

forskningspolitik.se

Forskning om Forskning

nr 3 2009

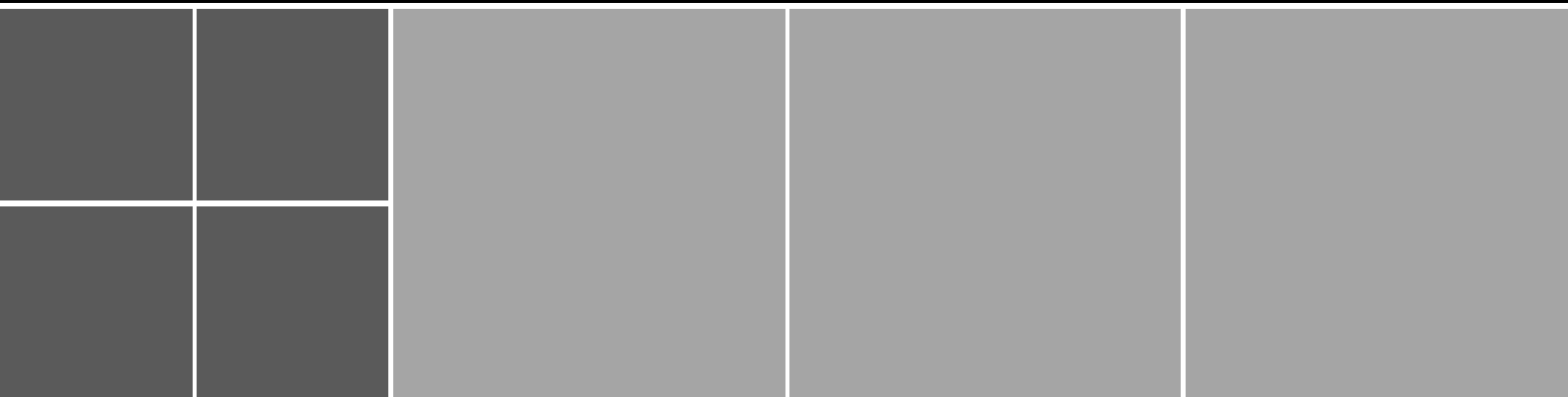
(2009-07-30)

Ulf Sandström

Linköpings universitet

Visiting scholar at Aalto University, Helsinki Finland

Bidrag till debatten om nepotism:
Särtryck från Läkartidningen och
Scientometrics



Fördelningen av medicinska forskningsanslag:

Män utan kontakter missgynnas

Vetenskapsrådets medicinska ämnesråd har kommit tillrätta med diskrimineringen av kvinnor när det gäller ansökningar om forskningsanslag, men knappast med »nepotismen«. En genomgång av 2004 års ansökningar visar att män utan rätta kontakter i viss utsträckning blir diskriminerade. Det är dags att lyfta fram jävsproblematiken i den forskningspolitiska debatten.

För ett år sedan fastslog en studie att forskningsmedel inom Forskningsrådet för arbetsliv och socialvetenskap (FAS) oftare gick till män än till kvinnor [1]. Vetenskapsrådet har i flera rapporter försökt att hantera frågan huruvida kvinnor och män behandlas likvärdigt av bedömnings-systemet [2, 3]. Vi menar att det finns anledning att ifrågasätta om analyserna genomförts med tillräckligt bra metoder och med rätt fokus.

Studier av hur resurser och positioner fördelas inom medicinsk forskning är relativt sällsynta och får sällan avgörande betydelse. Undantaget är en studie av mikrobiologerna Susanne Wenerås och Agnes Wold (härefter W & W) [4]. Den visade att kvinnor diskriminerades vid bedömningen av forskarassistenttjänster.

Författarna antydde dessutom att sökande med kontakter inom dåvarande Medicinska forskningsrådets (MFR) beredningsgrupper var oproportionerligt gynnade, något de benämnde »nepotism«. Dessa avslöjanden, som av MFR bemöttes med rökråd och av-värjningsmanövrar, fick omfattande

återverknings på diskussionen om kvinnor och forskning.

W & W kunde med hjälp av offentlig-hetsprincipen och en genomtänkt forskningsdesign dokumentera nepotismen och sexismen. De använde bästa tillgängliga metoder för att förklara bedömningsapparaters olika skevheter. Detsamma kan tyvärr inte sägas om flertalet av de analyser som följt i deras fotspår. Trots att det numera finns förbättrade analysmetoder tillämpas de inte av områdets utredare och »utvärderare«.

Senare studier har förvisso ett informationsvärde, men det är synnerligen begränsat till följd av metodernas svagheter. FAS och Vetenskapsrådet (VR) har kontrollerat för skillnader mellan män och kvinnor med avseende på status och erfarenhet, men för att uppnå de kvaliteter som W & W:s studie uppvisade krävs mer av forskningsdesign och analysmetod.

Förutom status måste hänsyn tas till faktiska akademiska prestationer, dvs kvantitet och kvalitet på utförd forskning. Forskarens kompetens är en central dimension i beredningsgruppernas bedömning av en ansökan. FAS undersökning [1] hade ambitionen att visa hur det ser ut tio år efter W & W, men man förmådde inte besvara frågan eftersom studien var feldesignad och helt ignorerade prestationer. Detsamma gäller VR:s utredningar.

Ansatsen som tillämpades av W & W var att hitta bibliometriska indikatorer på de sökandes kompetens (artiklar och citeringar). Utan kompetensaspekten



Foto: Tracy Kahn/Scanpix

Ansökningar till Vetenskapsrådet bedöms mera gynnsamt när det föreligger jäv, »familjära relationer«, mellan sökande och bedömare. Sökande utan sådana relationer får 15 procent lägre betyg.

riskerar sambanden mellan betyg och kön att bero på gruppernas genomsnittligt skilda prestationer. Sprängkraften i W & W:s arbete var att en kvinnlig sökande behövde tjugo gånger fler publiceringspoäng för att uppnå samma kompetensbetyg som sina manliga kolleger.

Med målet att följa upp och replikera

W&W:s studie samt att inkorporera de senaste landvinningarna inom det bibliometriska forskningsfältet har vi genomfört en studie av ansökningar till VR:s medicinska ämnesråd 2004 [5]. Som kontrollvariabler användes ett batteri av bibliometriska state-of-the-art-indikatorer (fältnormaliserad citeringsgrad, impact och antal publikationer).

Vår undersökning visar att det i kommittéernas bedömning av ansökningar inte längre finns signifikanta skillnader mellan män och kvinnor. Däremot finns tydliga effekter av nepotism, definierat som att någon i beredningsgruppen anmält jäv i relation till sökande. Data för denna variabel hämtas från ämnesrådets protokoll. Vi upptäckte även en intressant interaktionseffekt mellan kön och nepotism: Män och kvinnor med kontakter i beredningsgruppen samt kvinnor utan kontakter i



MARTIN HÄLLSTEN

fil kand, doktorand, Institutet för social forskning, Stockholms universitet
martin.hallsten@sofi.su.se



ULF SANDSTRÖM

docent, universitetslektor, Linköpings universitet
ulf.sandstrom@liu.se

»Räcker det att ledamöter med jäv lämnar rummet, eller stannar deras 'ande' kvar i rummet?«

beredningsgruppen fick högre betyg än män utan kontakter. Skillnaderna är betydande: om man räknar lägsta tilldelade betyget som 0 procent och högsta betyget som 100 procent får män och kvinnor med jävkontakt ca 15 procentenheter högre betyg än män utan kontakter.

Således blir, allt annat lika, män som saknar de rätta kontakterna i viss utsträckning missgynnade. De får relativt sett lägre betyg på sina ansökningar och bör ha sämre möjligheter att få bidrag när konkurrensen är stenhård.

Ämnesrådet för medicin har kommit tillrätta med diskrimineringen av kvinnor, men knappast med »nepotismen«. Varken FAS eller VR har ens närmat sig frågan om jävsrelationernas betydelse trots att detta är en synnerligen vital fråga i den forskningsetiska debatten. Räcker det att ledamöter med jäv lämnar rummet, eller stannar deras »ande« kvar i rummet?

Självklart finns andra och kanske effektivare modeller för att behandla ansökningar. Amerikanska National Science Foundation arbetar med starka »program officers« och externa bedömare. Möjligen skulle ett sådant system lösa den ständigt närvarande jävsproblematiken i svensk forskning. Under alla omständigheter är det dags att föra upp detta på dagordningen för den forskningspolitiska debatten.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

REFERENSER

1. Norström T. Forskningsanslag beviljades oftare för män än för kvinnor. *Läkartidningen*. 2007; 104:3273-5.
2. Glynn C, Hyenstrand P, Jacobsson C, Larsson M, Lundberg E, Wadskog D. Vetenskapsrådet och jämställdheten. Stockholm: Vetenskapsrådet; 2006.
3. Jacobsson C, Lundberg E. Jämställdheten i Vetenskapsrådets forskningstöd 2003-2007. Stockholm: Vetenskapsrådet; 2008.
4. Wennerås C, Wold A. Nepotism and sexism in peer-review. *Nature*. 1997;387(6631):341-3.
5. Sandström U, Hällsten M. Persistent nepotism in peer-review. *Scientometrics*. 2008;74(2):175-89. <http://forskningspolitik.se/DataFile.asp?FileID=150>

Fler debattinlägg på sidan 1912.

APROPÅ! Ett HSAN-ärende

Tur att Förnya-knappen finns ...

Redovisningen av fallet med fenytoin-överdosering, vilket beskrevs i *Läkartidningen* 24-25/2009 (sidan 1664) gav mig en tankeställare. Efter ett långt yrkesliv innanför sjukhusets skyddande murar har jag, efter pensioneringen, kastat mig ut i verkligheten och tjänstgör nu som vikarierande distriktsläkare (ej hyrläkare).

Vanligt förekommande på min mottagning är patienten, ofta till åren kommen, som kommer för sin regelbundna årskontroll. Vederbörande behandlas oftast för ett flertal åkommor såsom högt blodtryck, höga blodfetter, kärlkramp, andningsproblem, diabe-

tes, sömnproblem, depression etc. Patienten behandlas som regel med ett flertal läkemedel.

Vid korrekt handläggning borde jag nu, utifrån journaluppgifter, anamnes och klinisk undersökning, värdera om indikation fortfarande finns för vart och ett av de aktuella läkemedlen. Vidare borde jag, utifrån klinik och aktuella data för lever- och njurfunktion, värdera om doseringen av läkemedlen är korrekt. Åldersbetingade förändringar kan ha påverkat såväl farmakodynamik som farmakokinetik och därmed utgöra skäl för dosändring. Om något av läkemedlen sponeras eller om något läggs till, bör man bedö-

ma om detta påverkar effekt och metabolism av de övriga.

Hur ser då verkligheten ut? Efter journalläsning, samtal och klinisk undersökning har de tillmätta 30 minuterna för länge sedan runnit ut, och nästa patient väntar troligen otåligt på sin tur. Ändå återstår journalföring. Vad gör man då? Rådningen är den ack så bekväma *Förnya*-knappen i läkemedelsmodulen.

Kanske är det min tur att nästa gång sitta på de anklagades bänk.

Jan Eric Wirén

överläkare emeritus; vikarierande distriktsläkare, Vårdcentralen Husie, Malmö
jan.eric.wiren@telia.com

APROPÅ! Läkarens möte med patienten

Vart är vi på väg?

Då jag läser senaste numret av *Läkartidningen* väcks återigen tanken. I min vardag ser jag fler och fler exempel på hur patienter och läkare flyttas som pusselbitar mellan sjukhus och avdelningar. Gemensamma väntelistor, dålig kontinuitet då läkare tvingas att växla mellan arbetsplatser, utlokalisering av patienter till »fel« avdelningar, missar i information mellan vårdgivare då patienter flyttas runt, är saker som vi förväntas acceptera.

Två exempel från LT

26/2009: I en artikel om kirurgi på äldre läser jag: »Eftersom ortopedan-

ibland inte ser patienten efter operationen, får man lätt en glättad bild av verkligheten.« I ett referat av ett ansvarsärende läser jag: »När det var dags för operation träffade den opererande läkaren patienten först då hon redan var sövd.«

För mig är detta tydliga

tecken på hur strukturen i vården ökar riskerna för felbehandling och sänker kvaliteten på omhändertagandet. Visst finns det inte sällan goda skäl till att en operatör inte träffar sin patient pre- eller postoperativt, men att det kan vara en norm är för mig obegripligt.

Noggrann journalanteckning, sidomarkering

och preoperativ »time-out« (som en extra kontroll) är givetvis bra, men en personlig kontakt med patienten är att sätta ribben högre (där den ska vara).

För handen på hjärtat –

vem av oss som arbetar i vården skulle acceptera en elektiv operation utan att träffa kirurgen först?

Läkarens möte med patienten måste prioriteras bättre; vi ska inte stillatigande acceptera att förvandlas till gäster på våra avdelningar och mottagningar.

Anders Tuneskog

narkosläkare, Ljungby
anders.tuneskog@ltkronoberg.se

Persistent nepotism in peer-review

ULF SANDSTRÖM,^a MARTIN HÄLLSTEN^b

^a Department for Studies of Social Change and Culture, Linköping University, Linköping (Sweden)

^b Department of Sociology & Swedish Institute for Social Research, Stockholm University,
Stockholm (Sweden)

In a replication of the high-profile contribution by Wennerås and Wold on grant peer-review, we investigate new applications processed by the medical research council in Sweden. Introducing a normalisation method for ranking applications that takes into account the differences between committees, we also use a normalisation of bibliometric measures by field. Finally, we perform a regression analysis with interaction effects. Our results indicate that female principal investigators (PIs) receive a bonus of 10% on scores, in relation to their male colleagues. However, male and female PIs having a reviewer affiliation collect an even higher bonus, approximately 15%. Nepotism seems to be a persistent problem in the Swedish grant peer review system.

Introduction

Gender bias in the grant peer-review procedures of research councils is a crucial issue that generates serious discussion in core scientific journals [WESSELEY, 1998; BORNMANN & DANIEL, 2005]. Any type of bias would be detrimental towards strategies for scientific excellence. The proposed European Research Council is but one illustration of the importance of this debate. Since WENNERÅS & WOLD [1997]

Received February 12, 2007

Address for correspondence:

ULF SANDSTRÖM

Department for Studies of Social Change and Culture, Linköping University

Linköping, Sweden

E-mail: ulfsa@tema.liu.se

0138–9130/US \$ 20.00

Copyright © 2007 Akadémiai Kiadó, Budapest

All rights reserved

published the first ever analysis of peer-review scores for postdoctoral fellowship applications no studies have been able to use the same variables, including reviewer affiliation.

This study aims to trace the influence of gender and conflict-of-interest on scores, controlling for performance measures (bibliometrics), academic status (professor, assistant professor, and researcher), experience (years since dissertation), faculty discipline (medicine or not), university affiliation and committee assignment.

Research councils and peer review

An essential principle of the classical model for research councils is that scientists, unrestricted by and unaffiliated with external factors, should govern research. In Sweden, active researchers constitute a majority on the council's board, i.e. in all bodies where applications are assessed and evaluated and grants decided upon. Committee members are chosen through an electoral process involving all members of the research community in the university system. Peer review is used to guarantee the quality and diversity of basic research. A specific feature of the Swedish councils is that scientific committees are in command of the work. As a consequence, and in comparison to the situation at NSF and NIH (main U.S. financing bodies), programme officers in Sweden have little or no influence over granting procedures, initiatives and priorities (see [SANDSTRÖM & AL., 1997]).

Peer review is a documented, critical review performed by peers (persons having technical expertise in the subject matter to be reviewed) who are independent of the work to be reviewed. It is a tool that funding agencies (and programme managers) can use to obtain high-quality technical and scientific input for decisions on allocating resources. Peer review is intended to improve both the technical quality of projects in R&D and the credibility of the decision-making process. The independence of peers should make them more effective than internal reviewers.

Generally, the procedures among the councils in Sweden were similar during the 1990s. Each applicant (principal investigator) submitted a CV, bibliography and a research proposal. Applications were reviewed by one of 10 to 12 committees, each covering a specified research field or discipline. Each application was rated by several reviewers in the committee to which it was assigned. Reviewers graded the proposals and "track records" with a score, and that score formed the basis for the ranking used in funding decisions. Accordingly, review committees set up a conflict-of-interest protocol stating which members of the committee that have an affiliation with applicants. Affiliation might be maternal, kinship or supervisor relations. These protocols are open for public access according to Swedish legislation.

The gender gap

During the first half of 1990s, success rates for applicants to the Swedish Medical Research Council (MFR) were marked by a significant gender gap. Women often had 10% lower rates than men did. A study published in 1995 by Wennerås and Wold (later a full study was published in *Nature* [WENNERÅS & WOLD, 1997]) disclosed a gender bias in the evaluation of merits for postdoctoral fellowships at the MFR. Thereafter, a shift in policy was implemented. Practices among most other research councils changed, and the gap narrowed significantly during the second half of the 1990s (Figure 1).

Wennerås and Wold showed that female productivity was credited lower than male productivity in the evaluation of applications for postdoctoral fellowships (N = 114) submitted to the Swedish MFR. The conclusion was clear: “Our study”, they said, “strongly suggests that peer reviewers cannot judge scientific merit independent of gender.” (1997, p. 341). Affiliation with a committee member further increased the chances of being funded. The disclosure of discriminating procedures and nepotism resulted in an intensive debate.

A detailed examination of male and female success rates within the research councils from 1989 to 2000 reveal that success rates for males were generally 5 to 10 percentage points higher than the success rates for females. This difference decreased after the 1995 turning point (Figure 1), which indicates a positive effect on female funding chances.

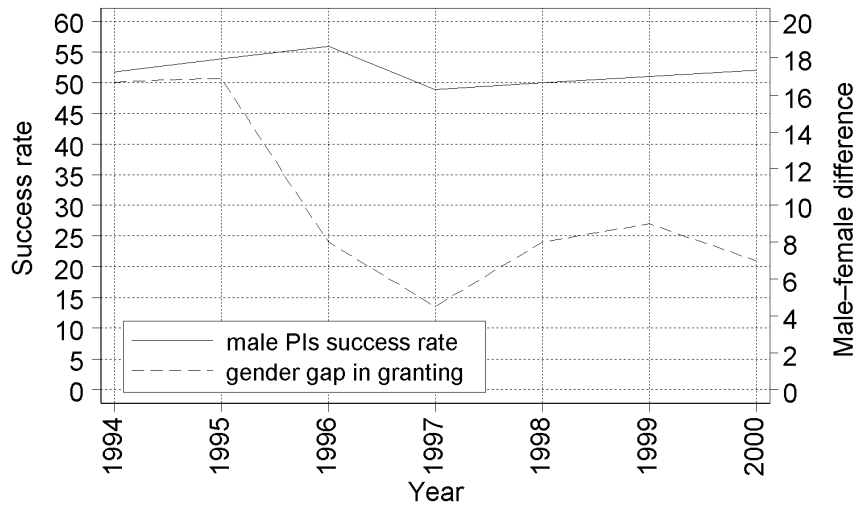


Figure 1. The Wold effect. Success rates of male applicants and the gender gap in the Medical Research Council 1994–2000

This change in policy from the MFR could be called “the Wold Effect” However, the success rates for male applicants remained higher even after 1995. Obviously, the differences are partially explained by male domination in upper academic positions. Hence, a positive difference in success rates between males and females is not by itself an indication of discriminating procedures.

According to WOLD & CHRAPKOWSKA [2004] policy changes already in 1997 were effective: “[I]ncreased knowledge on how prejudice influence peer decisions led to a reform of council procedure in 1997 so that it is now completely gender neutral” (our translation). We consider this as a controversial statement. A closer look at success rates for women indicate that structural factors for the group as distribution over committees and age/time elapsed since PhD etc. favour higher success rates, which leads toward closing the gender gap. Nevertheless, we will need a more comprehensive approach in order to fully understand the gender gap phenomena – our intention is to point at the nepotism factor which have not been investigated after the Wennerås & Wold article. Lack of data is a considerable problem for researching that question, but as already indicated we have reliable data produced by the scientific committees themselves.

Research on gender and Research Councils

Gender in science is becoming a rather well-researched question [WESSELY, 1998]. Productivity issues was one of the first to be discussed (the productivity puzzle) by COLE & ZUCKERMAN [1987], LONG [1992], KYVIK & TEIGEN [1996] and XIE & SHAUMAN [1998]. Further investigations have both challenged this discussion but also placed in wider context. There seem to be no significant differences when age, field and other variables are held constant; LEVIN & STEPHAN [1998], PRPIC [2002], ASMAR [1999], BLACK & HOLDEN [1998]. GANDER [1999] finds a higher marginal productivity among female researchers at liberal arts institutions. Results from a large sample of faculty from 57 universities in Netherlands indicated that factors affecting research productivity were identical for men and women, and that a family-related variable such as small children in the family, had no effects on research productivity. The territorial segregation issue, the gendered stratification of the scientific work force along disciplinary lines, has also raised interest. ROSSITER [1978] and KULIS & AL. [2002] are contributions to that literature. Women are overrepresented in fields like chemistry and microbiology, areas with high journal impact factors.

The paper by WENNERÅS & WOLD [1997] is a landmark in research on gender bias. Their article is one of the most frequently cited in this area of research (http://garfield.library.upenn.edu/histcomp/peer-review_to/list/au-lcs.html). It should be underlined that they showed that bias due to nepotism was as strong as the gender bias effect. Unfortunately, virtually none of the papers citing their article have attempted to replicate nepotism issues.

There are, however, a few exceptions to this observation: reports from funding agencies in the Netherlands and the UK have discussed fairness and funding in gender research from the perspective of modernising peer review. Other research projects have presented results from Denmark and the Netherlands [BROUNS, 2000]. However, none of these reports have been able to challenge the methodological excellence of the Wennerås & Wold paper. Especially, their combination of data extracted from applications (council database), bibliometric data (from CV and publication list) and conflict-of-interest (protocol from committees) are unusual, and a replication after ten years is warranted.

In a thorough analysis MOED [2005, CH. 20] shows that in a national research council (the council was held anonymous) proximity between applicant and expert committee was a strong determinant for priority rating of the application. Being member of the committee was considered as the highest level of proximity. Using a large data set, output and field normalised impact figures for 3,300 applicants and committee members were calculated. Proximity was indexed according to whether the applicant or co-applicants were members of the committee. The study is relevant in this context, although the Swedish case is somewhat different, as data on conflict of interest is available.

Confidence in the peer review system is essential to the academic community. The legitimacy of funding procedures is based largely on trust. Therefore, conclusions about the functions and performance of research councils should be scrutinised and well researched. The problem with several of the studies on gender bias in research councils is that they cover only small samples and short periods of time i.e. mostly cross-sectional studies. Although most of the work to date is careful and methodologically sound science, the lack of data is a general concern.

For the development of this area of research richness and quality of data is necessary. Generally, investigations on this subject have had problems in finding relevant data on the applicants and their research areas. Due to restrictions and economy of the research process, many of the studies have used only simple descriptors, e.g. age, time lapsed since dissertation, affiliation, publications, review grade, etc. as explanatory factors in relation to decisions taken by the councils. Now, we can add reliable data on reviewer affiliation, and advanced bibliometric analysis.

Table 1. Linear combinations of gender and conflict of interest coefficients

		male PI		female PI	
no association	<i>point estimate</i>	0.000		0.097	
	<i>confidence interval</i>	[-	-]	[0.027	0.167]
conflict of interest	<i>point estimate</i>	0.151		0.167	
	<i>confidence interval</i>	[0.032	0.271]	[0.051	0.282]

Source: Table 2, model 10.

Table 2. Linear regression models of normalised score. Sampling scheme adjusted estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
principal investigator is female	0.029 (0.023)	0.012 (0.020)	0.083 (0.044)	0.086* (0.041)	0.019 (0.014)	0.026 (0.013)	0.023 (0.014)	0.033* (0.013)	0.096** (0.036)	0.097** (0.036)
conflict of interest	0.058** (0.018)	0.061** (0.018)	0.048* (0.019)	0.052** (0.019)	0.097** (0.033)	0.007 (0.028)	0.183** (0.056)	0.091 (0.049)	0.077 (0.049)	0.151* (0.061)
quality (cpp/fcs)	0.011 (0.012)			-0.006 (0.008)	0.012 (0.012)			-0.006 (0.007)	-0.006 (0.008)	-0.006 (0.008)
number of publications		0.004** (0.001)		0.004** (0.001)		0.003** (0.001)		0.003** (0.001)	0.003** (0.001)	0.003** (0.001)
expected impact (jcs/fcs)			0.101** (0.032)	0.104** (0.032)			0.104** (0.023)	0.100** (0.024)	0.121** (0.030)	0.118** (0.031)
quality (cpp/fcs) * principal investigator is female	-0.010 (0.018)			-0.019 (0.015)				-0.014 (0.014)	-0.014 (0.014)	-0.011 (0.014)
number of publications * principal investigator is female		0.001 (0.001)		0.001 (0.001)				0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
expected impact (jcs/fcs) * principal investigator is female			-0.051 (0.037)	-0.043 (0.037)				-0.048 (0.032)	-0.048 (0.032)	-0.041 (0.032)
quality (cpp/fcs) * conflict of interest					-0.031 (0.023)			-0.023 (0.022)	-0.016 (0.022)	-0.011 (0.023)
number of publications * conflict of interest						0.003** (0.001)		0.003** (0.001)	0.004* (0.001)	0.003* (0.001)
expected impact (jcs/fcs) * conflict of interest							-0.112* (0.044)	-0.055 (0.043)	-0.052 (0.045)	-0.080 (0.046)
conflict of interest * principal investigator is female									-0.082* (0.037)	-0.082* (0.037)
Constant	1.120** (0.030)	1.102** (0.024)	0.996** (0.050)	0.966** (0.044)	1.117** (0.030)	1.105** (0.024)	0.994** (0.042)	0.974** (0.038)	0.950** (0.043)	0.953** (0.042)
Observations	280	280	280	280	280	280	280	280	280	280
R-squared	0.3763	0.4399	0.3985	0.4656	0.3783	0.4459	0.4046	0.4723	0.4767	0.4838

Standard errors in parentheses
 ** p<0.01, * p<0.05
 Note: all equations contain controls for academic status (3 categories), experience, medical faculty (3 categories), research committee (7 categories), and university affiliation (7 categories).

The 1990s were a dynamic period as regards Swedish research policy. New aspects to consider were introduced, e.g. gender, support of young researchers, interdisciplinary areas, strategic research etc. It was a period when research councils had to learn how to adjust to a changing environment. What has happened after ten years of time? Are female investigators still discriminated? Is reviewer affiliation no longer, as Wold has suggested, a disturbing factor in peer decisions on grant applications?

Data and methods

Our research draws on data from the sub-council for Medicine (a division of the Swedish Research Council) on the research grant applications submitted, reviewed and acted upon during 2004. A stratified sample was drawn from the 611 proposals for new projects, 63% of which had been submitted by male applicants. The Scientific Council for Medicine consists of eleven committees representing different medical disciplines or areas of research. Since female applicants were under-represented, and our primary interest is the effect of gender, we over-sampled women and adjusted our estimates accordingly. Our data include 280 applications, of which 118 are from female principal investigators (PIs). Beside the new projects a number renewals are reviewed each year. Renewals are projects that had funding the earlier two or three year period (see Figure 2).

Applicants submit an electronic research proposal to the Council. Each proposal is scored by 5 of the 6 reviewers on the committee to which the application is assigned. Reviewers assign a score between 1 and 7 on the following parameters: scientific competence, methodology and research question. Then an average for each parameter yields a final score on which the ranking is based. Policies of secrecy within the Scientific Council for Medicine classify “white sheets” from committee meetings as non-public documents. Consequently, we do not have the scores on scientific competence, the parameter used by Wennerås and Wold. Instead, we use the final ratings that are good estimates of competence scores (for 313 applications in year 2005, $r=0.90$).

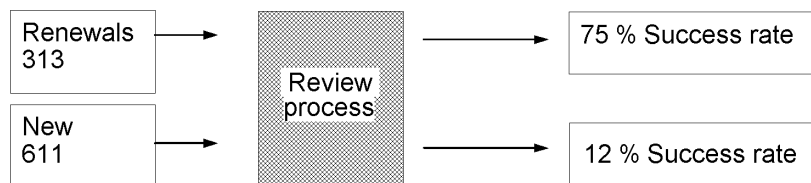


Figure 2. The review process of the Swedish Medical Research Sub-Council in 2004

Rating practises differ among committees. Our methodology to compensate for this is to normalise all scores within each committee in relation to the committee mean value. Wennerås and Wold were unable to normalise because they would have needed data on all projects to produce a committee average. However, as they had only data on applications for post-doctoral fellowships they could not create a normalised score per committee.

Figure 3 describes the distribution of grades for applications in the council's committees. It is evident that scoring practises differ across committees with respect to mean value, variance and overall shape of the distribution. A normalisation procedure is thus necessary to adequately compare applications from different committees. We normalise by dividing each score by its committee's mean value, hence the normalised grade range from 0.6–1.6. Note that the mean value is calculated for all applications in the committee, not only our sample. We will interpret the grade points as percentage points. The dependent variable (normalised grade) ranges between 0.66 and 1.67 (with mean 1.16 and standard deviation 0.178). Since the range is very close to 1, the coefficients*100 can be interpreted as percentage point changes.

The Swedish Research Council's conflict-of-interest policy specify that committees shall not review applications submitted by their own members, and that any such applications shall be re-assigned to other committees. Reviewers who declare a conflict of interest, normally due to supervisor, collaborative or kinship associations, are not permitted to score the proposal in question. Conflict-of-interest data are available by virtue of the principle of public access to official records.

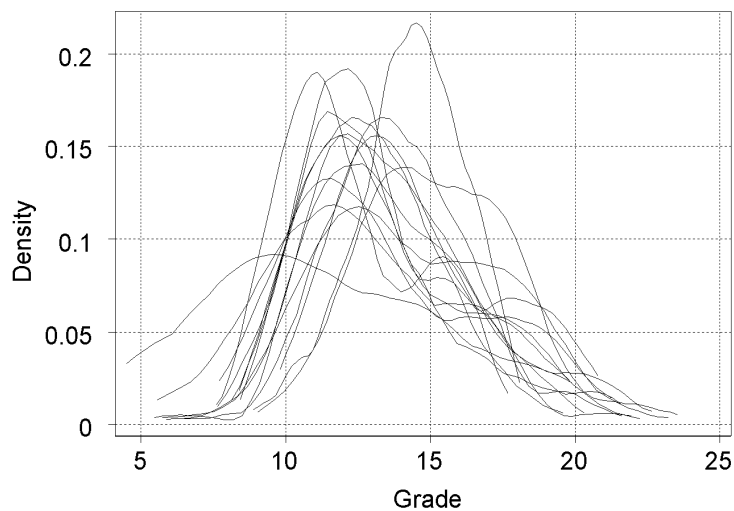


Figure 3. The distribution of grades in different committees prior to normalisation. Kernel density estimates

Full bibliometric information for publications from 1998 through 2004 was collected from the publication lists attached to the proposals. We double-checked against Internet Web of Science to have a secure basis for the bibliometric measures. To determine whether a certain volume of citations is high or low, it is necessary to measure against an international reference value. Journal citation score (JCS) is the average citation rate of all papers in journals where a researcher has published. Field citation score (FCS) is an average citation rate based on all journals within a specific subject code. These codes were produced by Thomson/ISI in order to classify journals. The JCS/FCS ratio gives the expected relative citation impact for journals, which is a more stable indicator of journal quality than the impact factor from Thomson/ISI. As an additional performance measure, we use number of papers (P).

The measured impact of a researcher, the number of citations per publication (CPP), concerns the actual citation rate. We apply standard bibliometric methods with field-normalised impact, e.g. CPP/FCS gives a relative figure of citation rates [VAN RAAN, 2006] Reference values, designed for each type of article (article, letter note or review) and for each year, is the basis for all subsequent ratios. For the analysis we use a two-three year citation window. We have not been able to correct for self-citations [GLÄNZEL & AL., 2006].

Having data on all journals in each category, we can determine whether citation performance is higher or lower than the field average. With the FCS indicator, we also have a more precise description of scientific field than Wennerås and Wold. They classified researchers in ten fields according to their main area of research. Web of Science classification of journals in which articles are published provides much more detailed information, and we can benefit from this field indicator.

Results: Nepotism in the medical sub-council

Are there systematic differences in ratings between men and women, and if so, which are the determining factors? First of all, we examine how grading procedures is related to gender of the applicant. To fully assess the impact of nepotism and gender on grading procedures, it is important to explore all interactions with productivity. Figures 4a–c show the relation between quintiles of productivity measures and awarded scores.

For citations per paper (Figure 4a), the awarded score increases with productivity for male PIs but not for female PIs. Instead, female PIs are rewarded the same score regardless of productivity.

In Figure 4b, male PIs seem to be reward higher grades than female PIs in the highest quintile of the impact score, but female PIs seem to have small advantage in the lower quintiles. In Figure 4c, awarded scores increase with number of publications to a similar extent for male and female PIs.

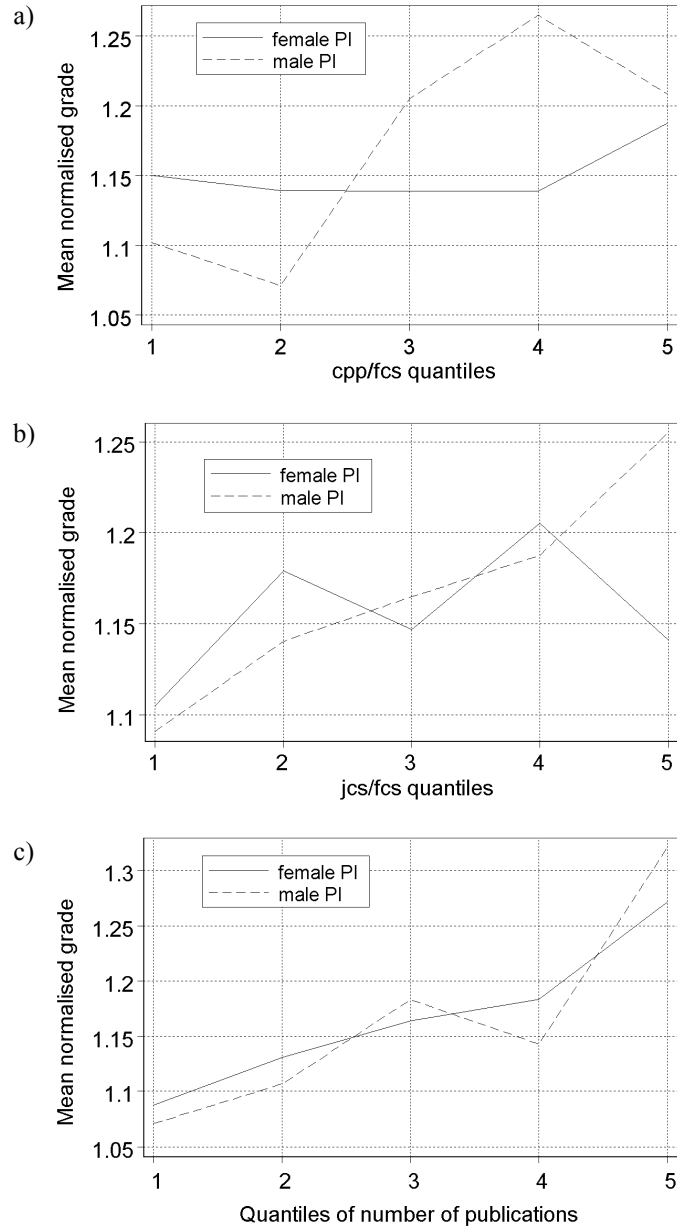


Figure 4. The association between grade and three different productivity measures in quantiles.
a) Citations per paper (CPP/JCS); b) Journal Impact (JCS/FCS); c) Number of publications
Separate estimates for male and female principal investigators

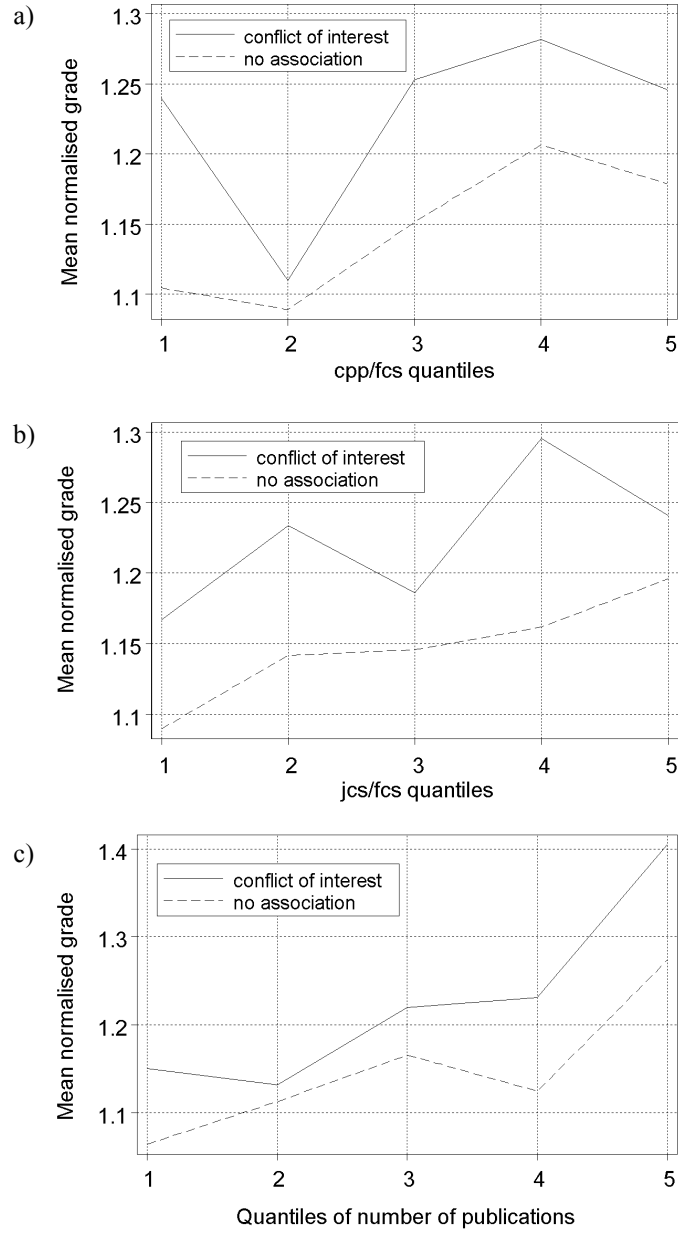


Figure 5. The association between grade and three different productivity measures in quantiles.
 a) Citations per paper (CPP/JCS); b) Journal Impact (JCS/FCS); c) Number of publications
 Separate estimates applicants with and without reviewer affiliation

Second, we examine how grading procedures is related to nepotism in Figures 5a–c. The pattern is very similar across productivity measures. Principal investigators with reviewer affiliation always receive a higher grade than applicants with no reviewer affiliation. There are no indications of interacting effects.

In the final step, we conduct regression of normalised scores analyses to elucidate the effect of nepotism and female gender. We specify models with control for individual characteristics and several productivity measures. All interaction effects between nepotism, gender and the productivity measures are explored (for descriptive statistics see Table 3 in Appendix). Since we have over-sampled female PIs, we use a weighted linear regression estimator. The results have been thoroughly checked and found robust. We conducted standard regression diagnostics (examination of residuals and influence statistics) and found no evidence that the results are due to the influence of outliers or other mis-specification problems. We also tested alternative specifications of the regression model (i.e. a piecewise linear spline regression model that may handle the nonlinearities of the productivity measures more adequately), but the same conclusions apply.

The explained variance range from 0.35 in the simplest specification to almost 0.5 with a full set of control variables and interactions, which we believe is quite high. In model 1–3 we examine each productivity variable and its interaction with gender separately. The effect of female PI is positive, albeit not significant, but the conflict of interest measure highly significant. It appears that nepotism is persistent in the practices of the research council. Impact score and number of publications influence awarded grades positively, whereas actual citations do not have any influence at all. Surprisingly, none of the productivity measures interact with gender: male and female PIs are judged similarly with reference to productivity. When we control for all productivity measures and interactions in model 4, the positive coefficient for female PIs becomes significant. It appears that female PIs receive a bonus compare to male PIs.

In model 5–7 we examine conflict of interest and its interaction with productivity variables separately. The effect of conflict of interest clearly varies by productivity measures. In model 4, there is no baseline effect of conflict of interest. However, for each publication PIs with research committee association receive a bonus of 0.4 percent. The conflict of interest effect is indifferent to actual citations (model 5). There is a quite strong interaction between impact score and conflict of interest. PIs with low impact have a very large nepotism bonus, whereas high impact PIs receive a lower bonus (it declines with eight percentage points for each normalised impact score). When we control for all productivity measures and interactions in model 8, only the interaction between conflict of interest and number of publications remain significant. Controlling for all interaction effects in model 9, the conclusions for model 4 and 8 remain.

However, the conflict of interest effect is not insignificant, but females receive bonus of around 9 percentage points (point estimate; it is at least 2.4 percentage points according to the confidence interval).

Finally, in model 10, we introduce an interaction term between gender and conflict of interest. The interaction term is significant and negative; the non-significance of the main effect in the previous model was due to a concealed interaction effect. Both the gender and the conflict of interest coefficient are positive and highly significant. We combine the coefficients in Table 1 for all combinations of gender and nepotism. Looking at the point estimates we find that female PIs with association are awarded the same grade as male PIs, whereas, female PIs without association are awarded a higher grade than male PIs. Taking the confidence intervals into consideration, the only significant difference is between males without reviewer affiliation and other applicants.

Discussion

There is a “productivism bias” in ordinary grant peer review. Our results show that expected impact score and number of publications influence awarded grades positively, whereas actual citations have no influence. The conflict of interest effect increases with number of publications. Males without reviewer affiliation are awarded lower scores than other applicants. The concluding words of WENNERÅS & WOLD [1997] still apply: “We see no reason why an applicant who manages to produce research of high quality despite not being affiliated with a prestigious research group should not be similarly rewarded.” Beyond this, we would conclude that the prestige of peer-review and the Scientific Council for Medicine is under threat if conflicts of interest overshadow the procedures. Still, after ten years, the system seems to be riddled with factors external to science. Nepotism matters.

*

Research was supported by the Bank of Sweden Tercentenary Foundation, grant no. 97-5085. Bibliometric data from the Swedish Research Council (VR) database. The authors thank Daniel Wadskog (VR) for bibliometric assistance. Data used for this article will soon be published at www.forskningspolitik.se/data.asp.

References

- BLACK, M. M., HOLDEN, E. W. (1998), The impact of gender on productivity and satisfaction among medical school psychologists, *Journal of Clinical Psychology in Medical Settings*, 5 (1) : 117–131.
- BORNMANN, L., DANIEL, H. D. (2005), Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees' decisions, *Scientometrics*, 63 (2) : 297–320.
- BROUNS, M. (2000), The gendered nature of assessment procedures on scientific research funding: the dutch case, *Higher Education in Europe*, 25 : 193–199.
- COLE, J. R., ZUCKERMAN, H. (1987), *Marriage, Motherhood and Research Performance in Science Scientific American*, 256 (2) : 119.
- GANDER, J. P. (1999), Faculty gender effects on academic research and teaching, *Research in Higher Education*, 40 (2) : 171–184.
- GLÄNZEL, W., DEBACKERE, K., THUIS, B., SCHUBERT, A. (2006), A concise review on the role of author self-citations in information science, bibliometrics and science policy, *Scientometrics*, 67 (2) : 263–277.
- KULIS, S., SICOTTE, D., COLLINS, S. (2002), More than a pipeline problem: Labor supply constraints and gender stratification across academic science disciplines, *Research in Higher Education*, 43 (6) : 657–691.
- KYVIK, S., TEIGEN, M. (1996), Child care, research collaboration, and gender differences in scientific productivity, *Science Technology & Human Values*, 21 (1) : 54–71.
- LEVIN, S. G., STEPHAN, P. E. (1998), Gender differences in the rewards to publishing in academe: Science in the 1970s, *Sex Roles*, 38 (11–12) : 1049–1064.
- LONG, J. S. (1992), Measures of sex differences in scientific productivity, *Social Forces*, 71 (1) : 159–178.
- MOED, H. F. (2005), *Citation Analysis in Research Evaluation*. Springer Verlag.
- PRPIC, K. (2002), Gender and productivity differentials in science, *Scientometrics*, 55 (1) : 27–58.
- ROSSITER, M. W. (1993), The Matilda Effect in science, *Social Studies of Science*, 23 : 325–341.
- SANDSTRÖM, U. & AL. (1997), "Does Peer Review Matter?" *Peers on Peers. Allocations Policy and Review Procedures at TFR*. Stockholm, Swedish Research Council for Engineering Sciences.
- VAN RAAN, A. F. J. (2006), Statistical properties of bibliometric indicators: Research group indicator distributions and correlations, *Journal of the American Society for Information Science and Technology*, 57 (3) : 408–430.
- WENNERÅS, C., WOLD, A. (1997), Nepotism and sexism in peer-review, *Nature*, 387 (6631) : 341–343.
- WESSELY, S. (1998), Peer review of grant applications: what do we know? *Lancet*, 352 (9124) : 301–305.
- WOLD, A., CHRAPKOWSKA, C. (2004), Förbjuden frukt på kunskapens träd. Atlantis. [in Swedish]
- XIE, Y., SHAUMAN, K. A. (1998), Sex differences in research productivity: New evidence about an old puzzle, *American Sociological Review*, 63 (6) : 847–870.

Appendix

Table 3. Descriptive statistics

Variable	observations	mean	stddev	min	max	percent	number
granted application	280			0	1	11.1	31
normalised grade score	280	1.163	0.178	0.669	1.676		
conflict of interest	280			0	1	21.4	59
principal investigator is female	280			0	1	42.1	118
quality (cpp/fcs)	280	1.203	0.878	0	10.899		
expected impact (jcs/fcs)	280	1.128	0.364	0.324	2.356		
number of publications	280	16.686	15.231	1	108		
assistant professor	280			0	1	35	97
professor	280			0	1	21.1	59
experience	280	11.821	8.145	0	41		
medical faculty	280			0	1	68.2	191
university: GU	280			0	1	11.8	32
university: LIU	280			0	1	9.3	25
university: LU	280			0	1	9.6	27
university: UMU	280			0	1	8.9	24
university: UU	280			0	1	7.5	21
university: other	280			0	1	13.2	37
committee: health and care	280			0	1	17.9	49
committee: clinical science	280			0	1	17.9	49
committee: medical chemistry	280			0	1	7.9	21
committee: microbiology	280			0	1	17.9	49
committee: psychiatry	280			0	1	8.9	24
committee: pharmacology	280			0	1	12.1	34