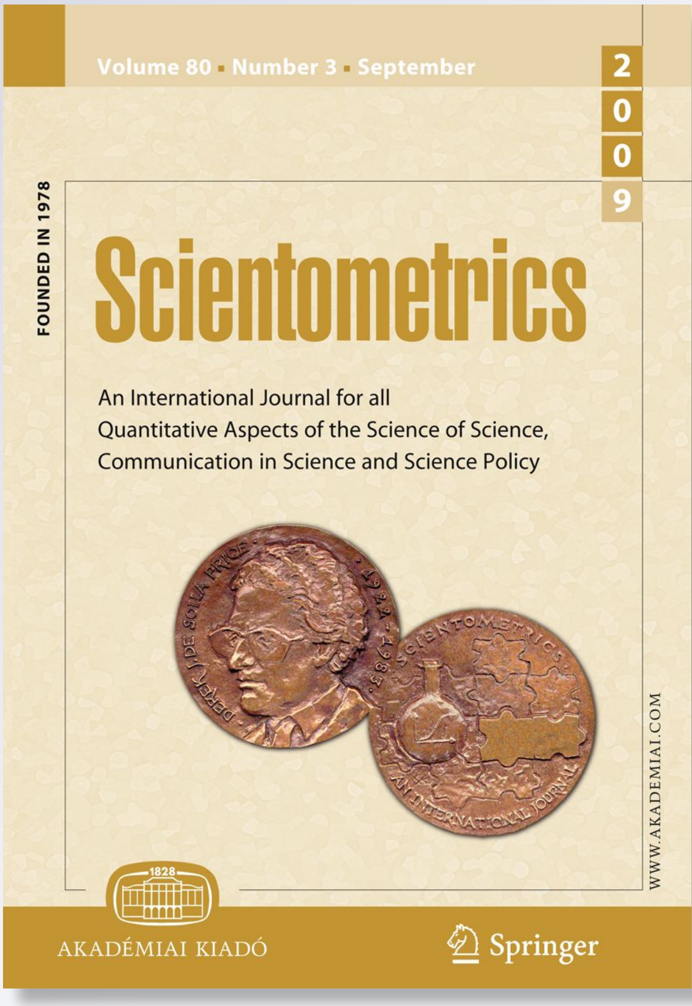
Research evaluation. Part II: gender effects of evaluation: are men more productive and more cited than women?

Kretschmer Hildrun, Pudovkin Alexander & Stegmann Johannes

Scientometrics
An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science Policy

ISSN 0138-9130

Scientometrics
DOI 10.1007/s11192-012-0658-0



Springer

Scientometrics
DOI 10.1007/s11192-012-0658-0

Research evaluation. Part II: gender effects of evaluation: are men more productive and more cited than women?

Kretschmer Hildrun · Pudovkin Alexander · Stegmann Johannes

Received: 23 January 2012

Abstract Productivity and citedness of the staff of a German medical research institution are analyzed. It was found in our previous study (Pudovkin et al.: Scientometrics, doi:10.1007/s11192-012-0659-z, 2012) that male scientists are more prolific and cited more often than female scientists. We explain in our present study one of the possible causes for obtaining this result with reference to Abramo et al. (Scientometrics 84(3): 821–833, 2009), who found in the small subgroups of star scientists a higher performance of male star scientists with respect to female star scientists; but in the remaining complementary subpopulations the performance gap between the two sexes is marginal. In agreement with Abramo et al. (2009), in our small subgroup of star scientists a higher performance of male star scientists with respect to female star scientists could be found. Contrasting, in the large complementary subgroup even a slightly higher performance of female scientists with respect to male scientists was identified. The last is even stronger expressed in favor of women than Abramo's result that the performance gap between the two sexes is truly marginal. In addition to Abramo et al. (2009), we already found in our previous study, special indexes characterizing the quality of papers (but not quantity) are not substantially different among sexes compared.

Keywords Gender · Evaluation · Productivity · Citations · Star scientists · Gender gap

Mathematics Subject Classification (2000) 62 · 68 · 94

K. Hildrun (✉)
Faculty of Business Administration/Business Computing, University of Applied Sciences, Bahnhofstrasse 15745, Wildau, Germany
e-mail: kretschmer.h@onlinehome.de

P. Alexander
Institute of Marine Biology, Far East Branch, Russian Academy of Sciences, Vladivostok, Russia 690041
e-mail: aipud@online.ru

S. Johannes
Berlin, Germany
e-mail: johannes.stegmann@alice.de

Published online: 17 February 2012

Introduction

Although participation of women in science is increasing in some countries and fields, the low participation rate of women in research activities has stimulated studies of the barriers faced by women in academia (Vetter 1981; Moore 1987; Leta and Lewison 2003). In recent years, increasing attention has been drawn to gender issues in academia (Butterwick 2005. The recent issue of *She Figures 2009* [published by the European Commission (2009)] indicates the urgency of the problem: averaged over all fields, despite the fact that more than half the European student population is female only 30% of European researchers and 18% of full professors are women. The more senior the position, the lower the presence of women (ETAN 2000; Hullmann 2001; Naldi and Parenti 2002, Naldi et al. 2004). Therefore, studying of the implications of evaluation parameters for the career development of women researchers is requested.

One of the goals of our previous study (Pudovkin et al. 2012) was related to gender differences obtaining the following results:

– Men are more productive than women.
– The parameters in which the differences between these groups are pronounced are the number of papers, sum of impact factor values of the journals in which the papers are published, cumulative number of cites to these papers, *H* index and some other indicators. These indexes are characterizing either productivity or both productivity and quality.
– While values of indexes characterizing the quality of papers (average citation rate per paper and similar indexes) are not substantially different among female and male scientists.

In our present study we intend to extend these findings from another point of view, i.e., we try to explain one of the possible causes for obtaining our former results with reference to Abramo et al. (2009). These authors have studied gender differences in research productivity of the entire population of research personnel in the scientific——technological disciplines of Italian university system. In particular the contribution of "star scientists" to overall sex differences in research productivity was analyzed. Abramo et al. (2009) could find out star, or "high-end", scientists play a preponderant role in determining higher performance among males. The term "star scientist" was coined by Zucker and Darby (1996). Abramo et al. (2009), identified the star (or high-end) scientists as those located in the top 10% of the rankings of scientific performance.

These authors have verified (cf. Abramo et al. 2009, page 821):

(a) There is a higher concentration of men among star scientists, and
(b) Additionally, there is a higher performance of male star scientists with respect to female star scientists in this subpopulation
(c) The performance gap between male and female star scientists is greater than for the complementary subpopulation
(d) In this complementary subpopulation the performance gap between the two sexes can be even seen as truly marginal.

Two evaluation parameters for scientific performance are used, both in our previous study (Pudovkin et al. 2012) and by Abramo et al. (2009):

– The number of papers (called "Output" by Abramo et al. 2009)
– Sum of impact factor values of the journals in which the papers are published (called "Scientific Strength" by Abramo et al. 2009)

On the other hand there are differences existing in data sizes, in performance indexes and in methodology:

– Whereas Abramo et al. (2009) have studied 29,036 researchers from different scientific disciplines and universities in Italy, the authors of the previous (Part I) and present (Part II) papers created a small data set only of research staff (30 female and 32 male scientists) of one institution [Deutsches Rheuma-Forschungszentrum (DRFZ)].
– Whereas the study by Abramo et al. (2009) is based on indexes characterizing either productivity or both productivity and quality of this productivity, the authors of the previous and present papers (Part I and II) have additionally used:

  *indexes characterizing citations
  *indexes characterizing only quality but not quantity

– In our previous and present studies the size of the differences between the compared groups are estimated by the difference index (DI). DI $= (x_1 - x_2)/\mathrm{SD}_{1\&2}$, x_1 and x_2 being the means in the compared groups 1 and 2, $\mathrm{SD}_{1\&2}$ being the averaged standard deviation. Statistical significance of the differences was estimated by the Student *t*-test. Abramo et al. (2009) have applied other statistical tests.

Our present study has five goals:

– To deliver similar verifications for the productivity patterns as done by Abramo et al. (2009) (cf. a, b, c, d above) although different sample sizes and methods are used. This is a kind of a "double" proof of the original approach by Abramo et al. (2009).
– Moreover, to deliver similar verifications for citation patterns and other indexes characterizing either productivity or both productivity and quality.
– Contrasting we will show there are minimum gender differences existing regarding values of indexes characterizing the quality of papers only. That means, the special role of the star or (high-end) scientists is rather diminished in this case.
– To deliver methods for visualization and comparison of the contrasting profiles of gender distributions in dependence on

  *Indexes of category 1: indexes characterizing either productivity or both productivity and quality
  *Indexes of category 2: indexes characterizing quality of publications independent on quantity

– To discuss the question: "Are men more productive and more cited than women?" on the basis of the new empirical results.

Data

The authors created a small data set of publications produced by the DRFZ in 2004–2008 (cf. our previous paper in this issue. Pudovkin et al. 2012). They extracted from the Web of Science (Thomson/Reuters [http://science.thomsonreuters.com/isi/]) all the publications by the staff of this institution and citation numbers to these publications. Consulting the web-site of this institution we identified the authors, their position and gender. There are 313 papers in the data base, authored and co-authored by 66 scientists of the DRFZ. The 313 papers were published in 96 journals, domestic and international. We identified 30 female and 32 scientists.

The authors performed an exercise in using different performance indexes (cf. Table 1 in our previous paper published in this issue. Pudovkin et al. 2012) in an attempt to see which of them are more informative in telling the difference in performance of male and female researchers. Two indexes, *e*-index (Zhang 2009) and *g*-index (after Egghe 2006) are added for the present study but the former two indexes ASI99 and ASI95 are not included based on the specialties of these indexes.

For comparison and visualization the two contrasting profiles of gender distributions we have classified these indexes (cf. Table 1 in Part I) into two categories:

Category 1: Indexes characterizing either productivity or both productivity and quality (NumP, sumIF, sumC, sumC/y, *H*-index, sumPRI, sumPRI × IF, G, E, ASI50, ASI75)
Category 2: Indexes characterizing quality of publications independent on quantity (avIF, avC, avPRI, avPRI × IF, av%75, avC/y, avNumAuth)

Furthermore, a new table is produced (Table 1 in the present study) for characterizing each index as follows:

(Remarks: the cumulative frequencies are calculated from down (corresponding to V_{max}) to up (corresponding to V_{min}), i.e., in the opposite direction as usual.)

– First column: index values (*V*)
– Second column: frequencies of female scientists (fe)

Table 1 Number of papers (NumP)

V	fe	CF	ma	CM	tot	CT	PF	PM	PT	COF	COM
1	7	30	7	32	14	62	100.00	100.00	100.00	1.00	1.00
2	3	23	3	25	6	48	76.67	78.13	77.42	0.99	1.01
3	2	20	3	22	5	42	66.67	68.75	67.74	0.98	1.01
4	1	18	4	19	5	37	60.00	59.38	59.68	1.01	0.99
5	4	17	0	15	4	32	56.67	46.88	51.61	1.10	0.91
6	3	13	1	15	4	28	43.33	46.88	45.16	0.96	1.04
7	2	10	1	14	3	24	33.33	43.75	38.71	0.86	1.13
8	1	8	1	13	2	21	26.67	40.63	33.87	0.79	1.20
9	1	7	0	12	1	19	23.33	37.50	30.65	0.76	1.22
10	1	6	0	12	1	18	20.00	37.50	29.03	0.69	1.29
12	1	5	0	12	1	17	16.67	37.50	27.42	0.61	1.37
14	1	4	2	12	3	16	13.33	37.50	25.81	0.52	1.45
16	1	3	0	10	1	13	10.00	31.25	20.97	0.48	1.49
17	0	2	1	10	1	12	6.67	31.25	19.35	0.34	1.61
20	1	2	0	9	1	11	6.67	28.13	17.74	0.38	1.59
22	0	1	1	9	1	10	3.33	28.13	16.13	0.21	1.74
24	0	1	1	8	1	9	3.33	25.00	14.52	0.23	1.72
27	0	1	1	7	1	8	3.33	21.88	12.90	0.26	1.70
28	1	1	1	6	2	7	3.33	18.75	11.29	0.30	1.66
40	0		1	5	1	5		15.63	8.06	0.00	1.94
41	0		1	4	1	4		12.50	6.45	0.00	1.94
42	0		1	3	1	3		9.38	4.84	0.00	1.94
52	0		1	2	1	2		6.25	3.23	0.00	1.94
64	0		1	1	1	1		3.13	1.61	0.00	1.94

- Third column: cumulative frequencies of females (CF)
- Fourth column: frequencies of males (ma)
- Fifth column: cumulative frequencies of males (CM)
- Sixth column: frequencies of scientists in total (tot)
- Seventh column: cumulative frequencies of scientists in total (CT)
- Eighth column: Female% $= 100 \times CF/CF_{max} = PF$
- Ninth column: Male% $= 100 \times CM/CM_{max} = PM$
- Tenth column: Scientists% $= 100 \times CT/CT_{max} = PT$
- Eleventh column: Concentration of females $= PF/PT = COF$
- Twelfth column: Concentration of males $= PM/PT = COM$

PF, PM and PT are called as relative cumulative frequencies.

In analogy to Abramo et al. (2009), the concentration of females (COF) is the ratio of the percentage of females in the subgroup of scientists with higher indicator values (100*CF/CT), to the percentage of females in the whole population ($100 \times CF_{max}/CT_{max}$).

The COM is the ratio of the percentage of males in the subgroup of scientists with higher indicator values ($100 \times CM/CT$), to the percentage of males in the population ($100 \times CM_{max}/CT_{max}$), i.e., COM = PM/PT and COF = PF/PT

Methods

Three methods are applied:

- Classification of the population (staff) into a subgroup of "high-end" scientists and the complementary subgroup
- Visualization and comparison of the contrasting profiles of gender distributions in dependence on category 1 or 2 indexes
- Statistical tests of gender differences

Classification of the population (staff) into a subgroup of "high-end" scientists and the complementary subgroup

As mentioned above, Abramo et al. (2009) identified the star (or high-end) scientists as those located in the top 10% of the rankings of scientific performance. However, because of our small sample size (30 female and 32 male scientists) it is not useful to identify the subgroup of high-end scientists as those located in the top 10% (i.e., 6–7 scientists only).

Inspired by the Pareto principle (known as the 80/20 rule: for many events, *roughly* 80% of the effects are coming from 20% of the causes), we propose in our case, to divide the whole population (62 scientists) per performance index slightly modified:

- into a subgroup of scientists located in the top about 25% of the rankings of scientific performance (called high-end scientists in our present paper)
- and in the remaining complementary subgroup (about 75% of the population).

Additionally, a method of visualization of gender distributions in the population is applied to identify if there are pronounced visible differences between men and women located in the field of high-end scientists. If this is the case, the population can be divided as mentioned above.

Visualization and comparison of the contrasting profiles of gender distributions in dependence on category 1 or 2 indexes

We will visualize the contrasting profiles of gender distributions in dependence on the different categories: category 1 and 2 with help of the SYSTAT graphic program:

*per index, per category
*per category

Statistical tests of gender differences

First

We will check in analogy to Abramo et al. (2009) if there is a higher concentration of men among high-end scientists using the concentration indexes COF and COM.

We will study separately:

*Category 1 indexes characterizing either productivity or both productivity and quality
*Category 2 indexes characterizing quality of publications independently on quantity

Second

We will test some verified assumptions by Abramo et al. (2009) but with other methods:

– If there is a higher performance of male high-end scientists with respect to female high-end scientists and
– If the performance gap between male and female star scientists is greater than for the complementary subpopulation
– If in this complementary subpopulation the performance gap between the two sexes can be even seen as truly marginal

In our previous study (Pudovkin et al. 2012) the size of the differences between the compared groups were estimated by the DI. DI $= (x_1 - x_2)/\mathrm{SD}_{1\&2}$, x_1 and x_2 being the means in the compared groups 1 and 2, $\mathrm{SD}_{1\&2}$ being the averaged standard deviation. Statistical significance of the differences was estimated by the Student *t*-test.

We are using the same method here for comparing male and female high-end scientists and comparing male and female scientists in the complementary subpopulation.

We will study separately again:

*Category 1 indexes characterizing either productivity or both productivity and quality
*Category 2 indexes characterizing quality of publications independently on quantity

Results

Visualization of some characteristic differences of gender distributions using category 1 or 2 indexes

Per index, per category

Depending on indexes, two examples are selected for presentation, one for category 1 (Fig. 1, left column), and another for category 2 (Fig. 1, right column). First row:

distributions of the relative cumulative frequencies PM (Male%) and PF (Female%) in dependence on the index values V.

Second row: distributions of the concentration measures COM and COF in dependence on the relative cumulative frequency PT (Scientists%).

Example for category 1, index NumP (Number of papers), cf. Table 1: Whereas the relative male cumulative frequency (PM) regarding the highest index values (Distance from $V = \text{NumP} = 40$ up to $V = \text{NumP} = 64$) is containing 15.63% of the male scientists the relative cumulative female frequency (PF) is equal to zero, i.e., the distributions of male and female cumulative frequencies are different from each other, cf. Fig. 1, first row (left pattern: the cumulative frequencies are counted from the right side of the abscissa to

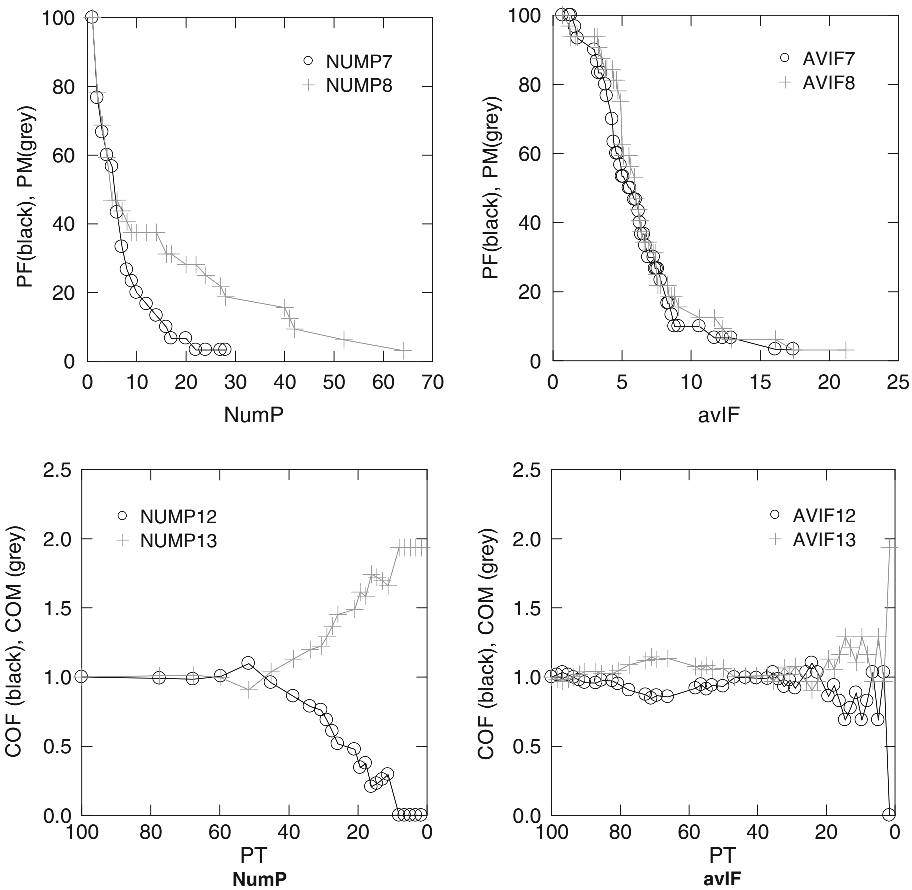


Fig. 1 Characteristic differences of gender distributions using category 1 (left column) or category 2 (right column) indexes. *First row* distributions of the relative male cumulative frequency (PM in *grey*: *plus*) and the relative female cumulative frequency (PF in *black*: *circle*) in dependence on the indexes. The left pattern shows the distributions in dependence on NumP (category 1 index) and the right the distributions in dependence on avIF (category 2 index) *Second row* distributions of the male and female concentrations (COM in *grey*: *plus* and COF in *black*: *circle*) in dependence on the relative cumulative frequency PT (Scientists%). The left pattern shows the distributions based on the index NumP (category 1 index) and the right the distributions in dependence on the index avIF (category 2 index)

the left). The relative cumulative frequency distributions (PF dots in black: ○ and PM dots in grey: +) are presented in dependence on the index' values $V = \text{NumP}$.

Example for category 2, index avIF [average journal IF (per paper, per author)]: In opposite to the category I indexes, using category II indexes the distributions for males and females are rather the same, cf. an example in Fig. 1, first row (right pattern) with the index avIF.

Concluding, the different distributions between males and females in category 1 indexes are rather diminished in the category 2 indexes

In Fig. 1, second row, the distributions of the male and female concentrations (COM and COF) are shown in dependence on the relative cumulative frequency PT (Scientists%). The left pattern shows the distributions based on the index NumP (category 1 index) and the right the distributions in dependence on the index avIF (category 2 index).

Concluding again, the different distributions between males and females in category 1 indexes are rather diminished in the category 2 indexes

Whereas in Fig. 1, first row, the distributions of PM and PF are presented in dependence on the values of a special index (for example, NumP or avIF), an option for better comparison of these distributions per category in total is given by replacement of this special abscissa by the *unique abscissa* PT as in Fig. 1, second row.

Per category

Four profiles of gender distributions in dependence on scientist's% (PT) will be established by overlay of all of the corresponding graphs into a single frame for:

*The distributions of the relative male cumulative frequency (PM in grey: +) and the relative female cumulative frequency (PF in black: ○), category 1 indexes (Fig. 2, first row, first column)
*The distributions of the relative male cumulative frequency (PM in grey: +) and the relative female cumulative frequency (PF in black: ○), category 2 indexes (Fig. 2, first row, second column)
*The distributions of the male and female concentrations (COM in grey: + and COF in black: ○), category 1 indexes (Fig. 2, second row, first column)
*The distributions of the male and female concentrations (COM in grey: + and COF in black: ○), category 2 indexes (Fig. 2, second row, second column)

This method of visualization of gender distributions in the population has confirmed there are pronounced visible differences between men and women, especially located in the field of high-end scientists (PT up to 25%). Thus, the population can be divided for further studies into a subgroup of the 25% high-end scientists and the complementary subgroup of the remaining 75% of the staff.

We could identify the profiles of the category 1 indexes (the two graphs, first column) are very different from the profiles of the category 2 indexes (the two graphs, second column), especially pronounced in the field with PT up to 25% (high-end scientists):

– Category 1 (profiles, first column): male scientists are overrepresented in the field with the highest indicator values, i.e. there are visible gender differences of distributions

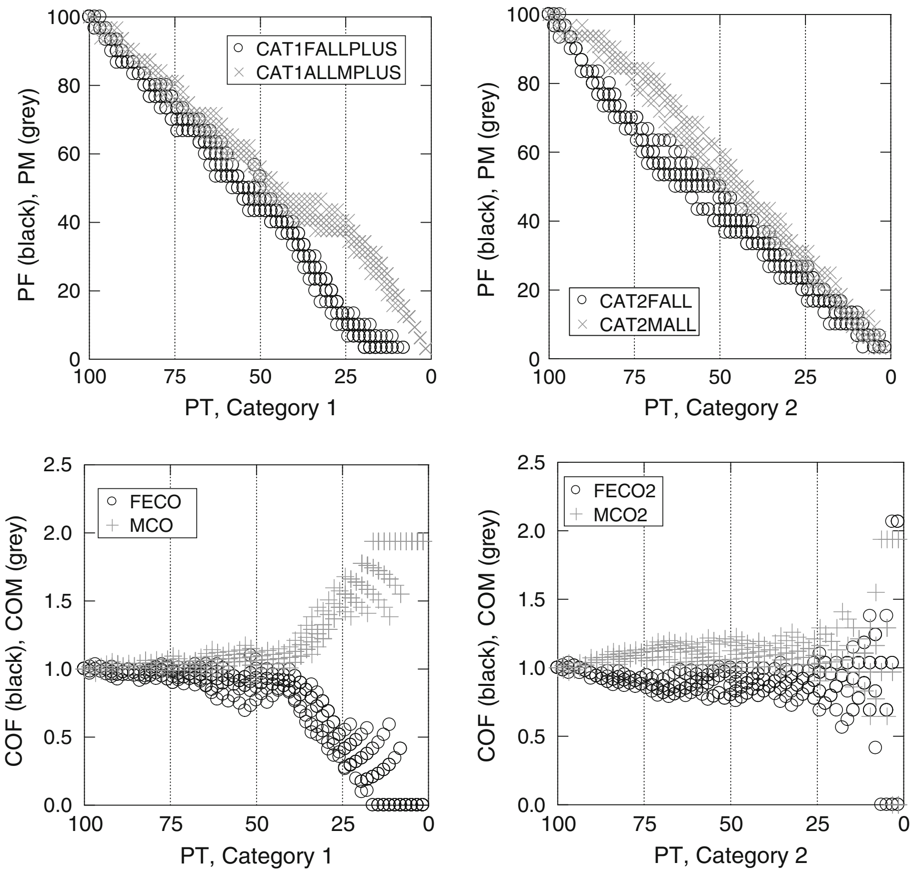


Fig. 2 Four profiles of gender distributions in dependence on scientist's% (PT). The profiles on the first column are related to the category 1 indexes and the profiles on the second column to the category 2 indexes. The *first row profiles* show distributions of the relative male cumulative frequencies (PM in *grey*: *plus*) and the relative female cumulative frequencies (PF in *black*: *circle*). The *second row profiles* show distributions of the concentrations (COM in grey: *plus* and COF in *black*: *circle*)

both of the relative cumulative frequencies (PM and PF) and of the concentration measures (COM and COF) in favour of men.

– Category 2 (profiles, second column): the relative cumulative frequency distribution of female scientists (PF) is rather equal to the male cumulative frequency distribution (PM) up to PT = 25% and there are similar distributions of the concentration measures COF and COM, i.e., the gender differences of distributions are rather diminished compared with the indicators in category 1.

Concluding, the method of visualization of gender distributions in the population has shown for the category 1 indexes, there are pronouncedly visible differences between men and women located in the field of high-end scientists (25% of the staff). Thus, the population can be divided for the following studies as mentioned above.

However, this kind of visible differences cannot be found in category 2 indexes. There is the question for further studies if this result is a specialty of the studied small German medical research institution. Nevertheless, for comparison category 1 indexes with category 2 indexes: The same method of division the population (25/75%) is applied.

Statistical tests of gender differences

In correspondence with Abramo et al. (2009) we have verified in our study, there is a higher concentration of men among high-end scientists (cf. Tables 2 and 3) with exception of the average journal impact factor (avIF, Table 3).

But the mean COM is higher for category 1 indexes ($COM_{mean} = 1.457$ in relation to women $COF_{mean} = 0.512$) than for category 2 indexes ($COM_{mean} = 1.107$ in relation to women $COF_{mean} = 0.886$).

Tables 4, 5 give the average values for the groups "Male"—M and "Female"—F and the DI. These groups with the DI (Statistical significance of the differences was estimated by the Student *t*-test) can be found under:

- Population (Staff: these values were already found and published in our previous paper, Part I in this issue)
- High-end subgroup
- Complementary subgroup

Table 2 Category 1 indexes for 62 scientists of the DRFZ

Parameter	M/sample size	F/sample size	COM	COF
NumP	12	4	1.453	0.517
sumIF	13	3	1.574	0.388
sumC	12	4	1.453	0.517
H-index	12	3	1.55	0.413
sumC/y	12	4	1.453	0.517
ASI50	12	5	1.367	0.608
ASI75	12	5	1.367	0.608
sumPRI	12	4	1.453	0.517
sumPRI × IF	12	4	1.453	0.517
G	12	4	1.453	0.517
E	12	4	1.453	0.517
Mean			**1.457**	**0.512**

Sample size of males in the population 32, sample size of females in the population 30

F woman, *M* man, parameter (meaning, cf. Table 1) *COM* concentration of males in the high-end subgroup, *COF* concentration of females in the high-end subgroup

Table 3 Category 2 indexes for 62 scientists of the DRFZ

Parameter	M/sample size	F/sample size	COM	COF
avC	9	7	1.09	0.904
avIF	8	8	0.969	1.033
av%75	10	6	1.211	0.775
avPRI	9	7	1.09	0.904
avPRI × IF	10	6	1.211	0.775
avC/y	9	7	1.09	0.904
avNumAuth	9	7	1.09	0.904
Mean			**1.107**	**0.886**

Sample size of males in the population 32, sample size of females in the population 30

F woman, *M* man, parameters (meaning, cf. Table 1) *COM* concentration of males in the high-end subgroup, *COF* concentration of females in the high-end subgroup

Table 4 gives the values for the category 1 indexes. One can see that the largest average contrast is observed between "Males" and "Females" in the high-end subgroup (DI = 0.99) and the smallest in the complementary subgroup (DI = −0.34). Moreover, the last shows even a slightly (not significant) higher performance of female scientists with respect to male scientists (DI is equal to −0.34).

The average contrast between "Males" and "Females" in the whole population (staff) is smaller (DI = 0.61) than the contrast in the high-end subgroup but very much larger than in the complementary subgroup.

The most expressed differences between "Males" and "Females" are seen in the high-end subgroup, in the following characteristics: sumC/y (DI = 1.39, $p < 0.001$) and sumPRI × IF (DI = 1.23, $p < 0.001$).

The parameters characterizing the quality of papers (category 2, Table 5) poorly differentiate the compared groups of "Males" and "Females" in the population (average

Table 4 Category 1 indexes for 62 scientists of the DRFZ

Parameter	Population (staff)			High-end subgroup			Complementary subgroup		
	M	F	Diff. index	M	F	Diff. index	M	F	Diff. index
Sample size	32	30	32 & 30						
Sample size				12	4	12 & 4	20	26	20 & 26
NumP	13.875	6.40	0.57**	32.08	19.50	0.84**	2.95	4.38	−0.51*
Sample size				13	3	13 & 3	19	27	19 & 27
sumIF	89.609	32.80	0.73**	191.82	87.78	1.10***	19.68	26.69	−0.35
Sample size				12	4	12 & 4	20	26	20 & 26
sumC	352.72	116.70	0.67***	856.25	378.00	1.11**	50.60	76.50	−0.38
Sample size				12	3	12 & 3	20	27	20 & 27
H-index	7.438	3.90	0.61**	15.83	10.67	1.00**	2.40	3.15	−0.47
Sample size				12	4	12 & 4	20	26	20 & 26
sumC/y	87.140	28.10	0.71**	208.91	86.13	1.39****	14.08	19.17	−0.31
Sample size				12	5	5 & 12	20	25	25 & 20
ASI50	8.906	3.80	0.55**	21.08	12.80	0.74*	1.60	2.00	−0.23
Sample size				12	5	5 & 12	20	25	25 & 20
ASI75	6.281	2.47	0.51**	15.17	8.80	0.61	0.95	1.20	−0.18
Sample size				12	4	12 & 4	20	26	20 & 26
sumPRI	875.2	370.8	0.57**	2058.33	1263.0	0.74*	165.4	233.6	−0.39
Sample size				12	4	12 & 4	20	26	20 & 26
sumPRI × IF	5378.0	1816.0	0.73***	12548.2	5451.5	1.23****	1075.0	1256.6	−0.18
Sample size				12	4	12 & 4	20	26	20 & 26
G	11.34	5.63	0.59**	25.33	16.00	1.02**	2.95	4.04	−0.44
Sample size				12	4	12 & 4	20	26	20 & 26
E	11.29	7.69	0.50*	21.37	16.62	1.15**	5.24	6.31	−0.35
Mean			**0.61**			**0.99**			**−0.34**

Line titles sample size—self-evident, averages for the groups NumP, sumIF, etc. (meaning, cf. Table 1)—self-evident DI: $(x_1 - x_2)/\mathrm{SD}$. Significance is indicated by asterisks * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

F woman, *M* man, *Diff. index* difference index

Table 5 Category 2 indexes for 62 scientists of the DRFZ

	Population (staff)			High-end subgroup			Complementary subgroup		
Parameter	M	F	Diff. index	M	F	Diff. index	M	F	Diff. index
Sample size	32	30	32 & 30						
Sample size				9	7	9 & 7	23	23	23 & 23
avC	19.972	17.203	0.19	37.88	39.66	−0.15	12.97	10.37	0.38
Sample size				8	8	8 & 8	24	22	24 & 22
avIF	6.706	5.983	0.20	11.88	10.14	0.44	4.98	4.47	0.30
Sample size				10	6	10 & 6	22	24	24 & 22
av%75	0.361	0.307	0.16	0.81	0.84	−0.19	0.16	0.17	−0.11
Sample size				9	7	9 & 7	23	23	23 & 23
avPRI	58.559	51.567	0.32	83.52	79.52	0.56	48.79	43.16	0.31
Sample size				10	6	10 & 6	22	24	24 & 22
avPRI × IF	354.686	295.978	0.31	545.21	616.84	−0.51	268.08	215.76	0.43
Sample size				9	7	9 & 7	23	23	23 & 23
avC/y	5.442	4.257	0.32	10.18	9.26	0.54	3.59	2.74	0.50*
Sample size				9	7	9 & 7	23	23	23 & 23
avNumAuth	9.128	8.302	0.22	13.25	12.93	0.13	7.52	6.89	0.36
Mean			**0.25**			**0.12**			**0.31**

Line titles sample size—self-evident; averages for the groups: avC, avIF, etc. (meaning, cf. Table 1)—self-evident DI: $(x_1 - x_2)$/SD. Significance is indicated by asterisks * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

F woman, *M* man, Diff. index difference index

DI = 0.25), in the high-end subgroup (average DI = 0.12) and in the complementary subgroup (average DI = 0.31).

Conclusions

Productivity and citedness of the staff of a German medical research institution are analyzed with 30 females and 32 males. This is a considerably higher percentage than mentioned above in the recent issue of *She Figures 2009*. In average only 30% of European researchers are women.

It was found that in the whole population male scientists are more prolific and cited more often than female scientists. But with reference to Abramo et al. (2009) we have verified star or "high-end" scientists play a preponderant role in determining this higher performance among males in relation to the category 1 indexes:

1. There is a concentration of men among high-end scientists (top 25% of the staff).
2. Additionally; the largest contrast of performance is observed between "Males" and "Females" in this star or high-end subgroup.
3. The smallest—but reversed contrast of performance between "Males" and "Females" was found in the complementary subgroup (75% of the staff). The last shows a slightly (not significant) higher performance of female scientists with respect to male scientists.

4. The average contrast in productivity and citedness between “Males” and ”Females” in the whole population (staff) is smaller than the contrast in the high-end subgroup but even inverted to the complementary subgroup.

There is a request for explanation why the concentration of women among high-end scientists is low. Age, child bearing, and glass ceiling are real phenomena we should take into consideration for women who are less productive.

However, unfortunately, there are practical difficulties existing rather all over the world. Going higher in the hierarchy, few women are found in science in general. On the one hand this could be influenced by performance indicators as in category 1 but additionally or basically this could be based on the role of women in the public society: Women get pregnant, bring up children. This leads to a deficiency of time for research work and subsequent fewer opportunities for a career. Men dominate the higher hierarchy positions. For increasing women’s production for example, a better system for the care of children has to be established in connection with other promotions of the family duties.

However, the differences between male and female scientists are significant in indexes related to the number of papers (category 1), while values of indexes characterizing the quality of papers (average citation rate per paper and similar indexes, category 2) are not substantially different among the sexes compared. This can be considered as a proof; women can deliver scientific results of the same quality as men but because of missing time there are difficulties to deliver the same quantity.

Last but not least, there is the question: Are men more productive and more cited than women?

Based on our empirical results we can confirm:

In the small subgroup of the top scientists men are more productive and more cited than women.

There is the question arising in what extent this statement obtained from the small subgroup can be truly transferred to the whole population although the larger complementary subgroup has given even a slightly reversed picture in favour of women?

Acknowledgement Part of this work by one of the authors (Kretschmer, H.) was supported by the 7th framework program by the European Commission, SIS-2010-1.3.3.1. Project full title “Academic Careers Understood through Measurement and Norms”, Project acronym ACUMEN.

References

Abramo, G., D’Angelo, C. A., & Caprasecca, A. (2009). The contribution of star scientists to overall sex differences in research productivity. *Scientometrics, 84*(3), 821–833.

Butterwick, S., & Dawson, J. (2005, January). Undone business: Examining the production of academic labour. *Womens Studies International Forum 28*(1):51–65.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics, 69*(1), 131–152.

ETAN (2000). Expert working group on women and science. Science policy in the European Union. Promotion excellence through mainstreaming gender equality. European Commission. Directorate—General for Research. Brusssels. http://www.cordis.lu/improving/women/documents.htm.

European Commission. (2009). *She Figures 2009. Statistics and indicators on gender equality in science*. Brussels: European Commission.

Hullmann, A. (2001). Bibliometric and patent indicators by gender: Is it feasible? http://www.cordis.lu/indicators/publications.htm.

Leta, J., & Lewison, G. (2003). The contribution of women in Brazilian science. A case study in astronomy, immunology and oceanography. *Scientometrics, 57*(3), 339–353.

Moore, K. M. (1987). Women’s access and opportunity in higher education: toward the twentyfirst century. *Comparative Education, 23*(1), 23–33.

Naldi, F., & Parenti, I.V. (2002, November). Biosoft sas: Bibliometric and patent indicators by gender: Two feasibility studies. Volume 1: Statistics, Volume 2: Methodology. http://www.cordis.lu/indicators/publications.htm.

Naldi, F., Luzi, D., Valente, A., & Parenti, I. V. (2004). Scientific and technological performance by gender. In H. F. Moed, et al. (Eds.), *Handbook of quantitative science and technology research* (pp. 299–314). Dordrecht: Kluwer Academic Publishers.

Pudovkin, A., Kretschmer, H., Stegmann, J., & Garfield, E. (2012). Research evaluation. Part I: Productivity and citedness of a German medical research institution. *Scientometrics*. doi:10.1007/s11192-012-0659-z.

Vetter, B. M. (1981). Women scientists and engineers: trends in participation. *Science, 214*, 1313–1321.

Zhang, C-T. (2009). The e-index, complementing the h-index for excess citations. *PlosOne*, *4*(5):e5429.

Zucker, L. G., & Darby, M. R. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Science of USA, 93*, 12709–12716.