Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 69, No. 1 (2006) 131–152

Theory and practise of the *g*-index

LEO EGGHE[a,b]

[a] *Universiteit Hasselt (UHasselt), Campus Diepenbeek (Belgium)*
[b] *Universiteit Antwerpen (UA), Campus Drie Eiken, Wilrijk (Belgium)*

The *g*-index is introduced as an improvement of the *h*-index of Hirsch to measure the global citation performance of a set of articles. If this set is ranked in decreasing order of the number of citations that they received, the *g*-index is the (unique) largest number such that the top *g* articles received (together) at least g^2 citations. We prove the unique existence of *g* for any set of articles and we have that $g \geq h$.

The general Lotkaian theory of the *g*-index is presented and we show that

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}}$$

where $\alpha > 2$ is the Lotkaian exponent and where *T* denotes the total number of sources.

We then present the *g*-index of the (still active) Price medallists for their complete careers up to 1972 and compare it with the *h*-index. It is shown that the *g*-index inherits all the good properties of the *h*-index and, in addition, better takes into account the citation scores of the top articles. This yields a better distinction between and order of the scientists from the point of view of visibility.

Received January 31, 2006

Address for correspondence:
LEO EGGHE
Hasselt University, Agoralaan, B–3590 Diepenbeek, Belgium
E-mail: leo.egghe@uhasselt.be

0138–9130/US $ 20.00

I. Introduction

Recently the physicist HIRSCH (2005)) introduced the so-called h-index – see also BALL (2005), BRAUN et al. (2005), GLÄNZEL (2006a,b), EGGHE & ROUSSEAU (2006). For any general "set of papers" one can arrange these papers in decreasing order of the number of citations they received. The h-index is then the largest rank $h = r$ such that the paper on this rank (and hence also all papers on rank $1,\dots,h$) has h or more citations. Hence the papers on ranks $h + 1$, $h + 2$, … have not more than h citations.

Although introduced by a physicist, this new science indicator has been well-received in scientometrics (informetrics). In the above mentioned references it was argued that the h-index is a simple single number incorporating publication as well as citation data (hence comprising quantitative as well as qualitative or visibility aspects) and hence has an advantage over numbers such as "number of significant papers" (which is arbitrary) or "number of citations to each of the (say) q most cited papers" (which again is not a single number). The h-index is also robust in the sense that it is insensitive to a set of uncited (or lowly cited) papers but also it is insensitive to one or several outstandingly highly cited papers. This last aspect can be considered as a drawback of the h-index. Let us discuss this point further.

Highly cited papers are, of course, important for the determination of the value h of the h-index. But once a paper is selected to belong to the top h papers, this paper is not "used" any more in the determination of h, as a variable over time. Indeed, once a paper is selected to the top group, the h-index calculated in subsequent years is not at all influenced by this paper's received citations further on: even if the paper doubles or triples its number of citations (or even more) the subsequent h-indexes are not influenced by this. We think it is an advantage of the h-index not to take into account the "tail" papers (with low number of citations) but it should (being a measure of overall citation performance) take into account the citation evolution of the most cited papers!

In order to overcome this disadvantage, whilst keeping the advantages of the h-index, we make the following remark: by definition of the h-index, the papers on rank $1,\dots,h$ each have at least h citations, hence these h papers together have at least h^2 citations. But it could well be (see examples further on) that the first $h + 1$ papers have together $(h + 1)^2$ or more citations (here we use the fact that, most probably, the top papers have much more than h citations) and the same might be true for ranks $h + 2$ (the top $(h + 2)$ papers having together at least $(h + 2)^2$ citations) or even higher.

Therefore, Egghe (2006a, c) introduced a simple variant of the h-index: the g-index.

Definition I.1: A set of papers has a g-index g if g is the highest rank such that the top g papers have, together, at least g^2 citations. This also means that the top $g + 1$ papers have less than $(g + 1)^2$ papers.

The following proposition (also remarked in EGGHE (2006a)) is trivial.

Proposition I.2: In all cases one has that

$$g \geq h \tag{1}$$

Proof: Since h satisfies the requirement that the top h papers have at least h^2 papers and since g is the largest number with this property, it is clear that $g \geq h$. □

An example shows the easy calculation of the h-index and the g-index. The data are the author's own citation data derived from the Web of Knowledge (WoK). It must be underlined, however, that the real citation data can be much higher due to several reasons:

– only source journals, selected by Thomson ISI are used,

– unclear citations (even to source journals, e.g. "to appear" etc.) are not counted in the WoK.

In the table below TC stands for the total number of citations for each paper on rank $r = 1,2,...$ and ΣTC stands for the cumulative number of citations to the papers on rank $1,...,r$ (for each r). The bold face typed numbers give the explanation for the h-index $h = 13$ and the g-index $g = 19$. Indeed $h = 13$ is the highest rank such that all papers on rank $1,...,h$ have at least 13 citations (and hence the papers on rank 14 or higher have not more than 13 citations). Also $g = 19$ is the highest rank such that the top 19 papers have at least $19^2 = 361$ citations (here $381 > 361$); on rank 20 we have $392 < 20^2 = 400$ citations.

Table 1. Ranking of the papers of L. Egghe according to their number of citations received (source: WoK)

TC	r	ΣTC	r^2
47	1	47	1
42	2	89	4
37	3	126	9
36	4	162	16
21	5	183	25
18	6	201	36
17	7	218	49
16	8	234	64
16	9	250	81
16	10	266	100
15	11	281	121
13	12	294	144
13	**13**	307	169
13	14	320	196
13	15	333	225
12	16	345	256
12	17	357	289
12	18	369	324
12	**19**	**381**	**361**
11	20	392	400

In the last section of this article we will compare the *h*- and *g*-indexes of the (active) Price medallists (updated calculations of the *h*-index as in GLÄNZEL & PERSSON (2005) and new calculations of the *g*-index) showing the advantage of the *g*-index above the *h*-index but in the next section we will give the mathematical theory of the *g*-index based on Lotka's law

$$f(j) = \frac{C}{j^{\alpha}} \tag{2}$$

$j \geq 1$, $C > 0$, $\alpha > 2$ (it will turn out that, if we let j to be arbitrary large – which we assume here for the sake of simplicity – we need to take $\alpha > 2$). In case of (2) we will show that (T = total number of sources (= papers here))

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} \ , \tag{3}$$

hence by GLÄNZEL (2006b) or EGGHE & ROUSSEAU (2006), since one showed there that $h = T^{\frac{1}{\alpha}}$, we have

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} h > h \ . \tag{4}$$

Also the relation of g with the total number A of items (= citations here) is given. Before this theory is developed we will, firstly, show the general existence theorem for the *g*-index: for any set of papers we always have that the *g*-index exists and is unique.
Note, cf. BRAUN et al. (2005), EGGHE (2006a), EGGHE & ROUSSEAU (2006), that any set of papers can be taken here, e.g. the papers of a scientist but also a year's production (articles) in a journal can be used.

II. Mathematical theory of the *g*-index

First we will give a mathematically exact definition of the *g*-index in continuous variables.

II.1 Mathematical definition of the g*-index*

Let $f(j)$ ($j \geq 1$) denote the general size-frequency function of the system (which can be more general than the papers-citation relation: we can work in general information production processes (IPPs) where we have sources that produce items – cf. EGGHE & ROUSSEAU (1990), EGGHE (2005)). We do not suppose f to be Lotkaian at this moment.

Let $g(r)$ ($r \in [0,T]$) denote the general rank-frequency function (the function $g(r)$ should not be confused with the *g*-index; we keep the $f(j)$ and $g(r)$ notation since this has been done in all previous articles and books on this topic – throughout the text it will be clear whether we deal with the function $g(r)$ or with the *g*-index g). The general (defining) relation between the functions $f(j)$ and $g(r)$ is as follows:

$$r = g^{-1}(j) = \int_j^\infty f(j')dj' \tag{5}$$

Indeed, if $r = g^{-1}(j)$ (the inverse of the function $g(r)$) then $g(r) = j$ and there are r sources with an item density value larger than or equal to j. Denote

$$G(r) = \int_0^r g(r')dr' \tag{6}$$

the cumulative number of items in the sources up to rank r (i.e. the top r sources).

Definition: The rank r is the *g*-index: $r = g$ of this system if r is the highest value such that

$$G(r) \geq r^2 \tag{7}$$

Note that this is the exact formulation of the *g*-index as proposed in Section I in practical systems.

II.2 Existence theorem for the g*-index*

Theorem II.2.1 Every general system has a unique *g*-index.

Proof: Define, for all $r \in]0,T]$

$$H(r) = \frac{G(r)}{r} \tag{8}$$

and we define $H(0) = \lim_{\substack{r \to 0 \\ >}} H(r) = g(0)$. We first prove that H strictly decreases on $[0,T]$.

Indeed

$$H'(r) = \frac{rg(r) - G(r)}{r^2} < 0$$

since

$$rg(r) < G(r) = \int_0^r g(r')dr'$$

since the function g is strictly decreasing (by (5)) for all values of $r \in\ ]0,T]$. Since

$$H(0) = \lim_{\substack{r \to 0 \\ >}} H(r)$$

we hence have that H strictly decreases on $[0,T]$. If $H(T) \geq T$ then $G(T) \geq T^2$ and since this is the largest possible value, we have the unique g-index $g = T$. Suppose now that $H(T) < T$. Define

$$F(r) = H(r) - r \tag{9}$$

$$F(r) = \frac{G(r)}{r} - r$$

Since $H(0) > 0$ (since the function g strictly decreases (by (5)) and by (8)) and $H(T) < T$ we have $F(0) > 0$ and $F(T) < 0$. Hence, since F is continuous, there is a value r such that $F(r) = 0$. By (8) and (9) we hence have the existence of a value r such that

$$G(r) = r^2$$

Note that this satisfies (7) and that it is the highest possible value that satisfies (7): indeed, H strictly decreases, so, for every value $r' > r$ we have

$$H(r') < H(r)$$

By (8):

$$\frac{G(r')}{r'} < \frac{G(r)}{r} = r$$

$$G(r') < rr' < r'^2$$

contradicting (7). Hence this unique r value is the g-index: $r = g$. Note that, except if $G(T) \geq T^2$, we can prove that the g-index always satisfies (7) with an equality sign instead of $\geq$. □

Now we will give formulae for the g-index in terms of parameters that appear in Lotkaian informetrics.

II.3 Formulae for the g*-index in Lotkaian systems*

If $G(T) \leq T^2$ then we know from the proof of Theorem II.2.1 that the g-index satisfies (7) with an equality sign:

$$G(g) = g^2 \tag{10}$$

Otherwise (if $G(T) > T^2$) we take $g = T$.

We have the following theorem.

Theorem II.3.1: Given the law of Lotka

$$f(j) = \frac{C}{j^{\alpha}} \tag{11}$$

$j \geq 1$, $C > 0$, $\alpha > 2$, we have that the *g*-index equals

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} \tag{12}$$

if $\left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} \leq T$ and $g = T$ if $g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} > T$.

Here T denotes the total number of sources.

Proof:

First proof:

The first (cf. (5))

$$r = g^{-1}(j) = \int_j^{\infty} f(j')dj'$$

$$= \frac{C}{\alpha - 1} j^{1-\alpha} \tag{13}$$

sources yield a total number of items (since $\alpha > 2$)

$$\int_j^{\infty} j' f(j')dj' = \frac{C}{\alpha-2} j^{2-\alpha} \tag{14}$$

(cf. also EGGHE (2005), Chapter II).

So, by (10) we have $r = g$ if

$$\frac{C}{\alpha-2} j^{2-\alpha} = g^2 \tag{15}$$

and if this g satisfies $g \leq T$ (otherwise take $g = T$).

By (13):

$$j = \left(\frac{(\alpha-1)r}{C}\right)^{\frac{1}{1-\alpha}} \tag{16}$$

(16) (for $r = g$) in (15) yields

$$g = \left[\frac{C}{\alpha - 2} \left(\frac{\alpha - 1}{C} \right)^{\frac{\alpha-2}{\alpha-1}} \right]^{\frac{\alpha-1}{\alpha}}$$

$$g = \left(\frac{\alpha - 1}{\alpha - 2} \right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}}$$

using that $T = \frac{C}{\alpha - 1}$ as follows from (13) by taking $j = 1$. This value is taken as the g-index if it is $\le T$ and we take $g = T$ if it is strictly larger than T.

Second proof:

Now we work directly with formula (10). Note that Lotka's law (11) is equivalent with Zipf's law

$$g(r) = \frac{B}{r^{\beta}} \tag{17}$$

$B, \beta > 0$, $r \in\ ]0,T]$ where we have the relations

$$B = \left(\frac{C}{\alpha - 1} \right)^{\frac{1}{\alpha-1}} \tag{18}$$

$$\beta = \frac{1}{\alpha - 1} \tag{19}$$

(cf. EGGHE (2005), Exercise II.2.2.6 but see also the Appendix in EGGHE & ROUSSEAU (2006) where a proof is given.).

Note that by (19) $\alpha > 2$ is equivalent with $0 < \beta < 1$. If that is the case, (10) gives

$$\int_0^g \frac{B}{r^{\beta}} dr = g^2$$

hence, since $0 < \beta < 1$

$$\frac{B}{1 - \beta} g^{1-\beta} = g^2$$

Hence

$$g = \left(\frac{B}{1-\beta} \right)^{\frac{1}{\beta+1}} \tag{20}$$

Now (18) and (19) in (20) again yield formula (12). □

Corollary II.3.2: If *g* is the *g*-index and *h* is the *h*-index of a Lotkaian system with exponent $\alpha > 2$, then

$$g = \left(\frac{\alpha-1}{\alpha-2} \right)^{\frac{\alpha-1}{\alpha}} h \tag{21}$$

(if this value is $\leq T$; otherwise $g = T$).

Proof: This follows readily from (12) and the fact that $h = T^{\frac{1}{\alpha}}$, see EGGHE & ROUSSEAU (2006) (also proved, approximatively, in GLÄNZEL (2006b)). □

Taking $j = 1$ in (13) and (14) we see that the total number of sources *T* equals $\frac{C}{\alpha - 1}$ and that the total number of items *A* equals $\frac{C}{\alpha - 2}$, hence

$$\mu = \frac{A}{T} = \frac{\alpha-1}{\alpha-2} \tag{22}$$

equals the average number of items per source (cf. also EGGHE (2005), Chapter II). Hence we have the following corollary

Corollary II.3.3: If μ is as above we have in case of (21)

$$g = \mu^{\frac{\alpha-1}{\alpha}} h \tag{23}$$

The *g*-index in function of α and *A* is as in Corollary II.3.4.

Corollary II.3.4: We have

$$g = \left(\frac{\alpha-1}{\alpha-2} \right)^{\frac{\alpha-2}{\alpha}} A^{\frac{1}{\alpha}} \tag{24}$$

(if this value is $\leq T$).

Proof: This follows readily from (22) and (12). □

We can also determine the item density *j* for which we have $r = g$. In practical cases this means the number of items in the source at rank *g*. Note that this is *h* for the *h*-index $r = h$, by definition of the Hirsch index.

For the g-index we have: if the value in (12) is $> T$ we have $j = g(T) = 1$ and if the value in (12) is $\leq T$ we substitute $r = g$ in (17), using (18) and (19) and the fact that

$T = \dfrac{C}{\alpha - 1}$, yielding

$$j = \left(\frac{\alpha-2}{\alpha-1}T\right)^{\frac{1}{\alpha}} \tag{26}$$

We immediately see that $j < h$ which is logical since $g > h$ and the item density is h in case $r = h$. Formula (26) presents a concrete formula for the item density cut-off place.

Note that, although α can be any value $\alpha > 2$, we do **not** have

$$\lim_{\substack{\alpha \to 2 \\ >}} j = 0 \ .$$

Indeed since the validity of (12) is limited to $g \leq T$ we have, by (12) that

$$\left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} \leq T$$

from which it follows that

$$\mu = \frac{A}{T} = \frac{\alpha-1}{\alpha-2} \leq T \tag{27}$$

This implies in (26) that

$$j = \left(\frac{T^2}{A}\right)^{\frac{1}{\alpha}} \geq 1 \ ,$$

by (27).

The case

$$\left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}} > T \tag{28}$$

(hence where we take $g = T$) occurs in the following case: from (28) it follows that

$$\frac{\alpha-1}{\alpha-2} > T \tag{29}$$

By (22) we have

$$\mu = \frac{A}{T} > T$$

hence

$$A > T^2 \qquad (30)$$

Equivalently, (29) gives the condition in α:

$$\alpha < \frac{2T-1}{T-1} \qquad (31)$$

Note that

$$\frac{2T-1}{T-1} > 2$$

so that (31) can occur (with the condition $\alpha > 2$).

Remark II.3.5: It might seem strange that $g = T$ is possible in this Lotkaian model. Note however that (22) implies that

$$\alpha = \frac{2A-T}{A-T}$$

So, for every T fixed, if we let $A \to \infty$ we have that $\alpha \to 2$ (but $\alpha > 2$) so that we are within the limitations of our theory. In this case we have

$$A = G(T) > T^2$$

and hence $g = T$.

In the next section we will apply the g-index to the publications and citations of the (still active) Price medallists and compare these g-indexes with the h-indexes of these same data.

III. Calculation and comparison of the *h*- and *g*-indexes of the (still active) Price medallists

In GLÄNZEL & PERSSON (2005), the h-indexes for the (still active) Price medallists are calculated. We could use these numbers and compare them with the here defined g-index. However for this we need to extend the tables in GLÄNZEL & PERSSON (2005) (since $g \geq h$) and it is hardly impossible to do this since we should do this for the maximal citing time August 2005 (since then the tables in GLÄNZEL & PERSSON (2005) were produced). So the easiest thing to do is to remake these tables for the present time

(January 2006) and make them long enough so that, on the same tables, the *h*- as well as the *g*-index can be calculated.

We have opted **not** to limit the publication year to 1986 or higher (as was the case in GLÄNZEL & PERSSON (2005)). Indeed, the *h*-indexes in GLÄNZEL & PERSSON (2005) seemed a bit unnatural in several senses. Garfield did not have the highest *h*-index (which we, normally, could expect) and Small scored lowest of all the Price medallists. The major reason for these observations is that, by limiting the publication year to 1986 or higher, one cuts away most publications (and perhaps the highest cited ones) of the relatively older scientists. Since we want to make a comparison of scientists (and not to draw conclusions on informetrics fields) we decided not to limit the publication year (except to the evident limit 1972 since before that date the ISI (now Thomson ISI) data do not exist).

For the same reason we count all publications even if scientists have published in different domains (e.g. T. Braun in chemistry and L. Egghe in mathematics). These publications were not used in the GLÄNZEL & PERSSON study.

Of course, by not limiting the publication period and the publication field, one might argue that there is a bias towards the older scientists. This is true but, with the *h*- and *g*-indexes, we want to indicate the "overall performance (visibility)" of the scientists as they are viewn today (in the sense of "lifetime achievement").

We base ourselves on the Web of Knowledge (WoK) and hence we are limited to the Thomson ISI data. This means that no citations to non-source journals or conference proceedings articles or books are counted. In addition, no citations to incomplete references are counted even if they are to source journal articles (e.g. a citation to *JASIST*, 2001, to appear): these are not collected in the WoK "times cited" data. So the actual *h*- and *g*-indexes can be somewhat higher but this effect plays for every scientist so that comparisons are still possible and also these limitations do not jeopardise the possibility to compare the *h*- and *g*-index.

The tables of citation data of the (still active) medallists are found in the Appendix. The table stops one line below the *g*-index since this is all we need. The number r denotes the rank of the publication and TC denotes the total number of citations to the paper on rank r. The number ΣTC denotes the cumulative number of citations to the first r ranked papers. Finally, also the table of r^2 values is presented as well as the publication year (PY) of the article on rank r. The *h*- and *g*-index determination is highlighted in the tables in the Appendix. Table 2 gives the results in decreasing order of h and g.

We leave the detailed (subjective) interpretation of Table 2 to the reader but it is clear that the *g*-index column is more in line with intuition and with the raw data in the Appendix than the *h*-index column. In other words, the *g*-index, as simple as the *h*-index (a single measure, containing publication and citation elements), contains more comparative information from the raw data than the *h*-index and resembles more the overall feeling of "visibility" or "life time achievement".

A possible interesting measure is g/h, i.e. the relative increase of *g* with respect to *h*. The result is presented in Table 3, in decreasing order of g/h. Here we see remarkable order changes with respect to the *h*- or *g*-orderings.

Table 2. *h*- and *g*-indexes of Price medallists in decreasing order

Name	*h*-index
Garfield	27
Narin	27
Braun	25
Van Raan	19
Glänzel	18
Moed	18
Schubert	18
Small	18
Martin	16
Egghe	13
Ingwersen	13
Leydesdorff	13
Rousseau	13
White	12

Name	*g*-index
Garfield	59
Narin	40
Small	39
Braun	38
Schubert	30
Glänzel	27
Martin	27
Moed	27
Van Raan	27
Ingwersen	26
White	25
Egghe	19
Leydesdorff	19
Rousseau	15

Table 3. g/h-values of Price medallists in decreasing order

Name	g/h
Garfield	2.19
Small	2.17
White	2.08
Ingwersen	2.00
Martin	1.69
Schubert	1.67
Braun	1.52
Glänzel	1.50
Moed	1.50
Narin	1.48
Egghe	1.46
Leydesdorff	1.46
Van Raan	1.42
Rousseau	1.15

IV. Conclusions and open problems

In this paper we studied the *g*-index being an improvement of the *h*-index. The *g*-index *g* is the largest rank (where papers are arranged in decreasing order of the number of citations they received) such that the first *g* papers have (together) at least g^2 citations. We show that $g \geq h$ and that *g* always uniquely exists. We present formulae for *g* in Lotkaian informetrics. We show that

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} T^{\frac{1}{\alpha}}$$

$$g = \left(\frac{\alpha-1}{\alpha-2}\right)^{\frac{\alpha-1}{\alpha}} h$$

if these values are $\leq T$; otherwise $g = T$.

Here α is the Lotka exponent and *T* denotes the total number of sources (in the citation application this means the total number of ever cited papers).

We then calculate the *h*- and *g*-indexes of the (still active) Price medallists. Different than in GLÄNZEL & PERSSON (2005) we do not limit the publication period (except for the fact that we do not use papers older than published in 1972 due to the fact that ISI has no data for them) nor do we limit the topic to informetrics, hence the complete careers (up to 1972) of the Price medallists are considered. It is found that the ranked *g*-index column resembles more the overall feeling of "visibility" or "life time achievement" than does the ranked *h*-index column.

We leave open the further exploration of the *g*-index, including the establishment of the *g*-index in function of time. In EGGHE (2006b) we were able to do this for the *h*-index based on the cumulative n^{th} citation distribution (see EGGHE & RAO (2001)) and in a forthcoming paper we will do the same for the *g*-index based on a time-dependent Lotkaian theory.

We also leave open the construction of other *h*- or *g*-like indexes and the comparison of these new indexes with the *h*- and *g*-index. It would also be interesting to work out more practical cases (in other fields) of *h*- and *g*-index comparisons. Such case studies can learn a lot on the advantages and/or disadvantages of the *h*-index and the *g*-index.

*

The author is grateful to Drs. M. Goovaerts for the preparation of the citation data of the Price medallists (January 2006).

Note added in proof

A small variant of the g-index is possible by not limiting it to $g \leq T$. In practical examples this means that, in these cases, fictitious articles with 0 citations have to be added.

Example: only 3 papers exist and are cited. The other ones are added with 0 citations until the g-index can be determined

rank	# citations	cum. citations	r^2
1	20	20	1
2	10	30	4
3	5	35	9
4	0	35	16
5	0	35	25
6	0	35	36

Here $g = 5 > 3$. Of course, $h = 3$ and $h \leq T$ always. This is considered by GLÄNZEL (2006a) to be a drawback of the h-index, giving a small h value to a small (but highly cited) article set (he calls it "small is not beautiful").

All results proved in this article remain the same: the Lotkaian model (12) for g is now always valid (also in this model g can be $>T$). In the existence theorem II.2.1, if $H(T) \geq T$ we go further in the fictitions ranks such that $\mathrm{H}(r) < r$ and the proof continues as in the case $\mathrm{H}(T) < T$.

In the case of the price medallist (Section III), no cases where $g > T$ were found.

References

BALL, P. (2005), Index aims for fair ranking of scientists. *Nature*, 436 : 900.

BRAUN, T., GLÄNZEL, W., SCHUBERT, A. (2005), A Hirsch-type index for journals. *The Scientist*, 19 (22) : 8.

EGGHE, L. (2005), *Power Laws in the Information Production Process: Lotkaian Informetrics*. Elsevier, Oxford (UK).

EGGHE, L. (2006a), An improvement of the h-index: The g-index. *ISSI Newsletter*, 2 (1) : 8–9.

EGGHE, L. (2006b), Dynamic h-index: The Hirsch index in function of time. *Journal of the American Society for Information Science and Technology*, to appear.

EGGHE, L. (2006c), How to improve the h-index: Letter. *The Scientist*, 20 (3) : 14, March 2006.

EGGHE, L., RAO, I. K. R. (2001), Theory of first-citation distributions and applications. *Mathematical and Computer Modelling*, 34 (1-2) : 81–90.

EGGHE, L., ROUSSEAU, R. (1990), *Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Science.* Elsevier, Amsterdam.

EGGHE, L., ROUSSEAU, R. (2006), An informetric model of the Hirsch index. *Scientometrics*, 69 (1) : 000–000.

GLÄNZEL, W. (2006a), On the opportunities and limitations of the H-index. *Science Focus*, 1, to appear.

GLÄNZEL, W. (2006b), On the H-index – A mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, 67 (2) : 315–321.
GLÄNZEL ,W., PERSSON, O. (2005), H-index for Price medallist. *ISSI Newsletter*, 1 (4) : 15–18.
HIRSCH, J. E. (2005), An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Scieneces of the United States of America,* 102 : 16569–16572.

Appendix

Tables of TC, *r*, ΣTC, r^2 and PY for each of the (still active) Price medallists and determination of the *h*- and *g*-index.

Garfield E.

TC	*r*	ΣTC	r^2	PY	TC	*r*	ΣTC	r^2	PY
625	1	625	1	1972	23	31	3146	961	1990
149	2	774	4	1980	20	32	3166	1024	1990
138	3	912	9	1977	19	33	3185	1089	1998
132	4	1044	16	1983	19	34	3204	1156	1998
132	5	1176	25	1981	18	35	3222	1225	1985
129	6	1305	36	1979	18	36	3240	1296	1979
127	7	1432	49	1996	18	37	3258	1369	1996
111	8	1543	64	1978	16	38	3274	1444	1979
109	9	1652	81	1975	15	39	3289	1521	1990
108	10	1760	100	1985	14	40	3303	1600	1976
107	11	1867	121	1984	13	41	3316	1681	1973
105	12	1972	144	1982	13	42	3329	1764	1973
104	13	2076	169	1986	13	43	3342	1849	1973
101	14	2177	196	1976	13	44	3355	1936	1998
96	15	2273	225	1973	13	45	3368	2025	1990
91	16	2364	256	1976	12	46	3380	2116	1973
89	17	2453	289	1974	12	47	3392	2209	2000
88	18	2541	324	1986	12	48	3404	2304	1998
87	19	2628	361	1987	12	49	3416	2401	1997
85	20	2713	400	1979	12	50	3428	2500	1996
80	21	2793	441	1985	11	51	3439	2601	1998
67	22	2860	484	1988	11	52	3450	2704	1997
63	23	2923	529	1999	10	53	3460	2809	1985
41	24	2964	576	1980	10	54	3470	2916	1984
29	25	2993	625	1990	9	55	3479	3025	1984
28	26	3021	676	1987	9	56	3488	3136	1975
27	27	3048	729	1987	9	57	3497	3249	1972
26	28	3074	784	1976	9	58	3506	3364	2002
26	29	3100	841	1992	9	59	3515	3481	1998
23	30	3123	900	1978	9	60	3524	3600	1990

Braun T.

TC	r	ΣTC	r^2	PY
125	1	125	1	1978
124	2	249	4	1989
78	3	327	9	1986
66	4	393	16	1975
57	5	450	25	1974
57	6	507	36	1990
55	7	562	49	1974
51	8	613	64	1989
43	9	656	81	1992
42	10	698	100	1974
38	11	736	121	1983
37	12	773	144	1995
37	13	810	169	1994
35	14	845	196	1980
35	15	880	225	1999
33	16	913	256	1988
32	17	945	289	1995
31	18	976	324	1975
31	19	1007	361	1995
28	20	1035	400	1977
27	21	1062	441	1973
27	22	1089	484	1988
27	23	1116	529	1987
26	24	1142	576	2000
26	25	1168	625	1994
25	26	1193	676	1973
25	27	1218	729	1972
23	28	1241	784	1978
23	29	1264	841	1973
23	30	1287	900	1994
23	31	1310	961	1987
22	32	1332	1024	1983
22	33	1354	1089	1982
22	34	1376	1156	1980
22	35	1398	1225	1987
21	36	1419	1296	1973
21	37	1440	1369	1973
20	38	1460	1444	1982
20	39	1480	1521	1982

Small H.

TC	r	ΣTC	r^2	PY
305	1	305	1	1973
239	2	544	4	1974
127	3	671	9	1978
109	4	780	16	1974
86	5	866	25	1977
80	6	946	36	1985
77	7	1023	49	1985
75	8	1098	64	1985
67	9	1165	81	1999
49	10	1214	100	1979
44	11	1258	121	1980
36	12	1294	144	1980
26	13	1320	169	1981
26	14	1346	196	1986
25	15	1371	225	1976
22	16	1393	256	1997
22	17	1415	289	1993
18	18	1433	324	1974
18	19	1451	361	1994
15	20	1466	400	1999
12	21	1478	441	1986
10	22	1488	484	1989
9	23	1497	529	1975
8	24	1505	576	1998
8	25	1513	625	1987
7	26	1520	676	1989
6	27	1526	729	1998
5	28	1531	784	1977
5	29	1536	841	1974
5	30	1541	900	1999
3	31	1544	961	1979
3	32	1547	1024	1995
2	33	1549	1089	1975
2	34	1551	1156	2004
2	35	1553	1225	2003
1	36	1554	1296	1973
1	37	1555	1369	2004
1	38	1556	1444	1997
1	39	1557	1521	1996
1	40	1558	1600	1992

Van Raan A.F.J.

TC	*r*	ΣTC	r^2	PY	TC	*r*	ΣTC	r^2	PY
108	1	108	1	1985	20	15	535	225	2001
51	2	159	4	1996	19	16	554	256	1998
49	3	208	9	1991	19	17	573	289	1998
41	4	249	16	1985	19	18	592	324	1994
35	5	284	25	1991	19	19	611	361	1994
32	6	316	36	1973	18	20	629	400	1998
31	7	347	49	1990	18	21	647	441	1993
30	8	377	64	1990	17	22	664	484	1993
25	9	402	81	1993	17	23	681	529	1985
25	10	427	100	1974	17	24	698	576	1980
23	11	450	121	1995	15	25	713	625	1993
22	12	472	144	1998	14	26	727	676	2001
22	13	494	169	1997	14	27	741	729	1994
21	14	515	196	2000	14	28	755	784	1991

Martin B.

TC	*r*	ΣTC	r^2	PY	TC	*r*	ΣTC	r^2	PY
156	1	156	1	1983	19	15	616	225	1985
74	2	230	4	1997	18	16	634	256	1986
52	3	282	9	1985	16	17	650	289	1996
38	4	320	16	1983	16	18	666	324	1986
35	5	355	25	2001	16	19	682	361	1985
33	6	388	36	1987	16	20	698	400	1984
33	7	421	49	1985	14	21	712	441	1991
30	8	451	64	1995	14	22	726	484	1984
29	9	480	81	1996	11	23	737	529	1986
28	10	508	100	1984	9	24	746	576	1994
24	11	532	121	1988	9	25	755	625	1989
23	12	555	144	1981	9	26	764	676	1987
22	13	577	169	1999	6	27	770	729	1982
20	14	597	196	1984	4	28	774	784	1992

Narin F.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
112	1	112	1	1997	29	22	1268	484	1994
95	2	207	4	1987	28	23	1296	529	1996
86	3	293	9	1976	28	24	1324	576	1977
82	4	375	16	1976	28	25	1352	625	1976
73	5	448	25	1977	27	26	1379	676	1984
71	6	519	36	1991	27	27	1406	729	1983
70	7	589	49	1972	26	28	1432	784	1988
63	8	652	64	1985	24	29	1456	841	1988
59	9	711	81	1992	23	30	1479	900	1995
55	10	766	100	1978	20	31	1499	961	1998
55	11	821	121	1973	19	32	1518	1024	1994
53	12	874	144	1975	18	33	1536	1089	1980
52	13	926	169	1991	18	34	1554	1156	1979
52	14	978	196	1981	17	35	1571	1225	1978
44	15	1022	225	1977	14	36	1585	1296	1996
41	16	1063	256	1980	13	37	1598	1369	1983
38	17	1101	289	2000	12	38	1610	1444	1986
37	18	1138	324	1980	10	39	1620	1521	1977
35	19	1173	361	1999	10	40	1630	1600	1972
33	20	1206	400	1989	9	41	1639	1681	1983
33	21	1239	441	1987					

Schubert A.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
124	1	124	1	1989	19	17	733	289	1986
90	2	214	4	2002	18	18	751	324	2000
78	3	292	9	1986	18	19	769	361	1993
59	4	351	16	1978	18	20	787	400	1986
57	5	408	25	1990	18	21	805	441	1984
40	6	448	36	1979	17	22	822	484	2001
33	7	481	49	1988	17	23	839	529	1988
32	8	513	64	1983	17	24	856	576	1982
27	9	540	81	1988	16	25	872	625	1982
27	10	567	100	1987	15	26	887	676	2002
27	11	594	121	1984	14	27	901	729	1993
26	12	620	144	2000	14	28	915	784	1989
26	13	646	169	1994	14	29	929	841	1985
23	14	669	196	1994	13	30	942	900	1992
23	15	692	225	1987	12	31	954	961	1996
22	16	714	256	1987					

Glänzel W.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
124	1	124	1	1989	22	15	533	225	2002
54	2	178	4	1988	22	16	555	256	1987
33	3	211	9	1988	20	17	575	289	1994
32	4	243	16	1995	19	18	594	324	1986
32	5	275	25	1983	18	19	612	361	1994
31	6	306	36	1995	18	20	630	400	1993
28	7	334	49	1995	18	21	648	441	1986
27	8	361	64	1988	18	22	666	484	1984
27	9	388	81	1987	17	23	683	529	2001
27	10	415	100	1984	17	24	700	576	1988
26	11	441	121	1994	16	25	716	625	1999
24	12	465	144	2001	16	26	732	676	1996
23	13	488	169	1994	15	27	747	729	1997
23	14	511	196	1987	15	28	762	784	1996

Moed F. H.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
108	1	108	1	1985	22	15	594	225	1991
56	2	164	4	1995	20	16	614	256	2001
54	3	218	9	1996	20	17	634	289	1999
54	4	272	16	1995	18	18	652	324	1989
49	5	321	25	1991	17	19	669	361	1985
41	6	362	36	1985	15	20	684	400	1999
35	7	397	49	1991	15	21	699	441	1993
31	8	428	64	1990	13	22	712	484	1998
26	9	454	81	2002	13	23	725	529	1993
26	10	480	100	1989	12	24	737	576	2002
24	11	504	121	1996	12	25	749	625	1993
23	12	527	144	1999	9	26	758	676	1999
23	13	550	169	1998	9	27	767	729	1996
22	14	572	196	2002	9	28	776	784	1996

Leydesdorff L.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
79	1	79	1	2000	15	11	290	121	1994
32	2	111	4	1998	13	12	303	144	1994
26	3	137	9	1986	13	13	316	169	1993
24	4	161	16	1989	13	14	329	196	1989
23	5	184	25	1990	11	15	340	225	2000
22	6	206	36	1987	11	16	351	256	1993
19	7	225	49	1989	11	17	362	289	1992
17	8	242	64	1996	10	18	372	324	1998
17	9	259	81	1991	10	19	382	361	1997
16	10	275	100	1997	9	20	391	400	1992

Egghe L.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
47	1	47	1	1990	15	11	281	121	1993
42	2	89	4	1985	13	12	294	144	1996
37	3	126	9	2000	13	13	307	169	1996
36	4	162	16	1992	13	14	320	196	1990
21	5	183	25	1992	13	15	333	225	1988
18	6	201	36	1991	12	16	345	256	2000
17	7	218	49	1986	12	17	357	289	1994
16	8	234	64	1995	12	18	369	324	1988
16	9	250	81	1988	12	19	381	361	1987
16	10	266	100	1986	11	20	392	400	2000

Rousseau R.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
25	1	25	1	1996	13	9	150	81	2002
18	2	43	4	2003	13	10	163	100	1999
18	3	61	9	1991	13	11	176	121	1996
16	4	77	16	1995	13	12	189	144	1996
16	5	93	25	1988	13	13	202	169	1993
15	6	108	36	1987	12	14	214	196	2000
15	7	123	49	1992	12	15	226	225	2000
14	8	137	64	1994	12	16	238	256	1990

Ingwersen P.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
120	1	120	1	1996	10	15	638	225	1992
93	2	213	4	1998	8	16	646	256	1999
83	3	296	9	1997	7	17	653	289	1999
79	4	375	16	1982	7	18	660	324	1995
52	5	427	25	2001	6	19	666	361	2000
37	6	464	36	1997	6	20	672	400	1993
31	7	495	49	1984	5	21	677	441	2000
29	8	524	64	1997	3	22	680	484	2001
29	9	553	81	1987	3	23	683	529	2000
19	10	572	100	2000	3	24	686	576	2000
17	11	589	121	1984	3	25	689	625	1994
15	12	604	144	1996	3	26	692	676	1994
14	13	618	169	1997	3	27	695	729	1992
10	14	628	196	2001					

White H.D.

TC	r	ΣTC	r^2	PY	TC	r	ΣTC	r^2	PY
128	1	128	1	1981	12	14	577	196	1996
106	2	234	4	1998	12	15	589	225	1981
103	3	337	9	1989	12	16	601	256	1981
45	4	382	16	1997	11	17	612	289	2003
37	5	419	25	1982	10	18	622	324	1990
28	6	447	36	1981	8	19	630	361	1986
22	7	469	49	1983	6	20	636	400	2001
21	8	490	64	1987	5	21	641	441	2004
20	9	510	81	2001	5	22	646	484	1986
15	10	525	100	1987	5	23	651	529	1984
14	11	539	121	1986	5	24	656	576	1977
14	12	553	144	1985	4	25	660	625	2003
12	13	565	169	2003	4	26	664	676	1986