TBA

????

TAILOR BASED ALLOCATIONS
for multiple authorship

Serge Galam

Centre de Recherche en Épistémologie Appliquée (CREA)
École Polytechnique, Paris
and
National Center for Scientific Research (CNRS)

Multiple authorship
and
the "Jesus multiplying breads" principle

S. Galam, MKD, Amsterdam, Oct 2010

Bibliometry

Individual

and

aggregated

Today major instrument

Allocate research funds

Promoting

Ranking

Recruiting

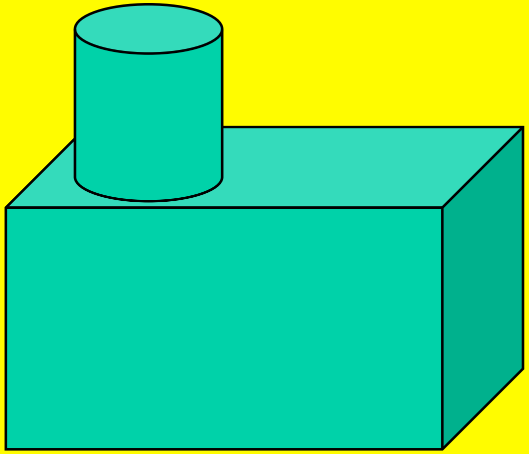
The now famous h-index has boosted the use of quantitative measures of scientist productions

Its incorporation within the Web of Science via a simple evaluation button have just turned upside-down the world of evaluation

Now widespread

And

unavoidable



The main bibliometrical resource for evaluating research is the publications and thus the number of publications

Once these numbers are known they are « treated » to produce a use value

The h-index focuses on ranking papers by their decreasing numbers of citations

Quantitative evaluations are based on individual outputs

Personal list of publications

A series of personal list of publications

One paper cosigned by 3 authors

Each one of the 3 authors includes the paper within its own list of publications

Counting the productions of papers through individual productions

The one paper becomes 3 papers

One paper cosigned by 3 authors is cited n times

Each one of the 3 authors incorporates the n citations in its numbers of citations

Counting the citations of papers through individual lists

The n citations are multiplied to 3n citations

The h-index of each one of the 3 authors is calculated using the same n citations

No matter the contribution of each author, each one gets the same credit for the common paper

Favors multiauthorship

Moreover

for the h-index it does not matter to have a paper which scores a high record of citations

Evaluation of the h-index

1 -> n+Δ1

2 -> n+Δ2

3 -> n+Δ3

…

k -> n+Δ k

…

n -> n+Δ n

n+1 -> n-1-Δ(n +1)

The minimum numbers of citations to have a h-index=n is to have n papers with at least n citations each

Δk = 0 for all k

More citations to each one of the n ranked papers does not modify h

So indeed we are violating the mere principle of scientific methodology

The “Jesus” principle : one -> several

Magic based on overruling the limitations of natural constraints

≠

Ruling within the limitations of natural constraints by the laws of conservation

The “science” principle : one -> one

Restoring the conservation of printed articles

All credits allocated to all authors of one single paper must add to ONE

How to proceed?

All numbers of citations used by all authors of one single paper must add to the actual number this paper got

Allocate a fractional number $g_{r,k}$ for the author at rank r for a k authors paper

$$\sum_{r=1}^{k} g_{r,k} = 1$$

The total number of papers of one author is no longer T the number of papers it has cosigned

$$\sum_{i=1}^{T} g_{r,k}^{i} = G < T$$

Allocate a fractional number of citations for the author at rank r for a k authors paper

$$g_{r,k}n$$

The total number of citations of one author is scaled down

$$\sum_{i=1}^{T} g_{r,k}^{i} n^{i}$$

The h-index of one author is evaluated according to the same previous scheme but now using for the paper number i the quantity

$$g_{r,k}^{i} n^{i}$$

$$h_g < h$$

How to implement the $g_{r,k}$ allocation?

The best is having the authors decide for each paper priori to its publication

Fair and accurate but for the future

What to do with existing papers?

We need an accepted scheme to set a unified standard

Can vary from one discipline to another

I do not intend in promoting one specific policy to favor or discourage conducting collaboration

Build a frame to capture the current practice and to allow adaptation to fit different policies

In condensed matter physics, usually

The first author has been the working force
The last author has been the driving force
The in between authors have made specific contributions

Various suggestions in the literature:

The simplest equalitarian fractional allocating : 1/k

The arithmetic allocating allocating : 2(k+1-r)/[k(k+1)]

The geometric fractional allocating : 2^(k-r)/(2^k-1)

The awkward equalitarian and boss: 1/2(k-1) and 1/2

My suggestion: allocate extra bonuses to first and last authors within a modified non-linear arithmetic ranking

$$\begin{cases} g_{1,k} = \dfrac{k+\delta}{S_k} \\ g_{k,k} = \dfrac{k-1+\mu}{S_k} \\ g_{r,k} = \dfrac{k-r}{S_k} \end{cases}$$

$$S_k = \frac{k(1+k)}{2} + \delta + \mu$$

$$2 \le r \le k-1$$

$\delta : ?$

$\mu : ?$

Two author case:

$$g_{1,2} = \frac{2+\delta}{S_2}$$

$$g_{2,2} = \frac{1+\mu}{S_2}$$

$$S_2 = 3 + \delta + \mu$$

2/3;1/3 → $\delta = 2\mu$

3/4;1/4 → $\delta = 1 + 3\mu$

1/2;1/2 → $\delta = -1 + \mu$

Fixing the two extreme ranks leaves one degree of freedom

Table 1: The various $g_{r,k}$ when $\delta = 2$ and $\mu = 1$ for $1 \le r \le k$ and $1 \le k \le 10$. It yields two third one third at $k = 2$. Last column corresponds to last author. Second column corresponds to second author. It thus does not incorporate the second author when it is also the last one. Accordingly, second author is defined only from $k \ge 3$. The same rule applies to all columns from $r = 2$ up to $r = 9$. Last case is defined only at $k = 10$.

r =	1	2	3	4	5	6	7	8	9	k
k										
1	1	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx
2	0.67	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	0.33
3	0.56	0.11	xxx	xxx	xxx	xxx	xxx	xxx	xxx	0.33
4	0.46	0.15	0.08	xxx	xxx	xxx	xxx	xxx	xxx	0.31
5	0.39	0.17	0.11	0.05	xxx	xxx	xxx	xxx	xxx	0.28
6	0.33	0.17	0.12	0.08	0.04	xxx	xxx	xxx	xxx	0.25
7	0.29	0.16	0.13	0.10	0.06	0.03	xxx	xxx	xxx	0.23
8	0.26	0.15	0.13	0.10	0.08	0.05	0.02	xxx	xxx	0.20
9	0.23	0.15	0.12	0.10	0.08	0.06	0.04	0.02	xxx	0.19
10	0.21	0.14	0.12	0.10	0.09	0.07	0.05	0.03	0.02	0.17

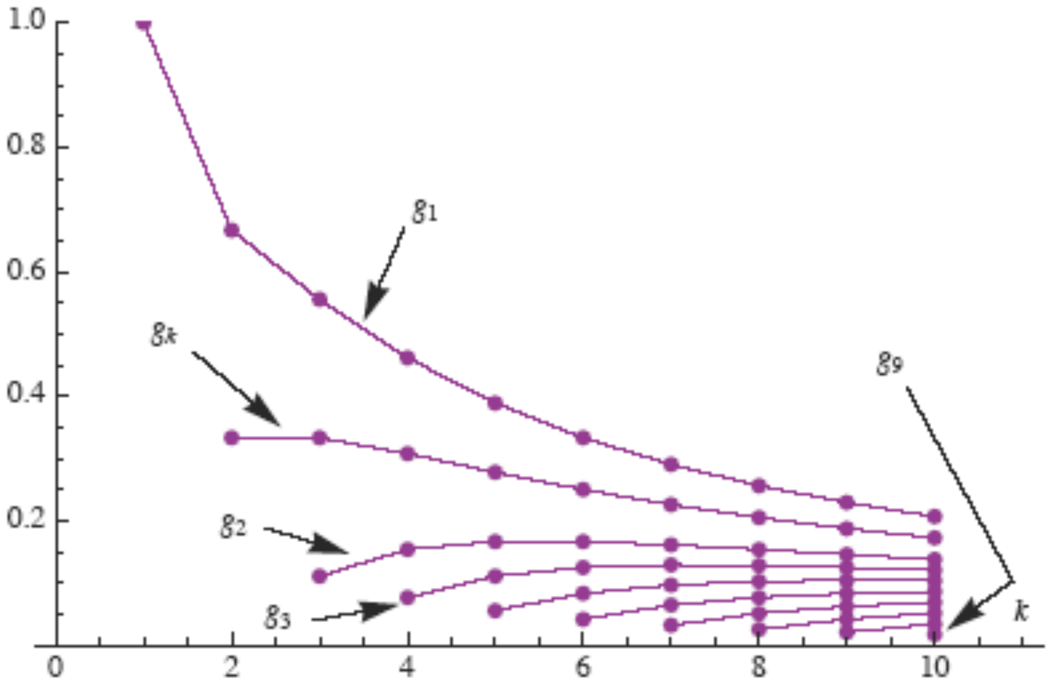
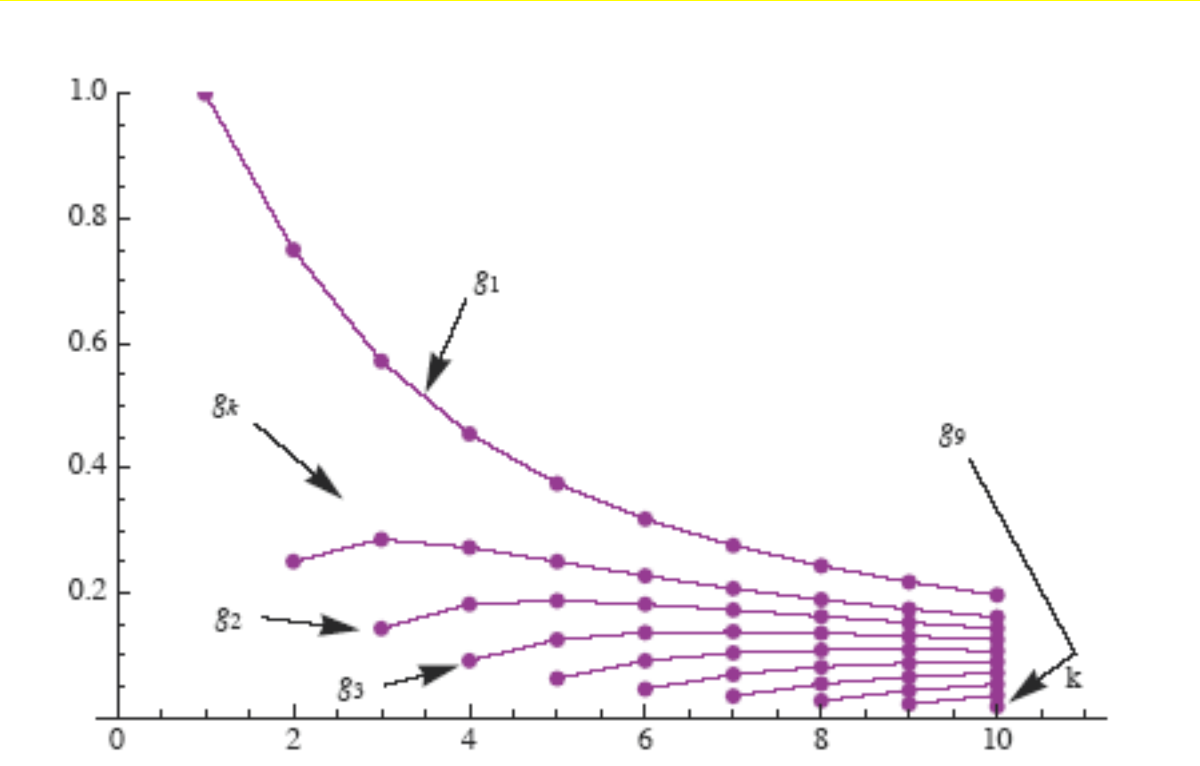
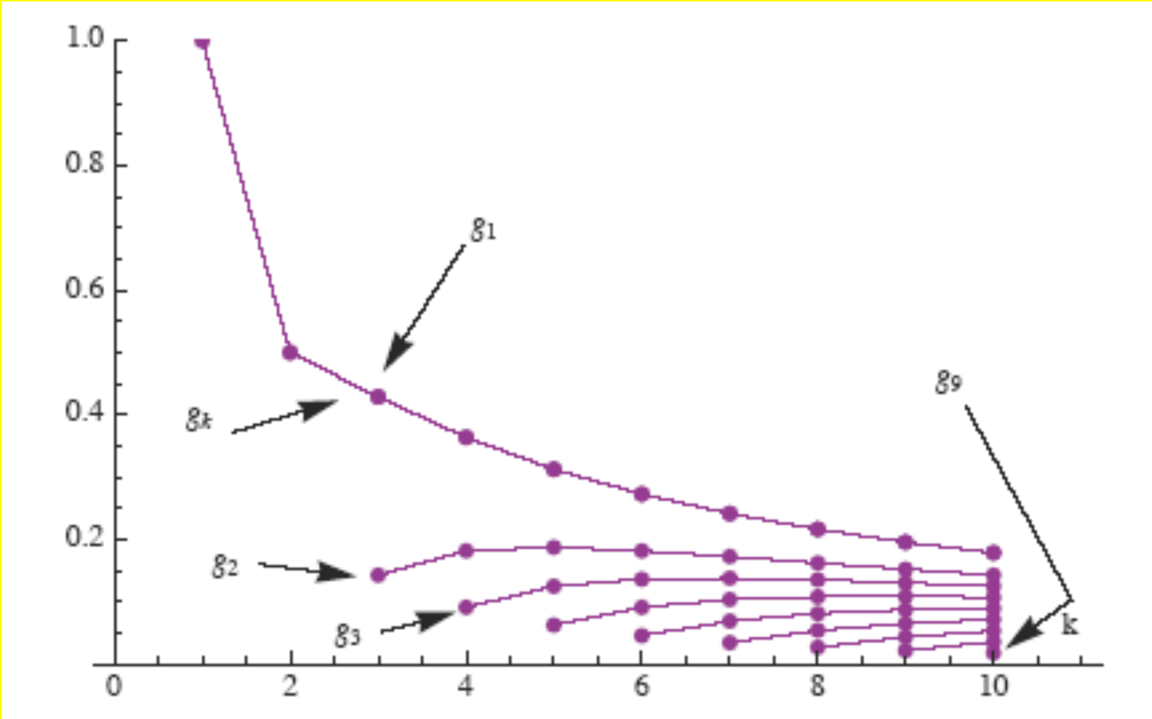


Figure 1: The various $g_{r,k}$ when $\delta = 2$ and $\mu = 1$ for $1 \le r \le k$ and $1 \le k \le 10$ from Table (1). The function $g_{k,k}$ represents the slot allocated to the last author of the list. It yields two third one third at $k = 2$. Every weight $g_{r,k}$ starts from $k = r + 1$.



$$\left.\begin{array}{l}\delta = 1 \\ \mu = 0\end{array}\right\} \text{-> 3/4, 1/4 at k = 2}$$



$$\left.\begin{array}{l}\delta = 0 \\ \mu = 1\end{array}\right\} \text{-> 1/2, 1/2 at k = 2}$$

One real case of an h-index

i	(k, r)	n^i
1	(2, 2)	187
2	(1, 1)	181
3	(3, 3)	179
4	(1, 1)	145
5	(1, 1)	145
6	(3, 2)	132
7	(1, 1)	132
8	(3, 2)	120
9	(2, 1)	104
10	(3, 1)	98
11	(2, 1)	94
12	(3, 3)	90
13	(3, 3)	81
14	(1, 1)	75
15	(2, 2)	72
16	(3, 3)	71
17	(3, 3)	68
18	(3, 3)	66
19	(2, 2)	63
20	(3, 3)	55
21	(1, 1)	51
22	(2, 2)	50
23	(2, 2)	48
24	(1, 1)	45
25	(1, 1)	43
26	(2, 1)	42
27	(1, 1)	39
28	(3, 2)	38
29	(2, 2)	38
30	(2, 2)	35
31	(2, 2)	35
32	(2, 2)	34
33	(6, 4)	33
34	(2, 1)	31

Calculating h_g

h = 33

9 x (1,1

4 x (1,2)

13 x (2,2)

1 x (1,3)

3 x (2,3)

9 x (3,3)

1 x (4,6)

→

(k, r)	$g_{n,k}^{2/3-1/3}$	$g_{n,k}^{3/4-1/4}$	$g_{n,k}^{1/2-1/2}$	$g_{n,k}^{arith}$
(1, 1)	1	1	1	1
(2, 1)	0.67	0.75	0.50	0.67
(2, 2)	0.33	0.25	0.50	0.33
(3, 1)	0.56	0.57	0.43	0.50
(3, 2)	0.11	0.14	0.14	0.33
(3, 3)	0.33	0.29	0.43	0.17
(6, 4)	0.08	0.09	0.09	0.14

Calculating h_g

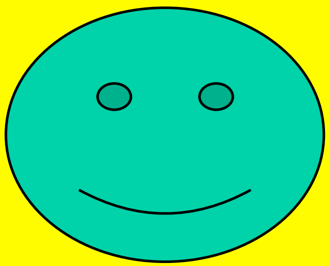
i	(k, r)	n^i	$g_{n,k}^{2/3-1/3} n^i$	$g_{n,k}^{3/4-1/4} n^i$	$g_{n,k}^{1/2-1/2} n^i$	$g_{n,k}^{arith} n^i$
1	(2, 2)	187	61.71	46.75	93.5	61.71
2	(1, 1)	181	181	181	181	181
3	(3, 3)	179	59.07	51.91	76.97	30.43
4	(1, 1)	145	145	145	145	145
5	(1, 1)	145	145	145	145	145
6	(3, 2)	132	14.52	18.48	18.48	43.56
7	(1, 1)	132	132	132	132	132
8	(3, 2)	120	13.20	16.80	16.80	39.60
9	(2, 1)	104	69.68	78.00	52.00	69.68
10	(3, 1)	98	54.98	55.86	42.14	49.00
11	(2, 1)	94	62.98	70.50	47.00	62.98
12	(3, 3)	90	29.70	26.10	38.70	15.30
13	(3, 3)	81	26.73	23.49	34.83	13.77
14	(1, 1)	75	75	75	75	75
15	(2, 2)	72	23.76	18	36	23.76
16	(3, 3)	71	23.43	20.59	30.53	12.07
17	(3, 3)	68	22.44	19.72	29.29	11.56
18	(3, 3)	66	21.78	19.14	28.38	11.22
19	(2, 2)	63	20.79	15.75	31.50	20.79
20	(3, 3)	55	18.15	15.95	23.63	9.35
21	(1, 1)	51	51	51	51	51
22	(2, 2)	50	16.5	12.5	25	16.5
23	(2, 2)	48	15.84	12	24	15.84
24	(1, 1)	45	45	45	45	45
25	(1, 1)	43	45	45	45	45
26	(2, 1)	42	28.14	31.50	21.00	28.14
27	(1, 1)	39	39	39	39	39
28	(3, 2)	38	4.18	5.32	5.32	12.54
29	(2, 2)	38	12.54	9.5	19	12.54
30	(2, 2)	35	11.55	8.75	17.5	11.55
31	(2, 2)	35	11.55	8.75	17.5	11.55
32	(2, 2)	34	11.22	8.50	17	11.22
33	(6, 4)	33	2.64	2.97	2.97	4.62
34	(2, 1)	31	20.77	23.25	15.50	20.77

$$h_{g_{2/3,1/3}} = 21$$

$$h_{g_{3/4,1/4}} = 19$$

$$h_{g_{1/2,1/2}} = 23$$

$$\leftarrow h = 33$$



Total number of articles

19.91, 18.94, 22.31

Instead of 40

With this scheme
having many citations
does make a difference
for every one of the
authors

k = 3 -> 0.56, 0.11, 0.33

20: -> 11.2, 2.2, 6.6

50: -> 28, 5.5, 16.5

k = 4 -> 0.46, 0.15, 0.08, 0.31

20: -> 9.2, 3, 1.6, 6.2

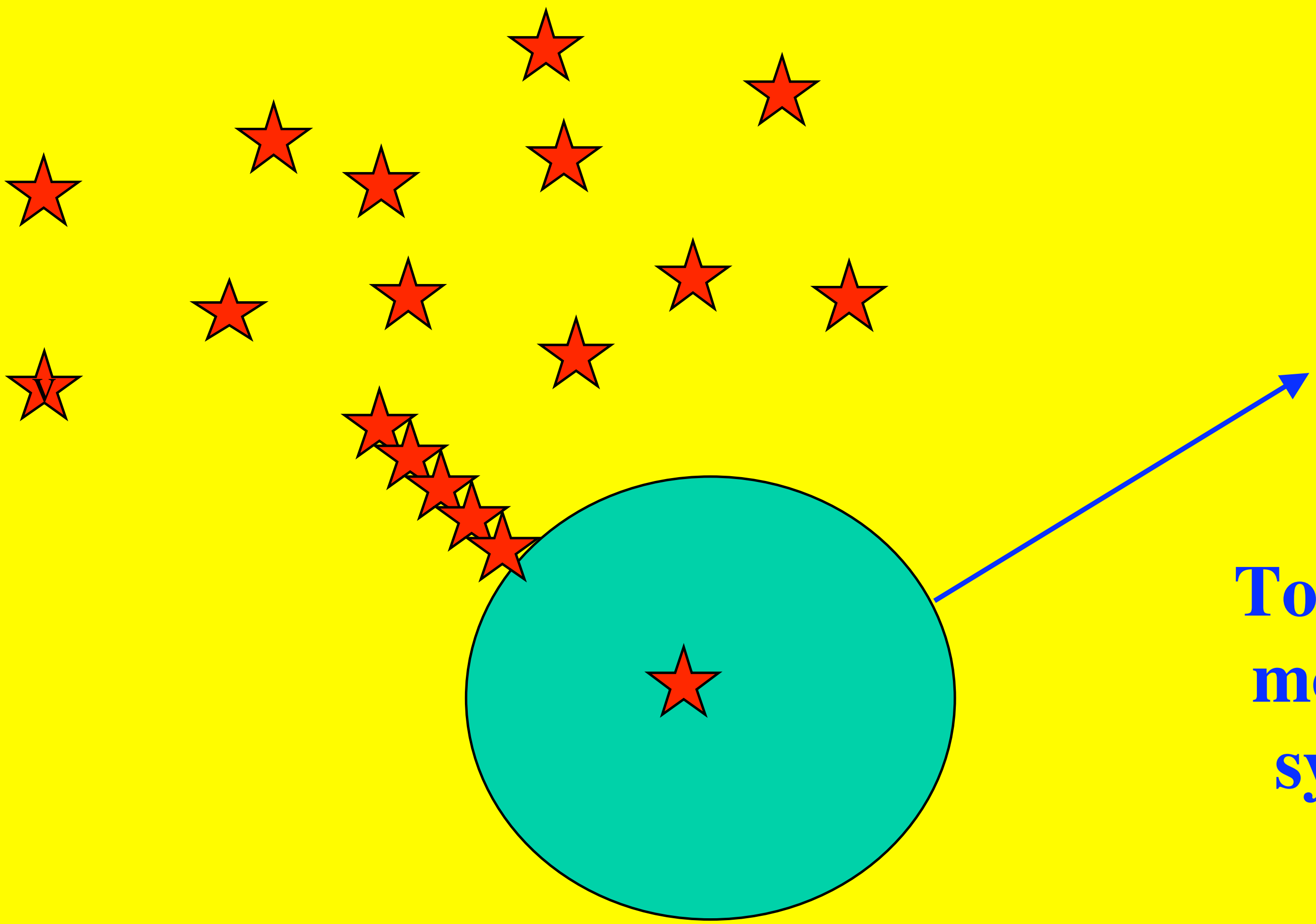
50: -> 23, 7.5, 4, 15.5

A reference:

S. Galam, Integrating multiple co authorship in the quantitative evaluation of individual scientific records

arXiv:1007.3708v1 [physics.soc-ph] 21 Jul 2010

To conclude:



…

Towards a more fair system?