Research Evaluation, 26(2), 2017, 115–123
doi: 10.1093/reseval/rvx008
Article

OXFORD

Gender and research funding success: Case of the Belgian F.R.S.-FNRS

Raphaël Beck* and Véronique Halloin

Fonds de la Recherche Scientifique—FNRS, Rue d'Egmont 5, Bruxelles 1000, Belgium

*Corresponding author. Email: raphael.beck@frs-fnrs.be.

Abstract

The influence of gender on the outcome of research evaluation activities and access to research funding has been heavily debated in recent decades. In this study, data from 6,393 applications submitted between 2011 and 2015 to the Belgian funding agency Fonds de la Recherche Scientifique - FNRS (F.R.S.-FNRS) were statistically analysed to highlight any possible effect of gender on success rates. Results show no significant influence of gender on success rates or the likelihood of getting funding for most of the funding schemes we analysed. Research credit (RC) was the only one where gender and success variables were statistically dependent, although mean success rates of male and female applicants were not significantly different. Average grades given by remote reviewers to male applicants were significantly higher in the frame of RC applications. Among RC applications, the difference in success rates was higher in Humanities and Social Sciences, followed by Exact and Natural Sciences, and finally Life and Health Sciences. Proportions of male researchers who apply were shown to be higher for most of the funding schemes analysed, mainly for grant applications (such as RC) where only tenure researchers are allowed to apply. Taken together, our results show that access to F.R.S.-FNRS funding is not gender-dependent for the majority of the funding schemes except one where men represent the vast majority of the applicants. Reasons that could explain this statistical dependence are under investigation and could be due to the lower grading of women by remote reviewers.

Key words: gender; research funding; peer review; funding agency; research evaluation.

1. Introduction

There has been intense debate in the scientific community over the past two decades regarding gender disparities amongst scientists and about the impact of gender on the likelihood and ease of pursuing a successful scientific career. One of the potential concerns is that women could be less likely to obtain funding or academic positions due to, amongst other things, a conscious (explicit) or subconscious (implicit) bias that would disadvantage them when it comes to peer-review and research evaluation processes. This bias would be caused by cultural stereotypes that lead to the implicit assumption that women are less competent than men, at least in certain scientific fields (Wenneras and Wold 1997; Kaatz, Gutierrez and Carnes 2014), such as science, technology, engineering, and mathematics (STEM; Leslie et al. 2015). Such biases are considered to be well founded in the scientific community. For example, in 2015, the British Nobel Prize winner Tim Hunt prompted numerous negative reactions and was asked to step down from various influential roles following his controversial pronouncements about female scientists (Nature Editorial 2015). However, no rational basis justifies the existence of these stereotypes, since male and female researchers' performances are not considered to significantly differ, and since these days young female researchers are sometimes even seen to be starting to outperform their male counterparts (van Arensbergen, van der Weijden and van den Besselaar 2012).

A recent publication stated that gender contributes to personal research funding in the Netherlands (van der Lee and Ellemers 2015), although this publication became rapidly controversial when other authors criticized the statistical methodology that was employed, suspecting a potential Simpson's paradox that could have biased the results obtained, since proportions of women were higher in disciplines where success rates were lower (Albers 2015; Volker and Steenbeek 2015). However, it seems in general well accepted that female scientists face discrimination and funding disparities (Bedi, Van Dam and Munafo 2012; Moss-Racusin et al. 2012; Shen

2013; Watson and Hjorth 2015), although men have been described as less receptive to this type of experimental evidence (Handley et al. 2015). Even if signs of improvement do exist, several members of the scientific community are calling for proactive reactions from major players in the field of research to close the gender gap (Boyle et al. 2015), while some others are less supportive of actions such as the implementation of quotas, due to several reasons including the absence of correlation between success rates of female applicants and the gender balance of evaluation panels (Vernos 2013). Therefore, literature about gender in research remains somehow contradictory, and since an influence of gender on peer review or grant allocation is not systematically reported by all authors (Marsh et al. 2009; Pohlhaus et al. 2011; Ceci et al. 2014), we wanted to find out if our own data could demonstrate any gender bias as regards the allocation of grants and resources.

Fonds de la Recherche Scientifique - FNRS (F.R.S.-FNRS) is a Belgian research funding agency that was founded in 1928 as a private foundation of public interest and that funds basic research in the French-speaking universities of Belgium using a bottom-up approach. Its annual budget is about 180 million euros, and it has approximately 2,000 fellows on its payroll. One of the key missions of the organization is research evaluation to allocate resources on the basis of scientific excellence. Within this context, F.R.S.-FNRS annually processes around 2,000 applications, of which around 1,200 applications for fellowships and 800 applications for grants. One of the key principles applied by F.R.S.-FNRS is that no strategic choices are made in terms of priority scientific fields, and therefore, success rates are equally distributed across all scientific disciplines for most of its funding schemes. Nevertheless, in the context of some funding schemes that are not analysed in this study, a top-down approach is applied instead of a bottom-up one, and some strategic choices are made in terms of funding (e.g. Fonds pour la formation à la recherche dans l'industrie et dans l'agriculture (FRIA) that funds more applied research in industry and agriculture and Fonds pour la recherche en sciences humaines (FRESH) that funds research in social sciences and humanities).

As a research funding agency, we have access to a lot of data about applications and funded research that we can use for analysis. F.R.S.-FNRS regularly carries out analyses and monitoring to improve its rules and regulations and to pick up on any potential bias in its evaluation processes. Up to now, we could never demonstrate any gender bias when it comes to funding allocation during specific calls for proposals. In this study, we statistically analysed data from several calls (6,393 applications in total over 5 years and 10 calls for proposals) and tried to identify any potential gender bias that could have occurred in the ex-ante evaluation processes that the Fund organizes.

The general F.R.S.-FNRS evaluation process is composed of two main steps: a first round of remote reviews of submitted proposals by international remote reviewers (not for doctoral applications) and a second round of panel meetings which results in a final consolidated ranking of applications received. Panel members have remote reviews at their disposal to help them making their final decisions but are not obliged to follow the opinion of remote experts nor the grades they gave. There are 14 different thematic scientific panels for one single funding scheme, and success rates are distributed equally across all scientific panels. During one call for proposals, F.R.S.-FNRS typically gathers around 2,500 remote reviews from 1,500 individual remote reviewers and 750 additional reviews from panel members. F.R.S.-FNRS uses a hybrid classification system to characterize scientific fields of proposals and persons. This system is mainly based on the ERC classification system but also includes some specific descriptor fields to take the particularities of research carried out in Belgian French-speaking universities into account. Globally, over 5 years, thematic scientific panels were composed of 18.4% of women and 81.6% of men. In total, 23.3% of remote reviewers were women.

2. Materials and methods

2.1 Data collection and control

Data were collected from the F.R.S.-FNRS SEMAPHORE database which contains all the application data that the Fund collects as part of the different calls for proposals it organizes (two main calls per year). All data were controlled for completeness and absence of duplicates. We focused on new applications (no renewals were considered, since success rates are generally very high for this kind of application), and only applications that went through the entire evaluation processes were considered (no withdrawn applications or ineligible ones). We focused on data from 2011 to 2015 for our analyses, for two main reasons. First, one of the aims of this study was to focus on relatively recent data to have a view of the current situation, and with this time window of 5 years, we could include data from 6,393 applications, which was enough to be able to perform robust statistical analyses. In addition, although F.R.S.-FNRS was founded in 1928, before the year 2011, the structure of the available data was slightly different and would have needed additional adaptations to be included in this study. Moreover, funding schemes as well as rules and regulations were significantly different before 2011, making comparisons difficult.

2.2 Funding schemes considered

Applications to several funding schemes were analysed. These could be divided into two distinct groups: fellowships (DOC and POSTDOC) that basically fund the salary of the applicant, and grants (Research Credit (RC), Research Project (RP)) that fund research carried out under the supervision of the applicant (only researchers holding tenure positions or assimilated can apply for grants). Here is a brief description of each of them:

Fellowships (F.R.S.-FNRS is the employer of funded applicants who are hosted within universities):

- DOC (research fellow): grants a fellowship to the successful applicant that allows him to carry out a PhD thesis. Duration: 4 years (2 years, renewable once for 2 additional years). Success rates are the same across all scientific fields for this funding scheme.
- POSTDOC (postdoctoral researcher): grants a 3-year postdoctoral fellowship to the successful applicant. Success rates are the same across all scientific fields for this funding scheme.

Grants (F.R.S.-FNRS provides money to tenure researchers who are hosted within universities to carry out their research):

- RC: small grant that provides 5,000–30,000 euros per year (initially over 1 year, and since 2015 over 2 years) to the successful applicant to fund his research (equipment and operating costs). Success rates are the same across all scientific fields for this funding scheme.
- RP: grant that provides 180,000–520,000 euros (over 2–4 years) to the successful applicant to fund his research (equipment, personnel, and operating costs). Success rates are the same across all scientific fields for this funding scheme.

2.3 Statistical analyses

Data were extracted from the F.R.S.-FNRS SEMAPHORE database to Microsoft Excel. All data are presented as means ± SDs. Where appropriate, statistical significance levels were reported using asterisks with the following meanings: *($p \leq 0.05$); **($p \leq 0.01$); ***($p \leq 0.001$). Most of the statistical analyses were carried out using GraphPad Prism software (GraphPad Software Inc., USA), R (R Development Core Team, http://www.R-project.org), and in some cases directly with Microsoft Excel.

To statistically analyse the data from different points of views and make sure to not miss any potential effect, our approach was the following: we first compared mean annual success rates (1 year = one value) by using mean comparison techniques (analysis of variance (ANOVA)). We then tested the independence of success and gender variables in contingency tables using Pearson's χ^2 analyses. We also performed the latter analyses scientific field by scientific field (using exact Fisher tests when sample sizes were too small) to identify in which subfield(s) a dependence could be observed. In addition, we also performed logistic regression (using R software) taking into account different variables (gender, year, subfield, and funding scheme) to test their influence on success and model the probability of success depending on these variables. We also performed unpaired *t*-tests where appropriate to compare means (mean grades awarded by remote reviewers and mean granted amounts of money).

3. Results

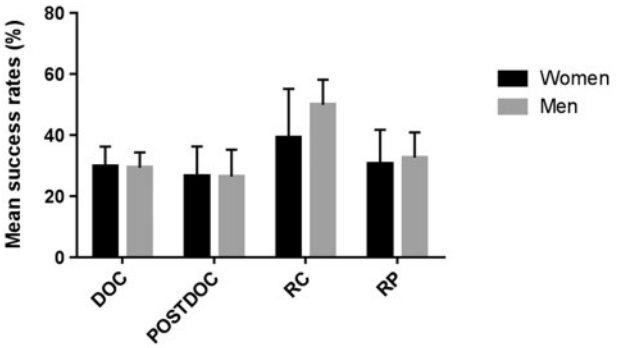
Since success rates applied by F.R.S.-FNRS differ depending on the funding scheme considered, data are always presented and analysed funding scheme by funding scheme to avoid any Simpson's paradox (Simpson 1951) during data analysis. Such a paradox states that an apparently significant relation between two variables may be due to a joint dependency on a third variable (Albers 2015). In this study, for instance, it could occur if the share of women or men amongst applicants happened to be higher or lower in subgroups (disciplines/funding schemes) where success rates are significantly higher or lower. However, success rates are, for most of the funding schemes and notably for all funding schemes analysed in this study, equally distributed across all scientific disciplines at F.R.S.-FNRS. For a given funding scheme, success rates may also vary slightly depending on the year in question because of budgetary constraints, but we observed that the proportions of female applicants were very similar over the different years considered. Therefore, we also performed logistic regression from the success data to different variables to test for their potential influence.

We focused on a set of funding schemes that include most of the applications we receive. These funding schemes are also very standard, since they exist in quite similar forms in most of the countries and funding agencies. Explanations about the funding schemes presented are given in the 'Materials and Methods' section.

We first calculated the annual success rates observed year by year, and then the mean success rates (±SD) over 5 years and 10 calls (five fellowship calls and five grant calls). Indeed, potential gender biases are often assessed by simply comparing success rates of women and men in the frame of different funding schemes. Results are shown in Fig. 1. No significant difference could be observed in terms of success rates, no matter which funding scheme was considered. Indeed, we performed a two-way ANOVA on data to compare means, and no statistically significant difference was found between female and male means ($F(1, 32) = 0.9561$, $P = 0.3355$). This indicates that on a global scale and in relation to a large number of applications, mean success rates amongst women do not significantly differ from mean success rates amongst men, no matter which type of funding scheme is taken into account.

In Table 1, the success rates and total number of female and male applicants for all applications considered are given, year by year and funding scheme by funding scheme. Although mean success rates are clearly decreasing with years, due to budgetary constraints and increasing amounts of applications, we cannot see a clear time tendency regarding the evolution of the differences (Δ) between rates of success of women and men.

The fact that female and male mean success rates do not significantly differ does not necessarily mean that the probability of success is independent of the gender of the applicant. Therefore, we performed several Pearson's χ^2 tests of independence to verify that gender and success variables were statistically independent for each funding scheme considered (Table 2). For most of the analysed funding schemes (three of four in total) and notably for all fellowship applications, both variables were statistically independent, meaning that gender was statistically independent of success during the different calls for proposals that were taken into account. However, for RCs (small grants that provide 5,000–30,000 euros per year (initially over 1 year, and since 2015 over 2 years) to the funded applicants to fund their research (equipment and operating costs) and for which applicants must be tenure researchers to apply), the alternative hypothesis (H1) showed true, meaning that both variables could be considered as statistically dependent. Over the 5 years considered, women had a success rate of 38.0% in the context of RC, whereas male applicants had a success rate of 49.8%. The general success rate for this funding scheme over the 5 years considered was 46.0%.



	Women			Men		
	Mean	SD	n	Mean	SD	n
DOC	29.81	6.39	5	29.27	5.01	5
POSTDOC	26.52	9.71	5	26.34	8.80	5
RC	39.19	15.94	5	49.89	8.21	5
RP	30.58	11.10	5	32.59	8.28	5

Figure 1. Mean success rates for female and male applicants to several F.R.S.-FNRS funding schemes (2011–5, n = 5 years, one call for proposals per instrument and per year). Data are presented as means (n = 5) ± SDs in a graph (above) and given in a table (below). A number of applications that were taken into consideration for calculations: DOC: five calls and 2,238 applications (1,186 women and 1,052 men); POSTDOC: five calls and 1,852 applications (830 women and 1,022 men); RC: five calls and 890 applications (292 women and 598 men); and RP: five calls and 1,413 applications (389 women and 1,024 men). In total, 6,393 applications were considered, of which 2,697 applications by women and 3,696 applications by men.

Table 1. Mean success rates, number of applicants, and difference (Δ) in terms of success rates (male−female) for female and male applicants to several F.R.S.-FNRS funding schemes, per year of application and funding scheme (2011–5, one call for proposals per instrument and per year). In total, 6,393 applications were considered, of which 2,697 applications by women and 3,696 applications by men

Funding scheme/year	n (women)	n (men)	Women success rates (%)	Men success rates (%)	Δ (Men−women) (%)
DOC	1,186	1,052	29.43	29.18	−0.24
2011	199	198	38.69	35.35	−3.34
2012	242	205	33.88	33.17	−0.71
2013	251	227	28.29	28.63	0.35
2014	252	219	24.21	26.03	1.82
2015	242	203	23.97	23.15	−0.81
POSTDOC	830	1022	25.06	24.56	−0.50
2011	118	141	41.53	35.46	−6.06
2012	143	159	29.37	35.85	6.48
2013	181	203	25.41	21.67	−3.74
2014	171	251	16.96	22.31	5.35
2015	217	268	19.35	16.42	−2.94
RC	292	598	38.01	49.83	11.82
2011	54	130	29.63	45.38	15.75
2012	53	117	66.04	64.10	−1.94
2013	52	89	40.38	43.82	3.44
2014	56	97	33.93	49.48	15.56
2015	77	165	25.97	46.67	20.69
RP	389	1,024	29.31	32.23	2.92
2011	69	190	42.03	40.00	−2.03
2012	63	189	42.86	41.80	−1.06
2013	81	223	25.93	32.29	6.36
2014	89	215	19.10	23.72	4.62
2015	87	207	22.99	25.12	2.13

Table 2. Results from the Pearson's χ^2 analyses performed on all contingency tables generated on the basis of the number of grants awarded and not awarded to female and male researchers, per funding scheme. n is the total amount of applications that were taken into consideration for the calculations.

Funding scheme	n	χ^2 sum	Degrees of freedom	$\chi^2_{df=1;\,P=0.05}$
DOC	2,238	0.016	1	3.841
POSTDOC	1,852	0.062	1	3.841
RC	**890**	**11.035**	**1**	**3.841**
RP	1,413	1.116	1	3.841

The statistical significance of the independence test (P = 0.05) is highlighted in bold.

We then performed a logistic regression analysis (logit) from the success data (success variable being defined as 1 if successful and 0 if unsuccessful) to variables 'gender' + 'year' + 'funding scheme' + 'subfield' to determine which ones had an influence on success (Table 3, top). Year and funding scheme were the only variables that were statistically significant. Gender was not statistically significant, suggesting no association between gender and success. Variable 'year' had the lowest P-value, suggesting a stronger association of this variable with the probability of success. Variable 'funding scheme' was also significant. However, these results were expected because success rates are decided by the governing board of the organization and largely vary from one funding scheme to another and from 1 year to another (e.g. success rates are higher in the frame of the RC funding scheme).

In addition, we also performed interaction analyses of some variables ('gender' * 'year', 'gender' * 'funding scheme' and 'gender' * 'subfield') using the logistic regression analysis (logit) from the success data to the different analysed variables. No statistically significant interaction between variables 'gender' and 'year' or between variables 'gender' and 'subfield' could be found. However, we found a significant interaction between variables 'gender' and 'funding scheme' for the RC funding scheme (Gender Men:Funding Scheme RC, estimate = 0.507, standard error = 0.175, z value = 2.895, Pr(>|z|) = 0.004, significance summary = **) (Table 3, bottom). This indicates that variables 'funding scheme' and 'gender' interact and that being a male applicant increases significantly more the probablity of getting a RC grant as compared to being a female applicant.

A potential explanation of the lower success rates of women, at least in the context of RC, could be that they receive lower grades from peer reviewers who are in charge of reviewing the proposals they submit, due, for example, to a biased peer review, as was previously suggested by some authors (Kaatz, Gutierrez and Carnes 2014). Indeed, applications to F.R.S.-FNRS for most of the funding schemes (all of the ones analysed here except for 'DOC', thus including RC) are assessed in two independent steps, a first review by several remote reviewers and a consolidated final evaluation by a thematic panel which meets once at the end of the process and relies on remote evaluations to give a final grade and ranking to each application. To test this hypothesis, we analysed the mean grades received by men and women within the context of their RC applications. To do so, grades given by reviewers were converted into figures (A+ (outstanding) = 100; A (excellent) = 95; A− (very good) = 90; B+ (good) = 85; B (average) = 80; B− (weak) = 75; C (insufficient) = 65). Results are shown in Fig. 2. Means were compared using an unpaired *t*-test and differed significantly (P value = 0.0004; $t = 3.541$; df = 888). On average, male applicants were awarded significantly higher grades by remote reviewers than female applicants in the context of RC applications.

Table 3. Results from the logistic regression analysis (logit model) performed on the 6,393 applications and modelling the probability of success or unsuccess depending on the variables 'gender', 'year', 'funding scheme', and scientific 'subfield' taken independently (top) or considering interactions (bottom). As regards discrete variables, the following values were selected as references: 'Gender': women; 'Funding scheme': doc; 'Subfield': 'Chemistry, materials'.

Coefficients:	Estimate	Standard error	z value	Pr(>|z|)
Variables considered independently ('gender', 'year', 'funding scheme', and 'subfield')				
Gender: men	0.082	0.059	1.391	0.164
Year	−0.196	0.020	−9.947	<2e-16***
Funding scheme: RC	0.718	0.089	8.096	5.70e-16***
Funding scheme: Postdoc	−0.192	0.073	−2.642	0.008**
Funding scheme: RP	0.087	0.080	1.081	0.280
Subfield: 'Sustainable development'	−0.185	0.252	−0.733	0.463
Subfield: 'Mathematics, Physics'	0.056	0.148	0.377	0.706
Subfield: 'Engineering'	−0.002	0.146	−0.015	0.988
Subfield: 'Earth sciences, Biology, Agronomy'	−0.023	0.138	−0.165	0.869
Subfield: 'Sociology, Anthropology, Political sciences, Communication'	−0.041	0.143	−0.288	0.773
Subfield: 'Psychology, Education sciences'	0.006	0.154	0.039	0.969
Subfield: 'Literature, Arts, Linguistics'	−0.036	0.141	−0.255	0.799
Subfield: 'History, Archaeology'	0.063	0.154	0.410	0.682
Subfield: 'Economics, Law'	−0.020	0.158	−0.126	0.900
Subfield: 'Genetics, Biochemistry, Cell biology'	−0.021	0.141	−0.152	0.879
Subfield: 'Physiology, Microbiology, Immunology, Veterinary medicine'	0.040	0.145	0.273	0.785
Subfield: 'Neurosciences'	0.070	0.177	0.395	0.693
Subfield: 'Medicine, Pharmacy, Dentistry'	−0.036	0.149	−0.244	0.808
(Intercept)	394.370	39.737	9.925	<2e-16***
Interaction between 'gender' and 'funding scheme'				
Gender Men:Funding Scheme RC	0.507	0.175	2.895	0.004**

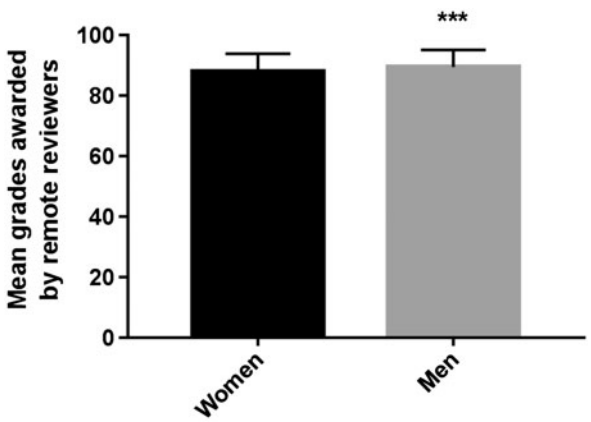
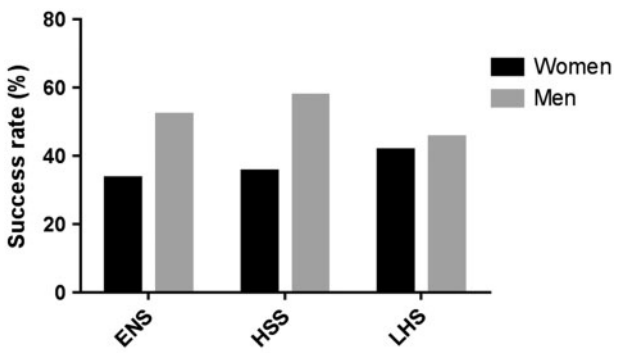


Figure 2. Mean grades awarded by remote reviewers to men and women in the context of their application to the RC funding scheme (2011–5). Data are presented as means ± SDs. A number of applications that were taken into consideration for calculations: five calls and 890 applications (292 women and 598 men).



Scientific field	n	χ^2 sum	Degrees of freedom	χ^2 df=1 ; (P = 0.05)
ENS	**303**	**8.098**	**1**	**3.841**
HSS	**143**	**6.536**	**1**	**3.841**
LHS	441	0.581	1	3.841

Figure 3. Top: success rates (2011–5) of men and women for the RC funding scheme, per scientific field (ENS, HSS, and LHS). Below: results from Pearson's χ^2 tests of independence of gender and success variables for the three scientific fields analysed. The statistical significance of the independence tests (P=0.05) is highlighted in bold.

It was postulated that biases in evaluation occur mostly in scientific fields where women would be considered less competent than men, such as STEM (Leslie et al. 2015). We therefore decided to focus our analysis of RC success rates of women and men on scientific disciplines to pick up on any potential discipline-linked bias. F.R.S.-FNRS applications are mainly classified either as 'Life and Health Sciences' (LHS), 'Humanities and Social Sciences' (HSS), or 'Exact and Natural Sciences' (ENS), which also includes engineering. We performed a specific analysis on this latter category. Surprisingly, the biggest difference in terms of success rates was in HSS (35.29% for women vs. 57.61% for men), then in ENS (33.33% for women vs. 52.0% for men), and finally in LHS (41.61% for women vs. 45.36% for men). Statistical analyses (Pearson's χ^2 tests of independence) showed that for ENS and HSS, gender and success variables were statistically dependent. As regards LHS, both variables were independent (Fig. 3).

To further investigate this, we analysed interactions between success and gender variables subfield by subfield, to determine if interactions were observable in specific scientific subfields only or if it was general for all ENS and HSS. To do so, we used Fisher's exact

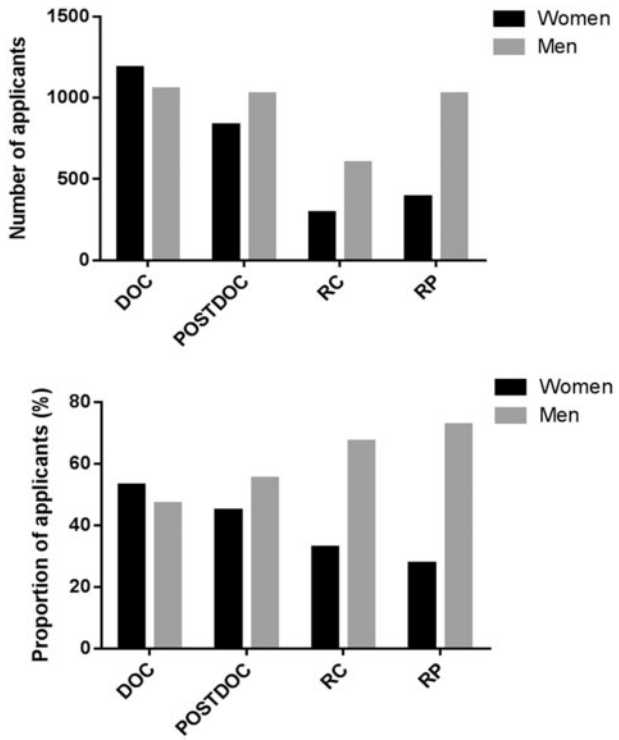


Figure 4. Absolute values (upper panel) and proportions (lower panel) of applicants to the different funding schemes that were taken into consideration in the context of this study (6,393 applications in total, of which 2,697 applications from women and 3,696 applications from men).

tests applied on data from all scientific subfields (instead of χ^2, since sample sizes were small for some of the subfields analysed). As shown in Table 4, we observed that success rates are higher for male researchers, in comparison with female researchers, in almost all scientific subfields. However, a significant dependence between success and gender variables was shown in only two scientific subfields, namely, 'Earth sciences, Biology, Agronomy' and 'History, Archaeology'. In these two subfields, the share of women amongst applicants is quite high. In addition, the total share of women amongst panel members in the two subfields is not the lowest but rather on the upper side, indicating that a lower representation of women amongst panel members in these fields is not a reasonable explanation for the statistical dependence observed. Furthermore, we also compared means of the grades given by remote reviewers in the different scientific subfields. Using *t*-tests, we could show significant differences of average grades given by remote reviewers only in Sociology, Anthropology, Political sciences, Communication (P value = 0.0181; $t = 2.477$; df = 36) and in 'History, Archaeology' (P value = 0.0083; $t = 2.915$; df = 21). Therefore, 'History, Archaeology' is the only scientific field for which we observed in the frame of RC applications both a significant difference between women and men applicants in terms of average grades given by remote reviewers and a statistical dependence between gender and probability of success.

As regards the total number of applications received and the proportions of men and women applying to fellowships, data showed that the proportion of male applicants increased with the mean seniority level of the requested fellowship. As regards grants (where only researchers with tenure positions are allowed to apply), proportions of male researchers applying were always higher than proportions of female researchers, and proportions also increased with the mean relative importance of the grant requested (Fig. 4). This could be due to the fact that, amongst researchers who are allowed to apply for such grants (tenure researchers from Wallonia-Brussels Federation (WBF) universities, mainly the academic staff), the proportion of male researchers is much higher than the proportion of female researchers (female researchers represent a proportion of just 26.5% of the academic personnel of WBF universities, in terms of full-time equivalents (FTEs)).

As at least in the context of one of the funding schemes amongst the grants (RC) we could demonstrate a statistical dependence between the gender and success variables, we wanted to analyse the average amounts of money granted to men and women who were awarded a grant from F.R.S.-FNRS in the past 5 years. Results are shown in Fig. 5. No difference could be demonstrated between the mean amounts granted to women and mean amounts granted to men, whatever the funding scheme considered. Means were compared funding scheme by funding scheme with unpaired *t*-tests, and no statistical significance could be established (values were the following: RC: P value = 0.9224; $t = 0.09744$; df = 407; RP: P value = 0.1879; $t = 1.319$; df = 442).

4. Discussion

The impact of gender on the access to research funding and publication has been the subject of intense debate over the past 20 years, and it was often postulated that female researchers were at a disadvantage compared to male researchers when it came to fellowship applications (Wenneras and Wold 1997), recruitment of academic personnel in universities (van den Brink, Brouns and Waslander 2006; van den Besselaar and Sandström 2016), research funding (van der Lee and Ellemers 2015), peer reviewing of articles (Walker et al. 2015) or peer review in general (Kaatz, Gutierrez and Carnes 2014), the amount of money received with awarded research grants (Bedi, Van Dam and Munafo 2012), and so on (Shen 2013). However, literature remains somewhat contradictory, since other studies fail to demonstrate any gender bias or support a contrary hypothesis following which no evidence of it can be found (Marsh et al. 2009; Pohlhaus et al. 2011; Ceci et al. 2014; Volker and Steenbeek 2015).

This study presents data analyses of a large number of fellowship and grant applications submitted to F.R.S.-FNRS during the 2011–5 period (5 years and 10 calls, n = 6,393). Results from our analyses show that for the majority of the applications (97.8% of them, 6,250 of 6,393) and funding schemes that F.R.S.-FNRS handles, with the exception of one, no statistically significant difference between male and female applicants can be observed in terms of success rates. This tends to demonstrate that gender does not play a significant role in the process of selecting fellowship and grant awardees by the different evaluators involved in the review processes (about 1,500 different experts per call for proposal, each of them assessing on average 1.5 applications per call). In addition, logistic regression from the success data to different variables including gender failed to show significance of the gender variable on the probability of success, indicating that gender variable is most likely not a good predictor of the success of an application in a logistic model.

Nevertheless, for one single funding scheme (RC) of the four we studied, we could demonstrate through Pearson's χ^2 analyses that

Table 4. Results from Fisher's exact tests performed to verify the independence between success and gender variables across all scientific subfields in the context of RC applications. The last subfield (sustainable development) was excluded from the tests, since there were not enough applicants in this field. n is the total amount of applications that were taken into consideration for calculations. Succes rates of women and men are also given, along with total shares of women amongst applicants and amongst panel members

Scientific subfield	n	Success rates					
		Share of women among applicants (%)	Share of women among panel members (%)	Women (%)	Men (%)	P value (Fisher's exact test)	Significance summary
Chemistry, Materials	70	27.1	18.7	31.6	47.1	0.2873	
Mathematics, Physics	65	15.4	12.0	40.0	58.2	0.3211	
Engineering	48	14.6	5.3	42.9	48.8	>0.9999	
Earth sciences, Biology, Agronomy	**120**	**35.0**	**17.3**	**31.0**	**52.6**	**0.0338**	*
Sociology, Anthropology, Political sciences, Communication	38	28.9	18.7	36.4	55.6	0.4756	
Psychology, Education Sciences	35	28.6	32.0	30.0	56.0	0.2642	
Literature, Arts, Linguistics	17	41.2	36.0	28.6	40.0	>0.9999	
History, Archaeology	**23**	**65.2**	**26.7**	**40.0**	**100.0**	**0.0072**	**
Economics, Law	30	26.7	12.0	37.5	54.5	0.6817	
Genetics, Biochemistry, Cell biology	147	34.7	14.7	43.1	45.8	0.8619	
Physiology, Microbiology, Immunology, Veterinary medicine	146	41.8	13.3	37.7	49.4	0.1796	
Neurosciences	46	34.8	13.3	37.5	46.7	0.7557	
Medicine, Pharmacy, Dentistry	102	32.4	15.1	48.5	39.1	0.3981	
Sustainable development	3	66.7	22.7	0.0	100.0	/	/

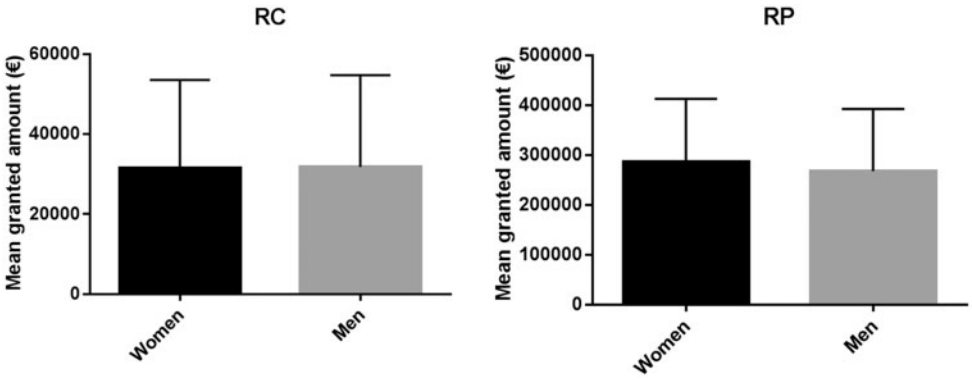


Figure 5. Mean granted amounts (€) to female and male beneficiaries of grants from F.R.S.-FNRS (2011–5). Data are presented as means ± SDs. RC: n = 409 (111 women and 298 men); RP: n = 444 (114 women and 330 men).

the success probability variable was statistically linked to the gender variable, even if the mean success rates of men and women over 5 years did not significantly differed. This indicates that even if mean success rates are statistically comparable, the probability of success variable is linked to the gender of the applicant. In our logistic regression model, we could identify a statistically significant interaction between gender and funding scheme variables at least in the frame of RC applications, for which success rates are significantly higher than for other funding schemes. This indicates that being a male applicant increases significantly more the probability of being funded in the frame of the RC funding scheme. It is interesting to note that RC is a 'grant' funding scheme and that in the context of the ex-ante evaluation of such applications, the quality of the applicant's CV represents a less important and weighted criteria than for fellowships where the applicant's CV is often considered as more important.

The statistical dependence between gender and success variables was demonstrated to be true only in ENS and HSS, and particularly in 'Earth Sciences, Biology, Agronomy' and in 'History, Archaeology'. This is of interest, since gender bias is generally thought to mainly affect specific fields such as STEM where women would be considered as less competent than men (Leslie et al. 2015). Our results do not confirm these assumptions, or only very partially. The fact that a statistically significant dependence does exist between gender and success variables in these two subfields in the context of RC does not necessarily mean there is a causative link between gender and success. It could also be due to the fact that for some reason, RC applications from women in these two subfields were of less good quality. Statistical analyses of performance are difficult to achieve, since for one of the subfields (History, Archaeology), bibliometric indicators are not relevant, whereas for the second one (Earth sciences, Biology, Agronomy), differences between subfields would make statistical analyses fairly unreliable. In addition, it is important to highlight the fact that among all applications that were taken into account in the analyses (6,393 in total), only 143 (2.2%) were concerned by a statistical link between gender and success variables, which is a very small proportion of the total.

Our results show that this statistical link is associated with lower mean grades awarded by remote reviewers to women as compared to men, since our analyses demonstrated that in the frame of RC applications, female applicants were awarded significantly lower grades than male applicants. This is of interest, since a bias in the peer-review procedure could have caused the observed link between gender and success variables. Peer review was indeed often accused by many authors of being biased (Demarest, Freeman and Sugimoto 2014; Kaatz, Gutierrez and Carnes 2014; Walker et al. 2015), even leading to initiatives such as testing double-blind peer review in some journals (Nature Medicine Editorial 2015). However, it is true that only the grades awarded by reviewers were analysed and not the comments they provided, which could possibly contain scientifically unjustified negative remarks and grades for female applicants as compared to male applicants, but observed differences could also be due to the fact that in the frame of this particular funding scheme (890 applications of the 6,393 analysed), applications by female researchers were of a slightly lower quality due to random reasons. A future step could be to identify the reason(s) that explain these differences. When RC applications are broken down in scientific fields, average grades given by remote reviewers significantly differ in only two scientific fields, namely, 'Sociology, Anthropology, Political sciences, Communication' and 'History and Archaeology'. As we could observe a statistical link between gender and success regarding RC applications only in 'History, Archaeology' and 'Earth sciences, Biology, Agronomy', this means that the observed statistical link between gender and success is not fully explained by significant differences in terms of grades given by remote reviewers. It also means that a significant difference between female and male applications in terms of average grades given by remote reviewers does not necessarily lead to a statistical dependence between success and gender variables in a particular subfield (case of Sociology, Anthropology, Political Sciences, Communication). Taken together, these results indicate that the observed statistical dependence between gender and success regarding RC in some scientific fields could only be partly explained by differences in terms of grades given by remote reviewers.

As regards grants, proportions of male researchers who apply are systematically higher than proportions of female researchers. This could be easily explained by the fact that only tenure researchers and academics from WBF universities are allowed to apply for such grants. Indeed, in WBF, the proportion of male researchers amongst the academic personnel of the universities is much higher than the proportion of women. For instance in 2014, statistics from the Conseil des Recteurs des Universités francophones show that for all WBF universities, the total amount of FTE among the academic personnel was 568.62 for women and 1,575.03 for men, with 2,143.64. This means female researchers only represent a proportion of 26.5% of the academic staff of WBF universities. Among the F.R.S.-FNRS tenure researchers, women represent a proportion of 32.7% (as in 2016). Until now we have no formal explanation as regards the observed statistical link between gender and success variables when it comes to RC, but the fact that amongst grant applicants and academic personnel in WBF, the proportion of men remains much higher than the proportion of women is an interesting lead to be kept in mind and further investigated, in addition to the higher grades awarded by remote reviewers to male applicants as compared to female.

The mean amounts of money granted to applicants who are awarded a grant by F.R.S.-FNRS do not statistically differ depending on the gender of the applicants. This is of interest, since it is in contradiction with results obtained in the context of other analyses, for example carried out using data from Wellcome Trust applications (Bedi, Van Dam and Munafo 2012).

One could argue that a potential weakness of this study is that it does not include any performance measurement of female applicants in comparison with male applicants. However, since almost no statistically significant difference could be found in the analyses we carried out, performance analyses would have been of limited interest here. In addition, we wanted this study to be broad and not to focus on specific scientific fields, and performance analysis across different scientific fields is often difficult or irrelevant.

The issue of imposing quotas of the under-represented gender amongst reviewers and in particular amongst panel members who are in charge of the final evaluation of proposals is regularly raised, mainly because people might think it could solve potential gender issues arising in the evaluation processes. Although carefully considering this question and paying close attention to the representation of women among its evaluation bodies, F.R.S.-FNRS has never applied any formal quota amongst reviewers or panel members for several reasons, including the difficulty of finding enough women scientists in some scientific fields and the artificial aspect of imposing quotas. Moreover, the question of the efficiency of quotas was already studied, and it was demonstrated that in the case of European Research Council (ERC), for instance, there was no correlation between the success rates of female applicants and the gender balance of evaluation panels. Findings from this author were published in *Nature* and suggest that a quota system for staffing evaluation panels will not lead to more grants for women. She even mentions that 'worse, quotas would place greater demand on the small pool of female scientists who would serve on these panels—possibly enough to hamper their career progress' (Vernos 2013). Our results tend to confirm her assumptions. Indeed, with no mandatory quotas imposed and a clear imbalance amongst panel members (18.4% of women and 81.6% of men in total during the 5 years considered in this study), we observed no difference in terms of success rates of women in comparison with men. In our opinion, good practices such as family-friendly policies, the close monitoring of gender bias in success rates and the representation of women in evaluation bodies, transparency in evaluation processes, and so on, could represent an alternative to formally imposed quotas. As a consequence of this study, we will henceforth monitor the evaluation processes of RC applications more closely, especially the remote peer-reviewing step and in subfields where a statistical correlation between gender and success has been demonstrated.

Acknowledgements

The authors would like to thank Fonds de la Recherche Scientifique-FNRS for providing data and support. They also thank Catherine Hall for English proofreading of the manuscript.

Funding

This work was supported by the Fonds de la Recherche Scientifique-FNRS.

Conflict of interest statement. None declared.

References

Albers, C. J. (2015) 'Dutch Research Funding, Gender Bias, and Simpson's Paradox', *Proceedings of the National Academy of Sciences of the United States of America*, 112/50: E6828–9.

Bedi, G., Van Dam, N. T., and Munafo, M. (2012) 'Gender Inequality in Awarded Research Grants', *The Lancet*, 380/9840: 474.

Boyle, P. J. et al. (2015) 'Gender Balance: Women are Funded More Fairly in Social Science', *Nature*, 525/7568: 181–3.

Ceci, S. J. et al. (2014) 'Women in Academic Science: A Changing Landscape', *Psychological Science in the Public Interest, Supplement*, 15/3: 75–141.

Demarest, B., Freeman, G., and Sugimoto, C. R. (2014) 'The Reviewer in the Mirror: Examining Gendered and Ethnicized Notions of Reciprocity in Peer Review', *Scientometrics*, 101/1: 717–35.

Handley, I. M. et al. (2015) 'Quality of Evidence Revealing Subtle Gender Biases in Science is in the Eye of the Beholder', *Proceedings of the National Academy of Sciences of the United States of America*, 112/43: 13201–6.

Kaatz, A., Gutierrez, B., and Carnes, M. (2014) 'Threats to Objectivity in Peer Review: The Case of Gender', *Trends in Pharmacological Sciences*, 35/8: 371–3.

Leslie, S. J. et al. (2015) 'Expectations of Brilliance Underlie Gender Distributions Across Academic Disciplines', *Science*, 347/6219: 262–5.

Marsh, H. W. et al. (2009) 'Gender Effects in the Peer Reviews of Grant Proposals: A Comprehensive Meta-Analysis Comparing Traditional and Multilevel Approaches', *Review of Educational Research*, 79/3: 1290–326.

Moss-Racusin, C. A. et al. (2012) 'Science Faculty's Subtle Gender Biases Favor Male Students', *Proceedings of the National Academy of Sciences of the United States of America*, 109/41: 16474–9.

Nature Editorial. (2015) 'Sexism has No Place in Science', *Nature*, 522/7556: 255.

Nature Medicine Editorial. (2015) 'A Double-Blind Option for Peer Review', *Nature Medicine*, 21/3: 199.

Pohlhaus, J. R. et al. (2011) 'Sex Differences in Application, Success, and Funding Rates for NIH Extramural Programs', *Academic Medicine*, 86/6: 759–67.

Shen, H. (2013) 'Inequality Quantified: Mind the Gender Gap', *Nature*, 495/7439: 22–4.

Simpson, E. H. (1951) 'The Interpretation of Interaction in Contingency Tables', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 13/2: 238–41.

van Arensbergen, P., van der Weijden, I., and van den Besselaar, P. (2012) 'Gender Differences in Scientific Productivity: A Persisting Phenomenon?', *Scientometrics*, 93/3: 857–68.

van den Besselaar, P., and Sandström, U. (2016) 'Gender Differences in Research Performance and Its Impact on Careers: A Longitudinal Case Study', *Scientometrics*, 106/1: 143–62.

van den Brink, M., Brouns, M., and Waslander, S. (2006) 'Does Excellence have a Gender? A National Research Study on Recruitment and Selection Procedures for Professorial Appointments in The Netherlands', *Employee Relations*, 28/6: 523–39.

van der Lee, R., and Ellemers, N. (2015) 'Gender Contributes to Personal Research Funding Success in The Netherlands', *Proceedings of the National Academy of Sciences of the United States of America*, 112/40: 12349–53.

Vernos, I. (2013) 'Research Management: Quotas are Questionable', *Nature*, 494/7439: 39.

Volker, B., and Steenbeek, W. (2015) 'No Evidence that Gender Contributes to Personal Research Funding Success in The Netherlands: A Reaction to van der Lee and Ellemers', *Proceedings of the National Academy of Sciences of the United States of America*, 112/51: E7036–7.

Walker, R. et al. (2015) 'Personal Attributes of Authors and Reviewers, Social Bias and the Outcomes of Peer Review: A Case Study', *F1000Res*, 4: 21.

Watson, D., and Hjorth, J. (2015) 'Denmark: Women's Grants Lost in Inequality Ocean', *Nature*, 519/7542: 158.

Wenneras, C., and Wold, A. (1997) 'Nepotism and Sexism in Peer-Review', *Nature*, 387/6631: 341–3.