Jointly published by Akadémiai Kiadó, Budapest
and Springer, Dordrecht

Scientometrics, Vol. 79, No. 3 (2009) 507–516
DOI: 10.1007/s11192-007-2025-0

Influence of individual researchers' visibility on institutional impact: an example of Prathap's approach to successive h-indices

RICARDO ARENCIBIA-JORGE[a] RONALD ROUSSEAU[b,c]

[a] *Network of Scientometric Studies for Higher Education, National Scientific Research Center, Avenida 25 y Calle 158, Cubanacán, Playa, PO Box 6414, Havana City, Cuba*
[b] *KHBO (Association K.U.Leuven), Industrial Sciences and Technology, Zeedijk 101, B-8400 Oostende, Belgium*
[c] *K.U.Leuven, Steunpunt O&O Indicatoren, Dekenstraat 2, B-3000, Belgium*

This study applies Prathap's approach to successive *h*-indices in order to measure the influence of researcher staff on institutional impact. The twelve most productive Cuban institutions related to the study of the human brain are studied. The Hirsch index was used to measure the impact of the institutional scientific output, using the *g*-index and *R*-index as complementary indicators. Prathap's approach to successive *h*-indices, based on the *author–institution* hierarchy, is used to determine the institutional impact through the performance of the researcher staff. The combination of different Hirsch-type indices for institutional evaluation is illustrated.

Introduction

During the latest two years, the Hirsch index (*h*-index) has been one of the most discussed topics related to the use of citation analysis for institutional evaluation [GLÄNZEL, 2006; HIRSCH, 2005; JIN, 2006]. The combination of productivity and impact is a major characteristic of the *h*-index, offering the possibility of measuring the lifetime achievement of researchers and scholars, based on their scientific output.

Several authors have published different adaptations to the original proposal [EGGHE, 2006A; JIN & AL., 2007; KOSMULSKI, 2006] and applied it in different contexts, including ranking institutions. The possibility of using the *h*-index as a basis for a series of successive *h*-indices is one of these new developments, proposed simultaneously by András Schubert in Hungary, and Gangan Prathap in India [PRATHAP, 2006; SCHUBERT, 2007].

Schubert proposed a series of successive *h*-indices for the *journal–publishing group–country* hierarchy, where the *h*-index of the journals determines the *h*-index of each publishing group, and this in turn determines the *h*-index of each country. However, he first expressed the idea of using successive *h*-indices in the context of

Received December 10, 2007; Published online January 31, 2009

Address for correspondence:
RICARDO ARENCIBIA-JORGE
E-mail: ricardo.arencibia@cnic.edu.cu

0138–9130/US $ 20.00

evaluations on different aggregation levels, giving the *researcher-institution-country* hierarchy as an example [SCHUBERT, 2007].

Somewhat earlier, in a brief letter published in *Current Science*, PRATHAP [2006] proposed two levels for using the *h*-index in institutional evaluations, taking into account a first order *h*-index (h_1) and a second order *h*-index (h_2), where the institute's first order *h*-index is equal to h_1 if the institution (this is the group of all its researchers) has published h_1 papers, each of which has at least h_1 citations; and its second order *h*-index is h_2 if the institution has h_2 researchers, each having an individual *h*-index which is at least equal to h_2 [PRATHAP, 2006]. Note that the calculation of Prathap's h_2 is a special case of Schubert's idea of successive *h*-indices. His h_1-index, however, is not. In order to explain this we make the steps in each scientist's approach clear.

Schubert's approach to successive *h*-indices starts with units (e.g. articles) which are subdivided in groups according to some principle (e.g. being published in the same journal, or being published by the same researcher). Citations are collected and a (standard) *h*-index is obtained for each group (e.g. for each journal or for each author). This is Schubert's first level *h*-index. Next the groups are subdivided according to another principle (e.g. journals are brought together in publisher groups; researchers are brought together per institute). The first level *h*-indices are ranked for each group leading to a second-level *h*-index (e.g. an *h*-index for journals or institutes). These second-level *h*-indices are again brought together according to some principle (e.g. publishers and institutions are brought together per country). Each member of a new group (here: a country) has a second-level *h*-index, and the classical *h*-index of these second-level *h*-indices yields the third-level *h*-index. In principle this can go on, leading to many more levels of successive h-indices. Note that only the first ranked list consists of citations, all other lists are lists of *h*-indices. This is the main characteristic of the idea of successive *h*-indices.

Prathap's approach is different. His h_1 is the standard *h*-index of all articles published by researchers of an institute. His h_2 is the second-level successive *h*-index, where first articles are grouped per researcher, leading to a standard *h*-index for each researcher. Then researchers are grouped per institute, leading to a second-level *h*-index, which is Prathap's h_2. Prathap's main idea is to compare h_1 and h_2, using them as complementary indices.

Schubert's proposal was tested recently by Arencibia and colleagues, who described the use of successive *h*-indices at the micro level, on a *research-department–institute* hierarchy, showing that the new indicator offered an integral vision of the institute and its staff [ARENCIBIA JORGE & AL., 2008]. A slight adaptation of Schubert's proposal, using rational *h*-indices, was used in a case study related to economics departments in Ireland [RUANE & TOL, 2008]. Prathap's h_1 and h_2 and Schubert's successive *h*-indices have been modelled by EGGHE [2007].

The current study follows Prathap's approach in order to measure the influence of the researcher's staff on the institutional impact of several Cuban research entities in the field of human brain research.

Methods

The first step of the study was to choose a scientific discipline, with the aim to identify the main actors of the research network. In this case, the Cuban scientific production on brain research published in mainstream journals was chosen. All Cuban papers – defined here as an article with at least one Cuban address – belonging to the five subject categories of the Web of Science (WoS) related to brain research were retrieved. These five topics are: Clinical Neurology, Neuroimaging, Neurosciences, Psychiatry, and Psychology. A total of five years were analyzed, from January 2001 to December 2005. All data fields were exported to an *ad hoc* database developed by the SITKIS program [SCHILDT & MATTSSON, 2006]. A careful process of normalization was carried out in order to identify the correct name of all the authors and institutions comprised in the study.

The second step was the selection of the most productive Cuban institutions on brain research, and the calculation of different indicators based on a citation analysis of papers published during the period by these institutions. For each entity the following indicators were calculated:

- Published articles during the period 2001–2005 (denoted as A) and its proportion (%) with respect to the total number of Cuban articles on brain research.
- Cited articles (at least once in the WoS) during the period 2001–2005 (denoted as CA) and its proportion (denoted as% CA) with respect to the total number of published articles.
- Total number of citations received (denoted as TC), and the average number of citations received by each article (denoted as CxA = TC/A).
- Average Garfield Impact Factor (denoted as AvIF) of journals in which the articles were published, using the IF of the year in which the article was published.
- Total number of citations received by the articles of one institute comprised in the Hirsch core (denoted as TC Hcore).
- Determination of *h*-indices [HIRSCH, 2005].
- Determination of *g*-indices [EGGHE, 2006B].
- Determination of *R*-indices [JIN & AL., 2007].

We further obtained the total number of occurrences and the total number of different authors involved in the scientific production of each institution, as well as the collaboration ratio (denoted as *C*-ratio and defined further on).

The third step was the *h*-index calculation for all the authors comprised in the study and from this the h_2-index calculation for each of the most productive institutions, based on the *author–institution* hierarchy suggested by Prathap and Schubert [PRATHAP, 2006; SCHUBERT, 2007].

Finally, rankings are compared, and the usefulness of these different *h*-indices for institutional evaluation is discussed.

Using the Hirsch index to rank institutions

The Hirsch index takes into account the individual scientific output of scholars and researchers. However, the indicator can be applied to other aggregation levels. In this section, considering an institutional list of publications, ranked according to the number of citations received, the *h*-index of the institution is defined as the highest rank such that the first *h* publications received each at least *h* citations. All articles ranked between rank 1 and rank *h* form the Hirsch core. This step leads to Prathap's h_1 for each institute involved in the study. Recall that h-indices for research groups have been discussed before [VAN RAAN, 2006; EGGHE & RAO, 2008].

In case of ties the *g*-index proposed by Egghe, and the *R*-index proposed by Jin and colleagues were used [EGGHE, 2006A; JIN & AL., 2007]. Both indicators take the amount of citations received by the most cited articles into account.

The *g*-index is defined as the highest rank such that the cumulative sum of the number of citations received is larger than or equal to the square of this rank [EGGHE, 2006A,B]. The *R*-index involves only articles included in the Hirsch core, and is defined as the square root of the sum of citations received by these articles [JIN & AL., 2007]. Jin and colleagues found a strong correlation between the *g*-index and the *R*-index, but they suggest the *R*-index in conjunction with the *h*-index for practical evaluation processes [JIN & AL., 2007]. In this study, ties are solved by the *g*-index, and in case ties still exist the *R*-index is used.

Twelve Cuban institutions related to brain research published more than ten articles during the period 2001–2005 (see Table 1). These institutions were ranked according to the *h*-index. We note that only three have an *h*-index higher than the number of years comprised in the study.

The Cuban Neurosciences Center (CNC) has ten articles with at least ten citations. These articles are, moreover, highly cited, an aspect clearly revealed by the *g*- and the *R*-index. The International Center of Neurological Restoration (CIREN) and the University of Havana (UH) have equal *h*-indices. The first one has better results according to the *g*- and *R*-index, but the second one performs better for the indices: proportion of cited articles, average number of citations by article, and average impact factor of journals in which the articles were published.

Table 1. Ranking of the most productive Cuban institutions on brain research according to the Hirsch index (h_1)

Institution	A	%	CA	% CA	TC	CxA	AvIF	TC Hcore	*h*-index	*g*-index	*R*-index
CNC	53	13.0	38	71.7	471	8.887	3.241	375	10	21	19.36
CIREN	106	26.0	45	42.5	264	2.491	1.5	185	7	14	13.60
UH	23	5.6	18	78.3	156	6.957	2.402	123	7	12	11.09
HCQ-HAM	22	5.4	9	40.9	38	1.727	1.526	29	4	5	5.39
CIRAH	12	2.9	7	58.3	27	2.250	1.818	23	4	5	4.80
ISCMH	28	6.9	15	53.6	37	1.321	1.258	13	3	4	3.61
HM-CJF	14	3.4	7	50.0	19	1.357	1.192	13	3	3	3.61
INNN	65	15.9	21	32.3	35	0.538	0.892	8	2	3	2.83
HCQ-AMC	14	3.4	7	50.0	14	1.000	0.300	5	2	2	2.24
HPH	12	2.9	1	8.3	4	0.333	2.404	4	1	2	2.00
HCQ-MAD	24	5.9	16	66.7	18	0.750	0.276	3	1	2	1.73
CIMEQ	14	3.4	6	42.9	7	0.500	0.282	2	1	1	1.41

A: Total number of articles; %: Percentage of the total of articles; CA: Total number of cited articles; % CA: Percentage of cited articles; TC: Total number of citations received; CxA: Average of citations by article; AvIF: Average impact factor of the journals in which the articles were published; TC Hcore: Total number of citations received by the Hirsch core.

Some of the ties for the *h*-index have also ties for the *g*-index rank. In these cases, the *R*-index was used to determine the rank. Only two institutions have the same value for the h- and the R-index: the Higher Institute of Medical Sciences from Havana (ISCMH), and the Military Hospital "Carlos J. Finlay" (HM-CJF). Both institutions have similar values for the average of citations by article and the average impact factor of journals used to publish the articles. However, the first institution is more productive and more cited than the second one, and this is reflected in the *g*-index. We confirm Jin's findings that the *g*-index and the *R*-index are highly correlated. In our case the Pearson correlation coefficient between these two indices is 0.998.

Applying successive *h*-indices in order to determine Prathap's h_2

Successive *h*-indices offer a different point of view. This approach takes the researcher as the unit for the determination of the institutional impact [ARENCIBIA-JORGE & AL., 2008]. For this reason, the *h*-index calculation for the researchers from each institution is a necessary first step.

We identified all authors of Cuban articles on brain research, and determined the *h*-index for each of them. Based on this set of authors, the *h*-index of the most productive institutions (h_2) was obtained.

A total of 888 authors were identified in the 408 Cuban articles on brain research, 694 (78.2%) of which were Cubans (defined here as scientists with a Cuban address, the large majority of which are actually Cubans), and 194 (21.8%) belonging to 123 institutions from 24 other countries.

A core of 217 authors (24.4%) published more than two articles. Taking into account the years comprised in the study (5 years), they were considered the representative authors in brain research (most are Cubans but some are non-Cubans working in collaboration with Cuban authors). The distribution of the author's *h*-indices is shown in Table 2. For the most productive core (n = 217), 84.3% of the authors have a non-zero *h*-index ($h \geq 1$), but only 3.2% have $h \geq 5$. Authors with *h*-index equal to 1 (38.2%) and 2 (29%) clearly dominate. We note that 82.4% of these authors belong to the twelve most productive institutions. However, when all authors are analyzed, the proportion of authors with *h*-index equal to 0 increases from 15.7 to 43.4%, and the correlation between *h*-index and the total number of articles published by an author is higher for all authors ($r = 0.64$) than for the most productive core ($r = 0.54$). This result demonstrates clearly the necessity of producing a sizable number of articles, in order to achieve a high *h*-index.

Table 2. Distribution of authors according to their *h*-index

	The most productive core (n = 217)			All authors (n = 888)		
h-index	Authors	%	Cumulative %	Authors	%	Cumulative %
7	1	0.46	0.46	1	0.11	0.11
6	1	0.46	0.92	1	0.11	0.22
5	5	2.30	3.23	5	0.56	0.78
4	4	1.84	5.07	4	0.45	1.23
3	26	11.98	17.05	26	2.93	4.16
2	63	29.03	46.08	85	9.57	13.73
1	83	38.25	84.33	381	42.91	56.64
0	34	15.67	100.00	385	43.36	100.00
Total	217	100.00		888	100.00	

Inspired by Rao's approach [RAO, 2007] we tried to fit a simple discrete distribution for this group of *h*-indices and obtained an acceptable fit for the most productive group using a Poisson distribution with parameter 1.5 (chi-square fit; d.f. = 4; $p = 0.37$). Also a binomial distribution provided an acceptable fit. Yet, we were not able to find an acceptable fit (with a simple discrete distribution) for the more relevant group consisting of all authors. Although the mean of this group is almost equal to the variance (mean = 0.769, variance = 0.793) no Poisson distribution fits the data. Clearly binomial, geometric or discrete Lotka (power function) distributions do not have the properties to provide an acceptable fit. We leave it as an open problem to find a model and a fitting procedure for this type of data (not only for the particular case of Cuban authors).

Having obtained each scientist's *h*-index, we can proceed and calculate the institutes' h_2-index. The new ranking (see Table 3) of Cuban institutes based on their h_2-index, and in case of ties, based on the maximum value of their scientists' *h*-index and the total number of publications, differs little from the one presented in Table 1 (Spearman rank correlation equal to 0.91). In the new ranking, the position of the Cuban institutions involved in brain research depends on the number of authors from these institutions with a high h-index. If an institute has only one big star, then its h_2-index is just equal to one, no matter how important this institute is. For this reason an institute's h_2-index can be considered as an indicator of institutional impact.

One more time, CNC and CIREN have the best performance, with an h_2-index larger than 2. Ties in h_2 were solved taking into account the highest h_1 value of each institution (h_{1max}); and if these are also equal the rank is determined by the total amount of articles published during the period 2000–2005.

Table 3. Ranking of the most productive Cuban institutions on brain research according to successive h-indices

Institution	Authors	Occurrences	C-ratio	h_2-index	h_{1max}
CNC	145	302	5.7	4	7
CIREN	208	594	5.6	3	6
UH	94	113	4.9	2	5
CIRAH	40	49	4.1	2	4
HCQ-HAM	85	136	6.2	2	3
INNN	148	286	4.4	2	2
HCQ-AMC	36	55	3.9	2	2
ISCMH	83	137	4.9	1	3
HM-CJF	71	104	7.4	1	2
CIMEQ	69	132	9.4	1	1
HCQ-MAD	32	101	4.2	1	1
HPH	21	35	2.9	1	1

C-ratio: Collaboration ratio; h_2-index: Prathap's h_2-index;
h_{1max}: The highest h_1-index of each institution.

The second column of Table 3 refers to the total number of different authors of the institute involved (types in the linguistic and informetrics literature). The third column refers to the actual number of occurrences in the bylines of all these articles (tokens, in the linguistic and informetric literature). The ratio of the number of occurrences and the number of different articles (data can be found in Table 1) is termed the collaboration ratio. It is the average number of authors of the institute under study in articles where at least one member of the institute is a co-author.

The correlation between h_2-index and the rest of the indicators used in this study is in the majority of cases rather high (see Table 4).

The h_2-index correlates well with the total number of citations received by an institution, the R-index, the total number of citations received by the Hirsch core and the g-index.

Table 4. Correlation between the h_2-index and the other indicators calculated for the Cuban scientific production on brain research

Indicators	Pearson's correlation with h_2-index
TC	0.888
R-index	0.886
TC Hcore	0.884
g-index	0.882
h_1-index	0.863
CxA	0.737
Total of Authors	0.681
Total of Occurrences	0.634
Average Impact Factor	0.561
Collaboration ratio	–0.066

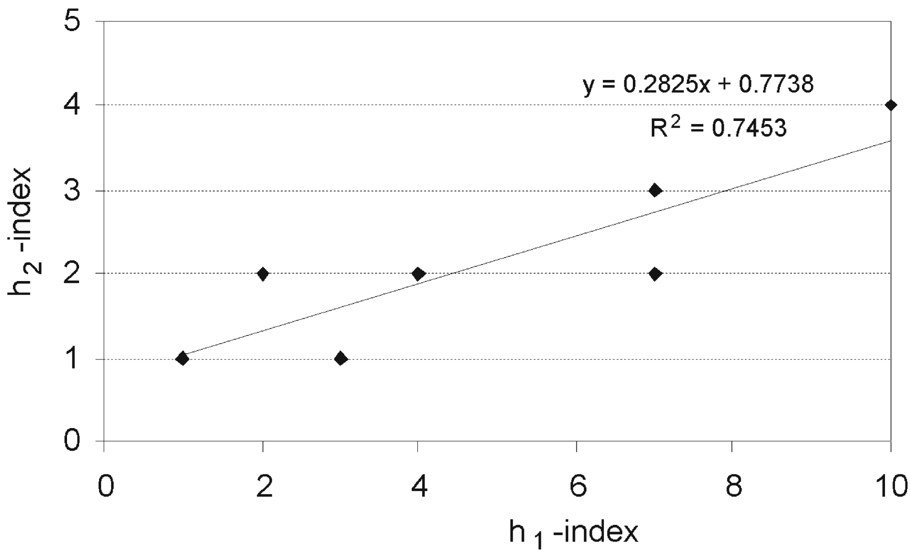

Figure 1. Linear regression between h_1-index and h_2-index

This correlation is lower for the average of citations by article, the total of authors and occurrences, and the average of the impact factor of journals where the papers were published. This relation with the impact factor is not surprising, as it has been stated again and again that the impact factor of journals used to publish should not be used for institutional evaluation, but nevertheless interesting. It shows why the use of *h*-indices or successive *h*-indices can be considered as an acceptable way of comparing institutes.

There is no correlation between h_2-index and the collaboration ratio, which implies that the impact of an institution, according to successive h-indices, is not related with the number of authors that usually collaborate in institutional research.

Particularly, the correlation between h_1-index and the h_2-index, the two main indicators of the rankings presented in this study, is strong ($r = 0.863$) (see Figure 1). This result suggests that these indicators are mutually dependent. Thus, if an institution has a large h_1-index, which means that it has many highly cited articles, the probability increases that many scientists with a high h-index work at this institute.

Conclusions

The application of the Hirsch index, Schubert's idea of successive h-indices and in particular Prathap's approach has been shown to lead to interesting new insights for the evaluation of Cuban institutions related to brain research. Prathap's approach allowed us to measure the performance of institutions from two different perspectives. On the one hand, it is possible to see the impact of the institutional scientific output; on the other hand, the institutional relevance is brought forward through the impact (as measured by their h-indices) of the researcher staff.

Both indicators were strongly related in the studied sample. Institutions with a large number of highly cited articles had the best probability also to employ staff with a high h-index. As this article is just a case study for one field in one country, more case studies are necessary in order to gain more experience in the use of successive h-indices and Prathap's approach.

We are confident though that in the near future, successive h-indices will be used as practical tools to measure the performance of the intellectual capital of research institutions, universities, scientific societies, countries, and other higher order units.

*

The authors thank Profs Felix de Moya Anegón and Gangan Prathap for their suggestions and advice. Research for this article was supported by the doctoral program on Documentation and Scientific Information developed by University of Granada, Spain, and the University of Havana, Cuba.

References

ARENCIBIA-JORGE, R., BARRIOS-ALMAGUER, I., FERNANDEZ-HERNANDEZ, S., CARVAJAL-ESPINO, R. (2008), Applying successive H indices in the institutional evaluation: a case study. *Journal of the American Society for Information Science and Technology*, 59 : 1–3.

EGGHE, L. (2006A), An improvement of the *H*-index: the *G*-index. *ISSI Newsletter*, 2 : 8–9.

EGGHE, L. (2006B), Theory and practice of the g-index. *Scientometrics*, 69 : 131–152.

EGGHE, L. (2007), Modelling successive h-indices. Preprint.

EGGHE, L., RAO, I. K. R. (2008), Study of different h-indices for groups of authors. *Journal of the American Society for Information Science & Technology*, 59 : 1276–1281.
GLÄNZEL, W. (2006), On the opportunities and limitations of the H-index. *Science Focus (in Chinese)*, 1 : 10–11.
HIRSCH, J. E. (2005), An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science of the USA*, 102 : 16569–16572.
JIN, B. (2006), H-index: an evaluation indicator proposed by scientist. *Science Focus (in Chinese)*, 1 : 8–9.
JIN, B. (2007), The AR-index complementing the h-index. *ISSI Newsletter*, 3 : 6.
JIN, B., LIANG, L., ROUSSEAU, R., EGGHE, L. (2007), The *R*- and *AR*-indices: complementing the h-index. *Chinese Science Bulletin*, 52 : 855–863.
KOSMULSKI, M. (2006), A new Hirsch-type index saves time and works equally well as the original *h*-index. *ISSI Newsletter*, 2 : 4–6.
PRATHAP, G. (2006), Hirsch-type indices for ranking institutions' scientific research output. *Current Science*, 91 : 1439.
RAO, I. K. R. (2007), Distributions of Hirsch-index and g-index: an empirical study, In: D. TORRES-SALINAS, H. F. MOED (Eds), *Proceedings of ISSI 2007*, Madrid, CSIC, 655–658.
RUANE, F., TOL, R. S. J. (2008), Rational (successive) h-indices: an application to economics in the republic of Ireland. *Scientometrics*, 75 : 395–405.
SCHILDT, H. A., MATTSSON, J. T. (2006), A dense network sub-grouping algorithm for co-citation analysis and its implementation in the software tool Sitkis. *Scientometrics*, 67 : 143–163.
SCHUBERT, A. (2007), Successive h-indices. *Scientometrics*, 70 : 201–205.
VAN RAAN, A. F. J. (2006), Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67 : 491–502.

Appendix A
Abbreviations of the most productive Cuban institutions in brain research

Abbreviation	Name of the institution
CNC	Cuban Neurosciences Center, Havana City
CIREN	International Center for Neurological Restoration, Havana City
UH	University of Havana, Havana City
CIRAH	Centre for Research and Rehabilitation of the Hereditary Ataxias, Holguin
HCQ-HAM	Clinical and Surgical Hospital "Hermanos Ameijeiras", Havana City
INN	Institute of Neurology and Neurosurgery, Havana City
HCQ-AMC	Clinical and Surgical Hospital "Armando Milián Castro", Villa Clara
ISCMH	Higher Institute of Medical Science of Havana, Havana City
HM-CJF	Military Hospital "Carlos J. Finlay", Havana City
CIMEQ	Medical Surgical Research Center, Havana City
HCQ-MAD	Clinical and Surgical Hospital "Manuel Ascunce Domenech", Camagüey
HPH	Psychiatric Hospital of Havana, Havana City