Brazilian Authors´ Scientific Performance: Does Gender Matter?

Pablo Diniz Batista[1] and Jacqueline Leta[2]

[1] *batista@cbpf.br*
Brazilian Center for Physics Research, Rua Dr. Xavier Sigaud, 150, CEP: 22290-180, Rio de Janeiro (Brazil)

[2] *jleta@bioqmed.ufrj.br*
Federal University of Rio de Janeiro, Av. Brigadeiro Trompowisky s/ nº, Prédio do CCS, Bloco B – sala 39, CEP 21941-590, Rio de Janeiro (Brazil)

Abstract
The number of women pursuing careers in science has recently increased remarkably. Nevertheless, they still have few chances to be successful in their careers. Low scientific performance is one of the most frequent reasons behind women's failure in science. The present study aims at presenting evidences upon men and women´s scientific performances. Two different databases were cross-examined to provide trends on publication and visibility (estimated by h-index) of 18,993 Brazilian authors. The results showed that the chances to be within the select group of the most productive Brazilian authors is much higher for male authors. As for those authors with 50 or less publications (90% of the sample), differences in probabilistic were not observed. Similar results were found for the h-index analysis. Regarding the relation between research field and scientific performance, we have found that research field plays an important role for male authors in terms of number of publications and h-index. However, the research field seems to have no effect upon female scientific performance. We hope the empirical evidences presented in this paper may contribute to the debate on gender differences within other scientific communities.

Introduction
Historically, science has always been perceived as a male activity. Such view was first described by Mead & Metraux (1957), who carried out an empirical study with over 30.000 high school students from the USA. This pioneering study revealed a stereotypical image about science's main actors - scientists. Accordingly, scientists were viewed as elderly or middle-aged males, wearing white coats and glasses, and performing dangerous experiments.
After 50 years, the stereotypical image of scientists found in Mead & Metraux's study was perpetuated worldwide (Finson, 2002). Thus, such collective and social imaginary upon scientists portrays the current global science scenario, in which women are still underrepresented. Causes for this situation are complex and need deeper and detailed investigation. However, scientists´ lasting social representation may not be disregarded as a potential cause.
Fortunately, women are bypassing social and/or cultural barriers. Throughout the last decades, we have been watching a remarkable increase in the number of women pursuing careers in science. In a recent report, UNESCO (2007a) presents a global map of the male and female researchers worldwide. In 40% of the 103 studied countries, women researchers still represent less than one-third of all researchers, but in 17% of the countries, they reached half of them. Yet, differences do exist among countries and regions. In Latin America and the Caribbean, females represent 43% of the total researchers, exceeding the world average. This share is similar to the one found for South East Asia countries, but contrasts with South Asia and Africa, where only 17% or 30% of the researchers are women, respectively. Curiously, countries with scientific tradition are not among those with gender parity instead of some "young" developing countries such as Brazil, Argentina, Venezuela, the Former Yugoslav Republic of Macedonia, Lithuania, and most of the Central Asia countries.
These recent gains, however, come together with an unbalanced number of women in senior academic and leadership positions. This is a widespread trend that seems to have no relation with the level of the country's scientific tradition or economic and social development. In the

Proceedings of ISSI 2009, pages 343–353.
Edited by B. Larsen and J. Leta.

USA, the global science leader, 39 out of 100 women employed in the research universities were full professors in 1995, whereas this ratio was 59 out of 100 for men (NSF, 1999). As for countries with tradition in science such as France and Germany, women represent 14% and 5.9%, respectively, of all professors in the highest top position (Hermann, 2002; Costas, 2002). Among countries with no tradition in science, India's women represent only 7% of all faculties in four important scientific and technological institutes (Gupta & Sharma, 2002).

Such unbalanced picture is not restricted to academic positions. Women are also underrepresented in scientific boards (UNESCO, 2007b). In England and Italy, for example, women, in average, represent 27% and 12% of the total members, respectively. As for a country with gender parity in science, Leta (2002) presented the unbalanced number of females and males in the Brazilian Academy of Sciences; from a total of 353, women made up only 26 academics. A similar picture was presented by McGregot & Harding (1996) to some select advisory board membership. Among their examples, special attention was drawn to the Committee for the European Development of Science and Technology constituted by 30 members, in which only one is a woman.

Thus, if in the one hand the number of women has been increasing in science, in the other hand they still have few chances to be successful in their careers. This current scenario illustrates a phenomenon described and discussed during the 80's, the vertical segregation or crystal ceiling (Schienbinger, 2001). But what would be the reasons behind such segregation? Rossi´s (1965) classical work points to some social and cultural aspects, such as marriage and maternity priority as well as some typical female behaviors and biologic characteristics, which may influence the process of career ascent. There is no doubt that these aspects are intertwined and they may be direct or indirectly involved with the differences in scientific performances of males and females, what ultimately explain the rise (and success) in scientific career

Among the vast international literature on gender and differences in scientific performances, female scientists are usually found to be lesser productive than their male peers (some recent examples: Pripc, 2002; Fox, 2005; Mauleón&Bordons, 2006). Besides, controversial data has been found regarding both the productivity and visibility (or quality?) of females' and males´ publications, estimated through number of papers and citations. Nonetheless, it is well-known that these variables weight directly on individual evaluations and they have been used as proxy measures of scientific performance for decisions on promotion for both male and female researchers. Moreover, according to Fox (2005), "publication productivity operates as both cause and effect of status in science. Publication productivity reflects women's depressed rank and status, and partially accounts for it". As for Bourdieu's (2003) sociological view, the higher the scientific capital, measured for example as number of papers, the faster and larger is the prestige and progress in the academic career.

Taking into account the central role of scientific performance variables for the meritocracy system in science, the present study aims at presenting empirical evidences upon male and female researcher performances in terms of publications and citations in a large scale study. Different from most gender studies on productivity, we carried out a set of analysis with data retrieved from almost 19,000 Brazilian authors. The results showed that male and female authors´ performance differences are clearly observed within the select group of the most productive Brazilian authors. On the other hand, as for those authors with 50 or less publications (90% of the sample), differences were not observed. Similar results were found for the h-index analysis. Regarding the influence of the research field on scientific performance, our data indicated that research field plays an important role for male authors, yet has no effect upon female authors' scientific performance. We believe such evidences may contribute to the debate on gender differences within other scientific communities.

Method and procedures

One of the major difficulties to study male and female authors´ participation in the scientific literature is to identify their sex. In most countries, the first name reveals or indicates the sex of any individual. However, scientific paper authors are commonly named by their last names and initials; in a very few number of journals, authors are named by their last and first full names. Thus, due to this methodological obstacle, most of the studies on gender and productivity are focused on small scientific communities, whose names and sex may be retreived either from journals or from some available database.

In order to overcome such limitation, all the quantitative analysis presented in this paper was based on two databases: CNPq and Brazilian Science Indicators (BSI). Since data from these two databases are complementary, they were cross-examined in order to investigate gender influences on the scientific performance of Brazilian researchers.

CNPq subset database versus BSI database

CNPq database covers information of around 90% of all the human resources engaged in the Brazilian science, including graduate and undergraduate students, post docs, technicians, and researchers from all fields, with or without a permanent position in a research institution. The subset database used in the present work comprises information about academic and scientific performances as well as general information, such as the full name and the gender of 51,233 junior and senior Brazilian researchers with a PhD degree.

BSI database's main source of information is the Thomson Scientific ISI Web of Science database (ISI). Built in 2005 by Batista *et al* (Batista, Campiteli, Kinouch & Martinez, 2006), BSI encompasses information about 188,909 Brazilian bibliographical references (original articles add up 150,323) published in the 1960 – 2004 period. Total references contain 207,917 authors, including Brazilians and foreigners. It is not possible to estimate the exact number of Brazilian and foreign authors of each reference. However, for each publication there is at least one author with a Brazilian address.

Sample Characteristics

In the present study, as our main target was to identify Brazilian authors – female and male – who published papers in ISI´s database, we combined data of scientific performance catalogued in BSI´s database with those of Brazilian researchers' general information, catalogued in CNPq´s database..

By crossing both databases information, we had to overpass a critical obstacle - the authors' names format. CNPq´s database presents the full name of each Brazilian researcher, while BSI´s database presents the abbreviation of his/her name. Thus, we had to build an algorithm, starting by a list of abbreviated names for all the 51,233 Brazilian names listed in CNPq´s database. In this stage, the abbreviation of each full name considered the last name followed by the initials of the first and middle names (whenever it was presented). The next stage was to check the concurrence of each abbreviation within BSI´s list of authors. To reduce problems with homonyms, all of the researchers with identical abbreviated names were excluded.

Table 1: Number of Brazilian authors according to gender and research field compiled in CNPq-BSI´s database and CNPq´s database .

Research Field	CNPq-BSI database			CNPq database			(a)/(b)
	Total (a)	*Male*	*Female*	*Total (b)*	*Male*	*Female*	
Biological sc.	3,749	1,750	1,999	6,806	3,002	3,803	55,1%
Exact science	4,800	3,404	1,396	9,934	7,070	2,857	48,3%
Health Sciences	3,289	1,972	1,317	6,969	3,821	3,146	47,2%
Agriculture	2,705	1,945	760	6,013	4,162	1,851	45,0%
Engineering	2,501	2,004	497	6,639	5,283	1,355	37,7%
Humanities	1,048	517	531	7,510	3,313	4,194	14,0%
Social Sciences	643	450	193	5,017	3,206	1,809	12,8%
Language & arts	258	109	149	2,335	868	1,465	11,0%
TOTAL	**18,993**	**12,151**	**6,842**	**51,223**	**30,725**	**20,480**	**37,1%**

This name crossing process resulted in a total of 18,993 names found in both BSI and CNPq databases. This list composed a new database, CNPq-BSI´s database, with 12,151 male authors (64%) and 6.842 female authors (36%). As seen in Table 1, female authors´ fraction varies within the research fields. Table 1 also indicates that the crossing process was much more efficient in recovering authors from biological sciences. For this field, 3,749 out of the 6,806 (55,1 %) names listed in CNPq´s database were found in BSI´s database. Conversely, the process was much less efficient among "soft sciences" (Humanities, Social Sciences Language & Arts).

The efficiency of this process can also be noticed by comparing the number of publications compiled in BSI´s database authored by at least one name listed in CNPq´s database (gray colors) and the total number of publications (in white) (Figure 1).

We found that from the total of 183,972 publications compiled in BSI database, 115,087 publications were authored by at least one researcher listed in CNPq´s database. Regarding authors´ gender, we found 72,631 publications authored by a male researcher listed in CNPq´s database (open triangle); 21,402 publications have a female reseacher (open diamond) and 21,054 publications have both a male and a female researcher listed in CNPq´s database (inverted open triangle). We also found that the number of publications issued by Brazilian male researchers are increasing faster than the females'. Nevertheless, considering the size of our sample, we noticed that, in average, a Brazilian male researcher publishes 7.71 publications during the whole period (1960 – 2004) while the average for female reseachers is 6.20.

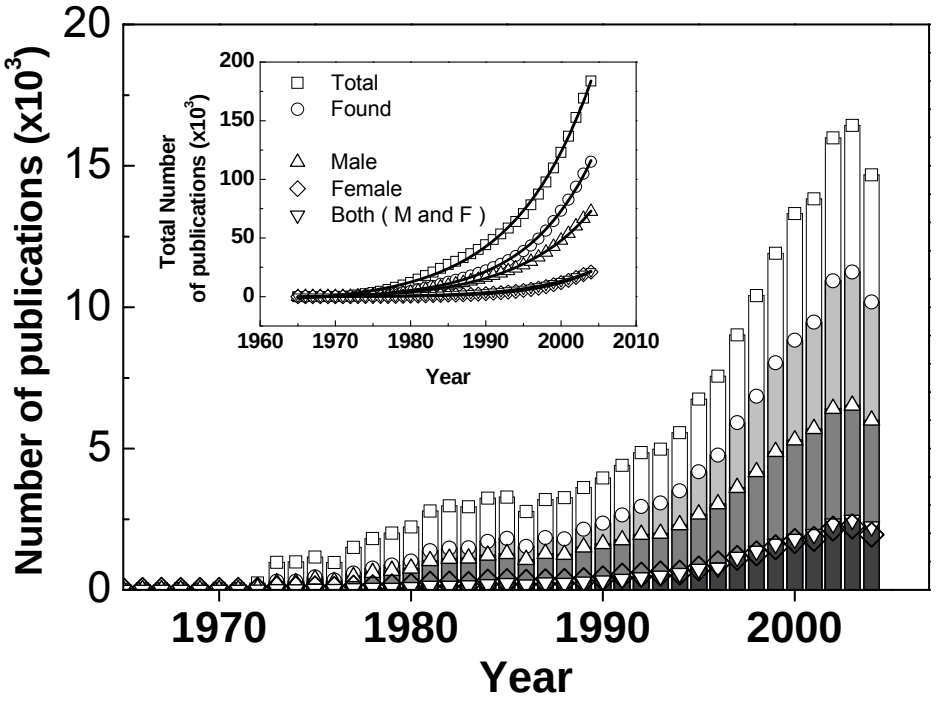

Figure 1: Annual distribution of Brazilian total publications in BSI´s database (□) and those with at least one author listed in CNPq´s subset database (○) - male authors (△), female authors (◇) and male and female authors (▽)

Inset: Cumulative number of publications for BSI database (□), CNPq-BSI database (○).

Results

In order to identify Brazilian researchers´ scientific performance, we focused the present study on two bibliometric indicators, namely the number of publications and h-index.

Gender variable behind scientific performance

Many studies have presented evidences that female researchers usually publish fewer papers than their male partners (examples: Zuckerman & Cole, 1975; Long, 1992; Symonds, Gemmell, Braisher3, Gorringe & Elgar, 2006). Such international trend seems to be partially true for the 18,993 Brazilian authors recorded in the CNPq-BSI database. In fact, since Brazilian authors are ordered by the number of publications they have published during the 1960 – 2004 period (Figure 2), we observed that male authors – open squares - are the majority among the first 100 most productive authors (from 10^0 to 10^2). Within this set of select authors, females are scarce (14%). Moreover, considering the 100 top male researchers and the top female researchers, we found they authored 15,900 and 8,063 publications, respectively.

The analysis of the complementary cumulative distribution (Figure 2 inset), in terms of number of publications, offers the probability to identify authors with more than "n" publications (Np). For instance, we noticed that the probability to find authors with $Np \geq 50$ (10% of all authors) is 4.1 % for male authors and 1.4 % for female authors. Thus, this analysis has also pointed to a different trend between male and female within the group of the most productive authors. Hence, the figure shows clearly that as large is the number of publications, higher is the probability to find a male author.

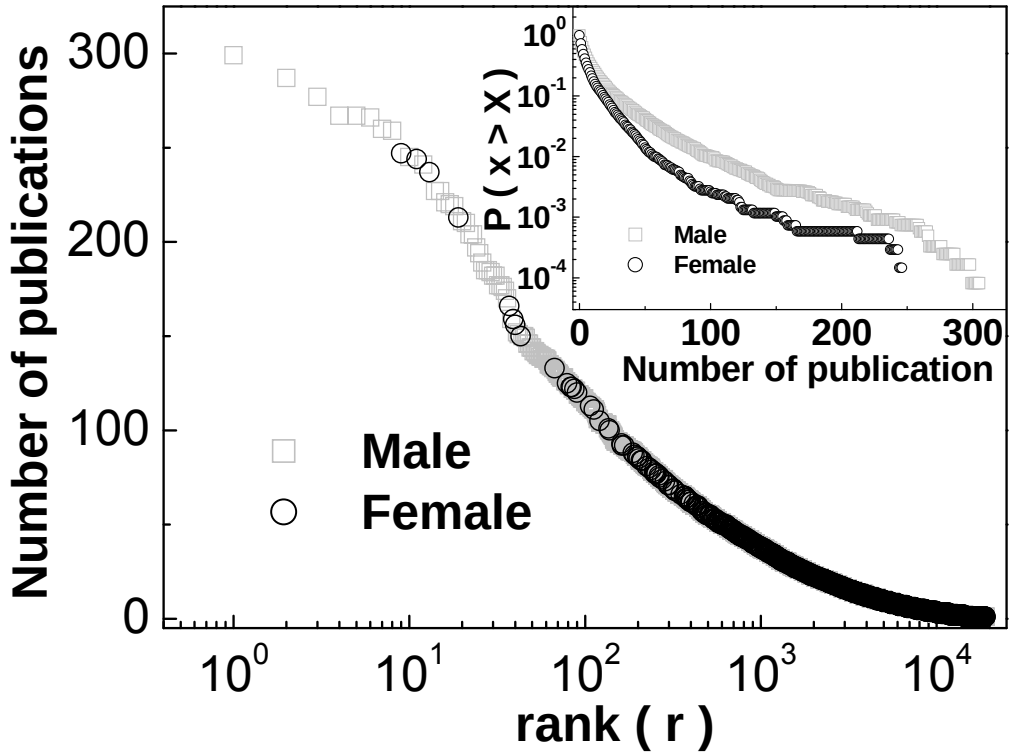

Figure 2: Ranking of Brazilian male and female authors according to the number of publications. Inset: Complementary cumulative distribution for Brazilian number of male and female authors´ papers.

Taking into consideration these trends, gender seems to emerge as a dependent variable related to the scientific performance among the select group of the 100 most productive authors as well as among the 10% Brazilian most productive researchers. However, we are not driven to a similar conclusion if we consider the majority of Brazilian authors, that is, those with 50 or less publications in the whole period (these authors totalize 90% of the whole sample). Among such largest set of authors, the probability of finding a male author is the same as finding a female author.

Similar findings were observed when we carried out an analysis upon citation data. Hence, as our goal was to get an individual scenario of our sample's performance on the number of papers and number of citations, we opted to present the h-index only. As coined by Hirsh (2005), the h-index considers both variables of an individual and so may allow individual comparisons. Despite criticism, h-index has been primarily used for evaluating individual performances (for example: Bornmann & Daniel, 2005; Mugnaini, Packer, & Meneghini, 2008). More recently, however, it has also been used for other evaluating purposes (Braun, Glanzel & Schubert, 2005).
In Figure 3, we present the distribution of Brazilian male and female authors in a decreasing h-index order. If we consider the first 100 authors with the highest h-index (from 10^0 to 10^2), we will find that most of them are male (84%). As it has previously been demonstrated, female authors are also scarce within this select group. Furthermore, considering only the first 100 male, as well as the first 100 female researchers with the highest value of h-index, we could find an average h-index of 23.34 for the former group and 14.87 for the latter.

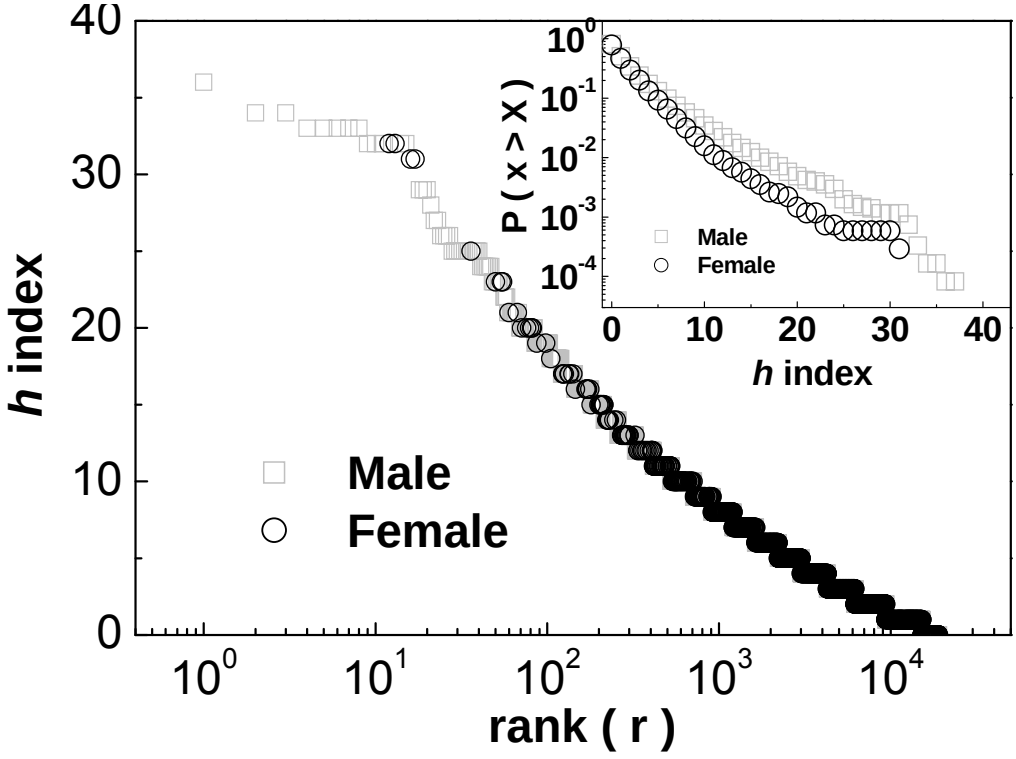

Figure 3: Ranking of Brazilian male and female authors according to h index value. Inset: Complementary cumulative distribution for Brazilian h index of male and female authors.

We have also observed that only 10% of the Brazilian authors present a higher than 5 h-index value. Within this set, the probability of finding a male author is 13%, whereas finding a female author, 9.2%. On the other hand, when we analyzed the performance of the other 90% of our sample whose h-index value was less than 5, we observed that the probability of finding a male author is the same as finding a female one. This result can be better noticed in Figure 3 – inset, where the Brazilian authors are graphically represented according to a complementary cumulative distribution in h-index terms.
Again, considering these trends, we also found evidences that gender may not be a dependent variable for scientific performance for the largest part of our sample.

Gender variable and performance among the fields

Although we have been dealing with a large sample set, it is important to remark that the data come from individual performances. In this case, different factors may be influencing directly or indirectly each performance. Among scientific and / or academic environment factors, research field seems to play an important role.
In fact, it is well-known that the scientific production process differs reasonably among research fields in various aspects; for example, publishing time, size of research group and collaborative work. With different characteristics, many analyses on research fields have pointed towards different trends and profiles. Within this context, we decided to investigate the relation between the research field and the performance of Brazilian male and female authors.
As indicated in Table 1, three research fields - Humanities, Social Sciences and Language & Arts - are not well-represented in the CNPq-BSI database. Therefore, we decided to exclude these fields from this analysis. Notice that Figure 4 presents the complementary cumulative

distribution of Brazilian authors, according to the number of publications (main figure) and h-index (the inset) among different research fields.

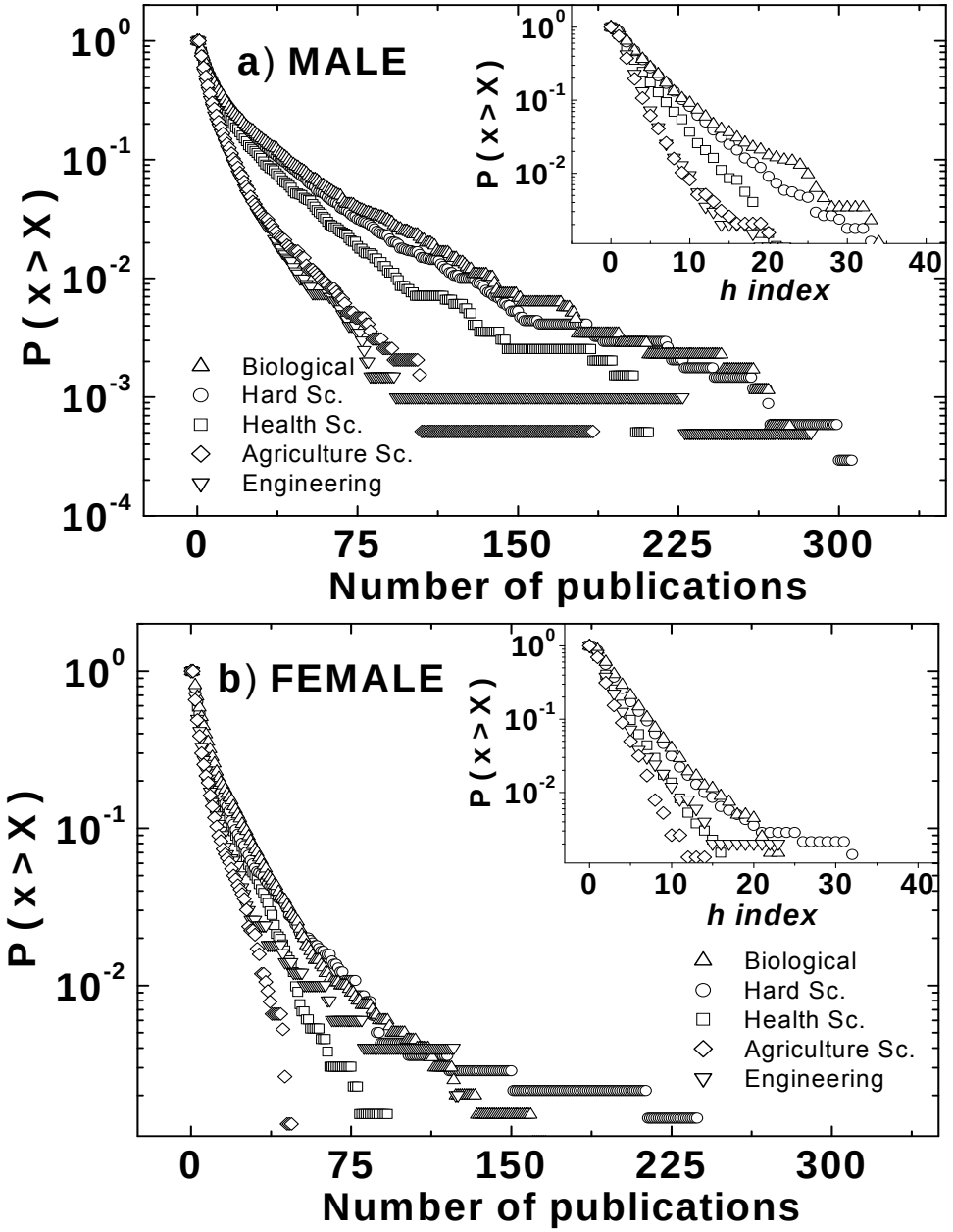

Figure 4: Complementary cumulative distribution for number of publications of Brazilian male (A) and female (B) authors (A) in different research fields.
Inset: Complementary cumulative distribution for h-index values of Brazilian male (A) and female (B) authors (A) in different research fields.

As previously described, this type of representation allowed us to identify the probability to find a male or a female author with an "n" number of publications or an "x" h-index value.
Taking into account the male set (Figure 4A), we noticed that three groups of authors were formed as the number of publications increased: authors of Biological & Hard Sciences, Health Sciences and Agriculture & Engineering. Considering, for example, authors with ≥ 14 publications (only 20% of authors out of the whole sample), we found 29.2% and 29.9% of

chances in finding authors of Biological and Hard Sciences, 25.00% of chances in finding Health Sciences authors and 13 % and 12% Agriculture and Engineering ones. The data indicates that the chances to find more productive authors are higher among Biological and Hard Sciences authors.

As for the female set (Figure 4B), the tendency of forming groups of authors is not as clearly observed as it is for the male set. Moreover, the chances to find a more productive one does not seem to be related to a specific research field, as it is clear to males. These findings suggest that the research field is a striking variable for male authors' scientific performance, but not for female authors. Nevertheless, looking Figure 4B carefully, the highest productive female authors are among Hard Sciences authors. Curiosly, women are underrepresented in this field (Table 1). Thus, being highly productive may be a compensatory mecanism to be compete for funds and positions.

Similar trends are observed in the h-index analysis (Figure 4 – inset). Again, for the male set - not for the female, though - we can clearly notice three groups. Thus, as the complementary cumulative distribution indicates, the probability to find male authors with an h-index ≥ 5 (only 17% of all authors) is higher among Biological and Hard Sciences authors. Yet, this tendency is not observed among the female set.

Discussion and Conclusion

Although we have carried out an analysis with almost 19,000 Brazilian authors, two main remarks should be observed. First, the efficiency of the crossing database process was around 37% for the whole list of Brazilian researchers catalogued in CNPq´s database. In other words, for each 100 Brazilian researchers listed in this database, only 37 were retrieved from BSI´s database and compiled into CNPq-BSI´s database. Thus, we cannot assure that similar results would be found out, if lacking Brazilian authors were included in the analysis.

A second remark is related to the format in which we performed most of the analysis. We have opted to present most of the data according to the complementary cumulative distribution, which allows one to clearly identify how often the random variables (number of publications and h-index value) are above or below a particular level. On the other hand, this analysis does allow one to observe quantitative differences among female and male authors´ scientific performance. Thus, any discussion upon these results must consider that they illustrate probability, not actual numbers.

Taking these two remarks into consideration, we have got evidences that gender differences in scientific performance, estimated by the number of publications and h-index value, are not as clear as it may appear in the international literature. We have found that within the largest fraction of Brazilian authors included in our sample, male and female authors have the same chances to publish n papers or to have x h-index value. A different scenario was only observed for the most productive authors, as well as for authors with the highest h-index value. Within these select groups, the probability to find male authors is much higher than it is to find a female author.

These findings arose many questions, such as: were females less productive because they were not in the fields that are more likely to have productive authors? Or they are less productive because they are novelties in these fields? Are there any other academic conditions behind this difference: salary, grants, laboratory infra-structures, research group?

As for the select group, we are also pushed to question: who are the authors in these two select groups? What are the main differences among these groups´s male and female authors? Is the different scientific performance influencing on their carreer ascent? Does the fact of being in these select groups confer female authors any rewards, even being in minority?

Questions such as these reinforce the notion that gender studies applied to the scientific community are too complex and involve many social and cultural aspects. Such complexity

can not be understood only by a qualitative analysis. On the other hand, such analysis may serve as a basis to confirm or not some hypoteses. In this paper, the main hypotesis came from the international literature "women are less productive than men". Nevertheless, our data indicated that this hypotesis was only true for the smallest fraction of our sample. Thus, details and characteristics of both female and male authors of this select group should also be investigated in details.

Hence, in the next steps, we plan to explore these questions not only by means of quantitative approach but also by a qualitative one. As for the whole sample, we intend to investigate the performance of males and females in different time periods. This would allow observing the changes in the performance of females. Also, within this approach we will investigate the weight of collaboration in the performance of males and females. As for the select group of most productive Brazilian scientists, the main idea is firstly to investigate their academic and scientific life. We intend to to this with all the 100 top male and female scientists as well as with those with an h-index value higher than 5 (17% of our sample). Such qualitative approach will be based on the analysis of authors' curricula which are available and freely accessed in CNPq´s on-line database (CNPq, 2008).

Finally, it is noteworthy to mention that this paper does not intend to present either the state of the art of gender studies to be exhaustive. Instead, this paper aims at contributing to the international debate on the current participation of women in the scientific environment. Questions about gender, either in science environment or not, are a multifaceted issue, involving many social and cultural aspects. Thus, studies with quantitative or qualitative approaches are still required to get a better understanding of its complexity.

Acknowledgments

We thank CNPq - especially the support of Felizardo P. da Silva, Silvana M. Cosac and Cristiano L. Kuppens - for the database set up for this study and Vicky Adler for the revision of this manuscript.

References

Batista, P.D., Campiteli M.G, Kinouch, O. & Martinez, SA. (2006) Is it possible to compare researchers with different scientific interests? *Scientometrics* 68, 179-189.

Bornmann, L. & Daniel, H. (2005) Does the h-index for ranking of scientists really work? *Scientometrics* 65, 391.392.

Bourdieu, P. Os usos sociais da ciência. Por uma sociologia clínica do campo científico. São Paulo: Editora UNESP, 2003.

Braun, T., Glanzel, W. & Schubert, A. (2005) A Hirsch-type index for journals. *The Scientist* 19, 8.

CNPq, Conselho Nacional de Desenvolvimento Cientifico e Tecnológico. Plataforma Lattes. (2008) Available at: http://lattes.cnpq.br/

Costas, I. (2002) Women in Science in Germany. *Science in Context* 15, 557–576.

Finson, K.D. (2002). Drawing a Scientist: What We Do and Do Not Know After Fifty Years of Drawings. *School Science and Mathematics*, 102, 335-345.

Fox, M.F. (2005) Gender, Family Characteristics, and Publication Productivity among Scientists. *Social Studies of Science* 35, 131–150.

Gupta, N. & Sharma, A.K. (2002) Women Academic Scientists in India. *Social Studies of Science* 32, 901–915.

Hermann, C. & Cyrot-Lackmann, F. (2002) Women in Science in France. *Science in Context* 15, 529–556.

Hirsch, J.E. (2007) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci* 102, 16569-16572.

Leta, J. (2003) As mulheres na ciência brasileira: crescimento, contrastes e um perfil de sucesso, Revista de *Estudos Avançados* 49.

Long, J.S. (1992) Measures of Sex Differences in Scientific Productivity. *Social Forces*, 71 (1), 159-178.

Mauleón, E. & Bordons, M. (2006), Productivity, impact and publication habits by gender in the area of Materials Science, *Scientometrics* 66, 199–218.

McGregot, E. & Harding, S. (1996) "*Science by whom*?" The Gender Dimension of Science and Technology, UNESCO, World Science Report.

Mead, M. & Metraux, R. (1957). Image of the scientist among high school students. *Science,* 126, 384-90.

Mugnaini, R., Packer, A.L. & Meneghini, R. (2008) Comparison of scientists of the Brazilian Academy of Sciences and of the National Academy of Sciences of the USA on the basis of the h-index. *Brazilian Journal of Medical and Biological Research* 41, 258-262.

NSF, National Science Foundation. *Women, minorities and persons with disabilities in science and engineering: 1998*. Arlington, VA, 1999 (NSF 99-338)

Pripic, K. (2002), Gender and productivity differentials in science, *Scientometrics*, 55 : 27–58.

Rossi, A. S. (1965) Women in Science: Why so Few? Social and Psychological Influences Restrict Women's Choice and Pursuit of Carriers in Science. *Science*, 148: 1196-1202.

Schienbinger, L. (2001) O feminismo mudou a ciência? Bauru, Edusc. (translation of *Has Feminism Changed Science? (1999)* Cambridge, Mass.: Harvard University Press.)

Symonds, M.R. E., Gemmell N. J., Braisher, T.L., Gorringe, K.L & Elgar, MA. (2006) Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance, *PlosOne*, Issue 1, 1- 5.

UNESCO. *A global perspective on research and development.* (2007a) UNESCO Institute for Statistics, Fact Sheet October. Retrieved November, 2008 from: http://www.uis.unesco.org/template/pdf/S&T/Factsheet07_No%20%205_ST_EN.pdf

UNESCO. *Science, Technology and gender. International report*. (2007b) Retrieved November, 2008 from: http://www.unesco.org/science/psd/focus/focus07/irstg-executivesummary-english.pdf

Zuckerman, H. & Cole, J.R., (1975) Women in American Science. Minerva, 13 (1), 82-102.