RelAC.

Country	Np	AC	Rank AC	RelAC	Rank RelAC	$RelQ1_{JIF}$	Rank RelQ1	RelAvgAC	Rank RelAvgAC
Chile	6380	856	23	13.42	1	42.70	25	1.08	2
Switzerland	25979	2709	10	10.43	2	60.55	2	0.53	32
Germany	98154	8576	3	8.74	3	54.00	12	0.59	22
Israel	12491	1012	19	8.10	4	52.87	17	0.58	23
Singapore	11287	873	21	7.73	5	59.64	3	0.50	39
France	68382	5241	5	7.66	6	54.99	9	0.47	46
United Kingdom (UK)	102766	7590	4	7.39	7	55.41	6	0.54	29
United States of America (USA)	365437	26631	1	7.29	8	55.15	7	0.60	18
Denmark	14944	1036	18	6.93	9	59.29	4	0.69	7
Japan	76638	4973	6	6.49	10	42.12	26	0.45	50
Austria	13384	856	22	6.40	11	54.40	11	0.60	19
Netherlands	35394	2188	12	6.18	12	61.21	1	0.58	26
Sweden	23716	1407	15	5.93	13	56.77	5	0.52	34
Belgium	19704	1082	17	5.49	14	55.01	8	0.49	41
Spain	53876	2896	9	5.38	15	52.41	18	0.56	28
Canada	60340	3224	7	5.34	16	53.70	14	0.61	15
Estonia	1744	92	44	5.28	17	45.81	23	0.53	33
Finland	11348	590	26	5.20	18	54.70	10	0.60	20
Italy	58707	3051	8	5.20	19	49.92	20	0.63	13
Australia	50495	2497	11	4.95	20	52.96	16	0.65	10
Ireland	7302	331	33	4.53	21	53.26	15	0.61	17
Hungary	6312	276	36	4.37	22	41.40	28	0.80	3

Saudi Arabia	9061	384	31	4.24	23	35.22	38	0.62	14
South Africa	10301	417	27	4.05	24	39.84	32	0.53	31
China	215731	8632	2	4.00	25	41.83	27	0.47	45
Russia	28981	1145	16	3.95	26	21.88	54	0.74	5
South Korea	50951	1966	13	3.86	27	38.72	34	0.56	27
Argentina	8454	321	34	3.80	28	41.22	29	0.46	49
Norway	11443	414	29	3.62	29	53.87	13	0.47	44
Ukraine	4869	173	39	3.55	30	25.78	50	0.65	9
Czech Republic	10602	372	32	3.51	31	39.72	33	0.54	30
United Arab Emirates	1456	49	54	3.37	32	40.04	31	0.77	4
Mexico	11681	391	30	3.35	33	34.54	39	0.64	12
New Zealand	8373	275	37	3.28	34	50.13	19	0.66	8
Portugal	12696	416	28	3.28	35	48.44	21	0.50	38
Taiwan	27586	887	20	3.22	36	47.28	22	0.46	47
Greece	10096	319	35	3.16	37	44.51	24	0.58	24
Poland	22872	685	25	2.99	38	30.51	44	0.61	16
Lithuania	1996	58	50	2.91	39	26.00	49	0.44	52
Croatia	3512	102	42	2.90	40	29.16	46	0.51	37
India	51751	1483	14	2.87	41	31.30	41	0.49	43
Bulgaria	2302	63	49	2.74	42	30.71	43	2.03	1
Slovenia	3888	106	41	2.73	43	38.01	35	0.58	25
Slovakia	3219	63	48	1.96	44	30.41	45	0.49	40
Indonesia	1563	30	59	1.92	45	41.20	30	0.41	55
Brazil	38309	715	24	1.87	46	30.78	42	0.46	48
Colombia	3422	56	51	1.64	47	35.36	37	0.44	51
Thailand	6583	92	43	1.40	48	36.29	36	0.42	54
Egypt	7932	87	45	1.10	49	27.23	48	0.70	6
Serbia	5105	55	52	1.08	50	28.23	47	0.51	36
Romania	8246	79	46	0.96	51	25.58	51	0.49	42

Turkey	26206	236	38	0.90	52	21.06	55	0.51	35
Pakistan	6446	42	55	0.65	53	21.89	53	0.44	53
Malaysia	9282	53	53	0.57	54	32.01	40	0.64	11
Iran	25824	147	40	0.57	55	24.18	52	0.59	21

Table 3 shows the size-independent AC values (RelAC) for the different countries whereas the AC values have been divided by the total number of papers of the specific country and multiplied with 100 in order to obtain percentages:

RelAC = AC/Np * 100

For comparison, analogous size-independent values for the $Q1_{JIF}$ ($RelQ1_{JIF}$) and AvgAC (RelAvgAC) indicators have been added. We could not calculate relative FC and WFC indicators, because we are not able to reproduce this variant of fractional counting based on our in-house database. As the results in Table 3 show (expectedly) the size-independent AC variant leads to top positions for smaller countries, like Chile or Switzerland. UK and USA are on positions seven and eight. Similar results are visible for $RelQ1_{JIF}$ and RelAvgAC: $RelQ1_{JIF}$ leads to top positions for the Netherlands, Switzerland, and Singapore and RelAvgAC puts Bulgaria, Chile, and Hungary on top positions. Whereas the top positions based on $RelQ1_{JIF}$ are reasonable (Netherlands, Switzerland, and Singapore are known as high performing small countries), the results based on RelAvgAc (Bulgaria and Hungary) seem questionable.

Similar to section 4.1, we calculated Spearman's rank correlation coefficients to compare AC and the different size-independent indicators. The correlation between AC and RelAc is r_s=0.76. Since this correlation coefficient is definitely lower than the coefficients in Table 2, the relative variant seems to be an informative additional indicator to the AC (Cohen, 1988; Kraemer et al., 2003). $RelQ1_{JIF}$ correlates on a similar level with RelAc (r_s=0.82) and definitely less with AC and (r_s=0.64). However, we obtain significantly lower coefficients for the correlations with RelAvgAC (between r_s=0.06 and r_s=0.23). As the top country positions for the RelAvgAC already revealed, the generally low coefficients point to less conforming results with the other indicators (with the size-dependent AC as well as the size-independent RelAC and $RelQ1_{JIF}$) and questions the validity of this indicator.

Table 4: Correlations between AC, RelAC, and $RelQ1_{JIF}$

	AC	RelAC	$RelQ1_{JIF}$	RelAvgAC
AC	1			
RelAC	0.76	1		
$RelQ1_{JIF}$	0.64	0.82	1	
RelAvgAC	0.08	0.23	0.06	1

The relative indicators RelAC and $RelQ1_{JIF}$ certainly offer an important additional perspective on country performance to AC and $Q1_{JIF}$. There is also another kind of relative indicators that can be informative: Indicators obtained by normalizing for the worldwide production, specifically AC divided by the sum of AC (RelSumAC) and $Q1_{JIF}$ divided by the sum of $Q1_{JIF}$ ($RelSumQ1_{JIF}$). Here, AC and $Q1_{JIF}$, respectively, are divided by the sum over all countries (AC or $Q1_{JIF}$, respectively). Finally, these indicators are multiplied by 100 in order to obtain percentages:

$$RelSumAC = AC/\sum AC * 100$$

$$RelSumQ1_{JIF} = Q1_{JIF}/\sum Q1_{JIF} * 100$$

These indicators can answer the question: Among all the papers published in reputable journals worldwide (measured by AC or $Q1_{JIF}$), how many of them come from a specific country? Note that RelSumAC and $RelSumQ1_{JIF}$ offer a relative perspective (relative to the world), but the indicators are size-dependent. Since ranks based on these indicators would lead to the same country positions as in Table 1 (AC and $Q1_{JIF}$), we present in Table 5 the indicator values only. As the results in the table show the USA is the largest producer of papers published in reputable journals with 23.87% ($RelSumQ1_{JIF}$) and 27.04% (RelSumAC).

Table 5: RelSumAC and RelSumQ1$_{JIF}$ by country (in percent). The table is sorted in decreasing order by RelSumAC.

Country	**Q1**	**AC**	**RelSumQ1$_{JIF}$**	**RelSumAC**
United States of America (USA)	201555	26631	23.87	27.04
China	90233	8632	10.69	8.76
Germany	53002	8576	6.28	8.71
United Kingdom (UK)	56942	7590	6.74	7.71
France	37602	5241	4.45	5.32
Japan	32281	4973	3.82	5.05
Canada	32401	3224	3.84	3.27
Italy	29305	3051	3.47	3.10
Spain	28237	2896	3.34	2.94
Switzerland	15729	2709	1.86	2.75
Australia	26740	2497	3.17	2.54
Netherlands	21666	2188	2.57	2.22
South Korea	19730	1966	2.34	2.00
India	16200	1483	1.92	1.51
Sweden	13463	1407	1.59	1.43
Russia	6342	1145	0.75	1.16
Belgium	10839	1082	1.28	1.10
Denmark	8861	1036	1.05	1.05
Israel	6604	1012	0.78	1.03
Taiwan	13043	887	1.54	0.90
Singapore	6732	873	0.80	0.89
Austria	7281	856	0.86	0.87
Chile	2724	856	0.32	0.87
Brazil	11791	715	1.40	0.73
Poland	6979	685	0.83	0.70
Finland	6207	590	0.74	0.60
South Africa	4104	417	0.49	0.42
Portugal	6150	416	0.73	0.42
Norway	6164	414	0.73	0.42
Mexico	4035	391	0.48	0.40
Saudi Arabia	3191	384	0.38	0.39
Czech Republic	4211	372	0.50	0.38
Ireland	3889	331	0.46	0.34
Argentina	3485	321	0.41	0.33
Greece	4494	319	0.53	0.32
Hungary	2613	276	0.31	0.28
New Zealand	4197	275	0.50	0.28
Turkey	5518	236	0.65	0.24
Ukraine	1255	173	0.15	0.18
Iran	6244	147	0.74	0.15
Slovenia	1478	106	0.18	0.11
Croatia	1024	102	0.12	0.10

Estonia	799	92	0.09	0.09
Thailand	2389	92	0.28	0.09
Egypt	2160	87	0.26	0.09
Romania	2109	79	0.25	0.08
Iceland	515	65	0.06	0.07
Bulgaria	707	63	0.08	0.06
Slovakia	979	63	0.12	0.06
Lithuania	519	58	0.06	0.06
Colombia	1210	56	0.14	0.06
Serbia	1441	55	0.17	0.06
Malaysia	2971	53	0.35	0.05
United Arab Emirates	583	49	0.07	0.05
Pakistan	1411	42	0.17	0.04
Luxembourg	471	35	0.06	0.04
Qatar	442	35	0.05	0.04
Panama	199	32	0.02	0.03
Armenia	329	30	0.04	0.03
Indonesia	644	30	0.08	0.03
Total	844419	98497	100.00	100.00

5 Discussion

According to Osterloh and Frey (2014), there are four reasons why "rankings are deemed to be necessary. First, it is argued that because of the high specialization of research and the lack of expertise in areas that are different from the own research field, it is efficient to rely on research rankings … Second, research rankings fuel competition among scholars, lead to more and better research, and promote what is called an 'entrepreneurial university' … Third, research rankings give the public a transparent picture of scholarly activity. They make scientific merits visible to people who have no special knowledge of the field like politicians, public officials, deans, university administrators, and journalists … Fourth, academic rankings make universities and departments more accountable for their use of public money" (Osterloh & Frey, 2014, p. 2). However, university rankings have always been heavily criticized (Huang, 2012; Mutz & Daniel, in press; Schmoch, 2015). For example, it is

criticized that different indicator values are weighted in a specific way in order to compute a sum score and that the weighting is not appropriately justified.

The NI does not use a mix of different indicators to compute a sum score. It is based on the number of papers (fractionally counted) published in reputable journals. Thus, the NI provides a ranking of institutions and countries based on a small sub-set of bibliometric data only. Starting from two short comments on the NI (Haunschild & Bornmann, 2015a, 2015b), we undertook an empirical analysis of the index using comprehensive country data. The analysis is based on data from the MPDL in-house database (which is based on WoS data). In a first step of analysis, we correlated the NI with other metrics which are simpler to generate than the NI. The resulting very large correlation coefficients point out that the NI produces very similar results as simpler solutions. Thus, the use of the NI is questioned by the empirical results. For example, the NI could be replaced by the $Q1_{JIF}$ indicator (which is used in the SIR and excellencemapping.net in a similar form, see Bornmann, Stefaner, de Moya Anegón, & Mutz, 2014) and also measures the amount of output published in high-quality journals. In a second step of analysis, two relative variants of the NI have been generated: one variant is size-independent (RelAC and $RelQ1_{JIF}$) and one variant is size-dependent (RelSumAC and $RelSumQ1_{JIF}$). The size-dependent variant produces the same country ranks as the original AC and the $Q1_{JIF}$. Therefore, we recommend that the relative and size-independent variants (RelAC and $RelQ1_{JIF}$) should be additionally presented by the NPG along with the fractionally counted versions RelFC and RelWFC. The size-dependent NI indicators favor large countries (or institutions) and the top-performing small countries (or institutions) do not come into the picture.

According to Campbell and Grayson (2015), the "NPG actively seeks constructive feedback from the researcher community we serve, and our aim is to iterate and improve the Nature Index in response to such feedback" (p. 1831). We hope that our empirical results and recommendations are helpful for the improvement of the NI.

Acknowledgements

The bibliometric data used in this paper are from an in-house database developed and maintained by the Max Planck Digital Library (MPDL, Munich) and derived from the Science Citation Index Expanded (SCI-E), Social Sciences Citation Index (SSCI), Arts and Humanities Citation Index (AHCI) prepared by Thomson Reuters (Philadelphia, Pennsylvania, USA). We thank the anonymous reviewers for their constructive comments on the initial version of our manuscript.

References

Aguillo, I. F., Bar-Ilan, J., Levene, M., & Ortega, J. L. (2010). Comparing university rankings. *Scientometrics, 85*(1), 243-256. doi: DOI 10.1007/s11192-010-0190-z.

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014). What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. *Journal of Informetrics, 8*(3), 581-593.

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71*(3), 349-365.

Campbell, N., & Grayson, M. (2014). Introducing the index. *Nature, 515*(7526), S52-S53.

Campbell, N., & Grayson, M. (2015). A response to 'Discussion about the new Nature Index'. *Scientometrics, 102*(2), 1831-1833. doi: 10.1007/s11192-014-1516-z.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates, Publishers.

Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education, 49*(4), 495-533.

Gonzalez-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegon, F. (2010). A new approach to the metric of journals' scientific prestige: the SJR indicator. *Journal of Informetrics, 4*(3), 379-391. doi: 10.1016/j.joi.2010.03.002.

Haunschild, R., & Bornmann, L. (2015a). Criteria for Nature Index questioned. *Nature, 517*(7532), 21.

Haunschild, R., & Bornmann, L. (2015b). Discussion about the new Nature Index. *Scientometrics, 102*(2), 1829-1830. doi: 10.1007/s11192-014-1505-2.

Huang, M. H. (2012). Opening the black box of QS World University Rankings. *Research Evaluation, 21*(1), 71-78. doi: DOI 10.1093/reseval/rvr003.

Kraemer, H. C., Morgan, G. A., Leech, N. L., Gliner, J. A., Vaske, J. J., & Harmon, R. J. (2003). Measures of clinical significance. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(12), 1524-1529. doi: 10.1097/01.chi.0000091507.46853.d1.

Levy, P. S., & Lemeshow, S. (2008). *Sampling of Populations: Methods and Applications*. Hoboken, NJ, USA: Wiley.

Lopez-Abente, G., & Munoz-Tinoco, C. (2005). Time trends in the impact factor of Public Health journals. *BMC Public Health, 5*. doi: 10.1186/1471-2458-5-24.

Mutz, R., & Daniel, H.-D. (in press). What is behind the curtain of the Leiden Ranking? *Journal of the Association for Information Science and Technology*.

Osterloh, M., & Frey, B. S. (2014). Ranking Games. *Evaluation Review*. doi: 10.1177/0193841x14524957.

Pudovkin, A. I., & Garfield, E. (2004). Rank-normalized impact factor: a way to compare journal performance across subject categories. In J. B. Bryans (Ed.), *ASIST 2004: Proceedings of the 67th Asis&T Annual Meeting, Vol 41, 2004: Managing and Enhancing Information: Cultures and Conflicts* (Vol. 41, pp. 507-515). Medford: Information Today Inc.

Rauhvargers, A. (2011). *Global university rankings and their impact*. Brussels, Belgium: European University Association (EUA).

Safón, V. (2013). What do global university rankings really measure? The search for the X factor and the X entity. *Scientometrics, 97*(2), 223-244. doi: 10.1007/s11192-013-0986-8.

Schmoch, U. (2015). The Informative Value of International University Rankings: Some Methodological Remarks. In I. M. Welpe, J. Wollersheim, S. Ringelhan & M. Osterloh (Eds.), *Incentives and Performance* (pp. 141-154): Springer International Publishing.

Sheskin, D. (2007). *Handbook of parametric and nonparametric statistical procedures* (4th ed.). Boca Raton, FL, USA: Chapman & Hall/CRC.

Tsigilis, N., Grouios, G., Tsorbatzoudis, H., & Koidou, I. (2010). Impact factors of the sport sciences journals: Current trends, relative positions, and temporal stability. *European Journal of Sport Science*, *10*(2), 81-90. doi: Pii 919150584
10.1080/17461390903125152.