Selecting Scientific Excellence through Committee Peer Review – A Citation Analysis of Previous Publications by Successful and Non-successful Post-doctoral Research Fellowship Applicants

Lutz Bornmann[*] and Hans-Dieter Daniel[**]

[*]*bornmann@gess.ethz.ch*
Swiss Federal Institute of Technology Zurich (ETH Zurich)
Professorship for Social Psychology and Research on Higher Education
Zaehringerstr. 24, CH-8092 Zurich

[**]*daniel@gess.ethz.ch*
University of Zurich, Evaluation Office
ETH Zurich, Professor for Social Psychology and Research on Higher Education

Abstract
We investigated committee peer review for awarding long-term fellowships to post-graduate researchers as practiced by the Boehringer Ingelheim Fonds (B.I.F.) – a foundation for the promotion of basic research in biomedicine. Assessing the validity of selection decisions requires a generally accepted criterion for research impact. A widely used approach is to use citation counts as a proxy for the impact of scientific research. Therefore, a citation analysis for articles published previous to the applicants' approval or rejection for a B.I.F. fellowship was conducted. Based on our model estimation (negative binomial regression model), journal articles that had been published by applicants approved for a fellowship award (n=64) prior to applying for the B.I.F. fellowship award can be expected to have 37% (straight counts of citations) and 49% (complete counts of citations) more citations than articles that had been published by rejected applicants (n=333). Furthermore, comparison with international scientific reference values revealed (a) that articles published by successful and non-successful applicants are cited considerably more often than the "average" publication and (b) that excellent research performance can be expected more of successful than non-successful applicants. The findings confirm that the foundation is not only achieving its goal of selecting the best junior scientists for fellowship awards, but also successfully attracting highly talented young scientists to apply for B.I.F. fellowships.

Introduction
If originality is the motor of scientific progress, organised scepticism – systematically practiced by peer review – is its brake (Ziman, 2000). The task of peers asked to evaluate scientific work is to recommend only those that meet the highest of scientific standards. Peer review is the principal mechanism for quality control in federal funding of academic science in the United States, for example, with increasing usage through the general trend towards the "soft money" system (Guston, 2003). Although it is the best available mechanism (Kostoff, 1997), it is not perfect. Peers are not prophets, but ordinary human beings with their own opinions, strengths, and weaknesses (Ehses, 2004). Every scientific institution that uses peer review has to deal with the following question: Does the peer review system implemented by my institution fulfil its declared objective to select the best scientific work?

We investigated committee peer review for awarding long-term fellowships to post-graduate researchers as practiced by the Boehringer Ingelheim Fonds (B.I.F.), a foundation for the promotion of basic research in biomedicine (Bornmann & Daniel, 2004, 2005). The foundation invites highly talented young scientists in biomedicine to apply for long-term fellowship awards. What the Board of Trustees of the B.I.F. looks for most of all is excellence in scientific performance.[1] Young scientists that demonstrate scientific excellence are selected for the fellowships, providing that sufficient funds are available for all of them (Fröhlich, 2001). Fellowship applicants that do not meet the high standards are rejected. As there is broad support for citation counts of scientific articles as a measure

[1] The research award for post-doctoral fellows consists of a three-year fellowship which is renewable for a further three-year term. Applicants should not be older than 31 years. Their scientific achievements must be of outstanding quality, having resulted in papers in or accepted by leading international journals (Boehringer Ingelheim Fonds, 1999).

Proceedings of ISSI 2005, pages 343–351.
Edited by P. Ingwersen and B. Larsen.

of the impact of scientific research (Cole, 2000; van Raan, 2004), our assumption is that non-successful applicants earn lower citation counts than approved applicants because their scientific performance previous to applying for a fellowship is lower. We explored this hypothesis by determining the impact of scientific research of 397 post-doctoral applicants (64 approved and 333 rejected applicants) for B.I.F. fellowship awards between 1990 and 1995 and then comparing it with the decisions made by the B.I.F. Board of Trustees.

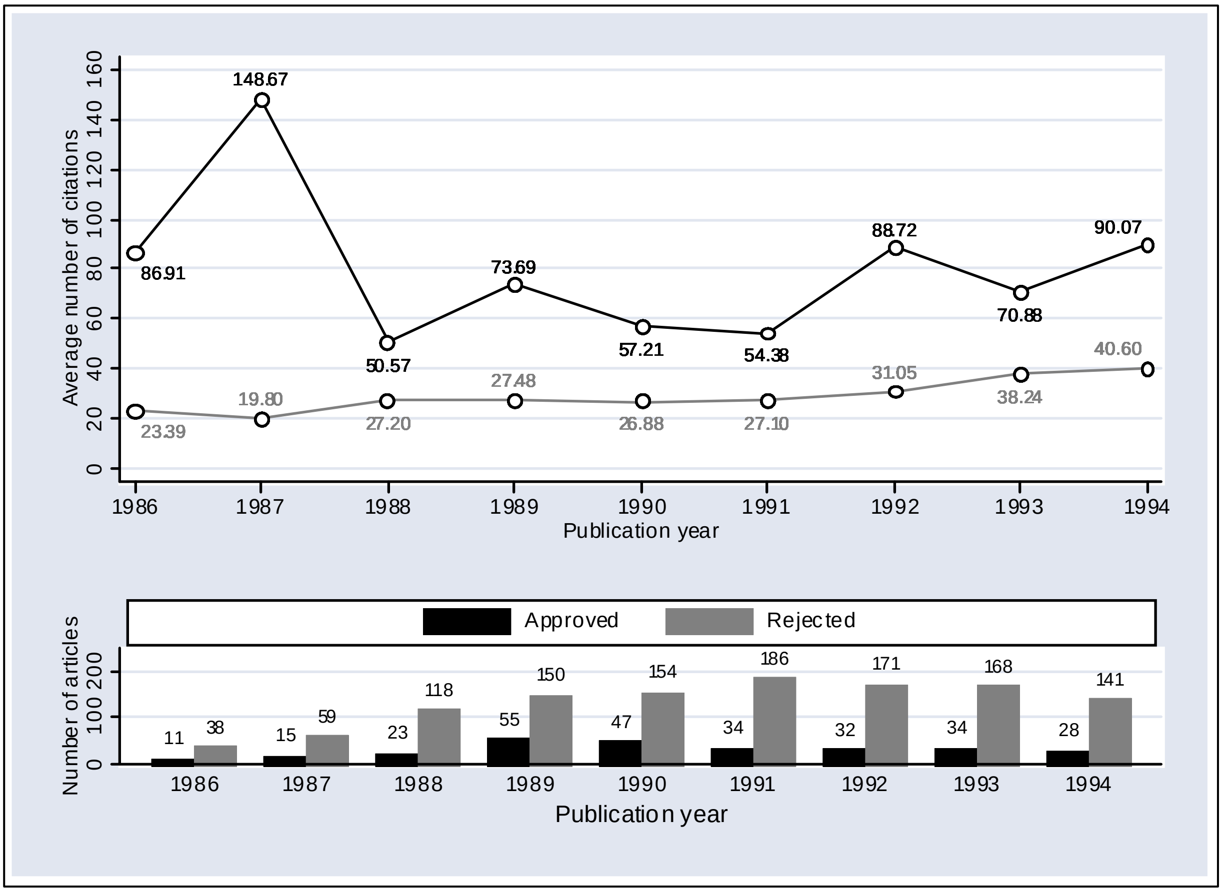
The data set on which the evaluation is based

All in all, 1,586 articles (full length articles, letters, notes, communications and reviews) had been published by 397 applicants previous to their applications to the B.I.F. (on average four articles). No articles had been published by 17 non-successful applicants.

Table 1. Journals in which ten or more articles from B.I.F. fellowship applicants had appeared (JCR impact factor in 1998, *n*=1,586)

Journal title	JCR impact factor in 1998	Number of articles
EMBO Journal	13.171	63
Journal of Biological Chemistry	7.199	52
FEBS Letters	3.581	33
Proceedings of the National Academy of Sciences USA	9.821	31
Nucleic Acids Research	4.878	29
Journal of Cell Biology	12.785	28
Biochemistry	4.628	27
Nature	28.833	25
Cell	38.686	22
Biochemical Journal	3.855	19
Brain Research	2.15	18
Neuroscience Letters	1.934	17
Molecular and cellular biology	9.571	17
Annals of the New York Academy of Sciences	0.959	15
European Journal of Biochemistry	3.249	14
Biochemical and Biophysical Research Communications	2.78	14
European Journal of Immunology	5.438	13
Development	9.712	13
Biological Chemistry	2.636	13
Virology	3.55	12
Journal of Molecular Biology	5.803	12
Journal of Immunology	7.166	12
Journal of Comparative Neurology	3.476	12
Experimental Cell Research	3.051	12
Biochimica et Biophysica Acta	2.478	12
Neurophysiology	0.114	11
American Journal of Physiology	3.077	11
Neuroscience	3.591	10
Journal of Cell Science	5.453	10
Human Genetics	2.826	10
Experimental Brain Research	2.018	10
European Journal of Cell Biology	2.485	10
Other journals (altogether 500 different journals, each with less than ten articles)		979

The bibliographic data of the articles were taken from the applicants' lists of publications, which were double-checked in the online databases BIOSIS Previews and Science Citation Index (SCI) (both databases provided by Thomson ISI, Philadelphia, Pennsylvania, USA), EMBASE (provided by Elsevier, Amsterdam, Netherlands) and Medline (provided by the National Library of Medicine, NLM, Bethesda, Maryland, USA). The vast majority of the applicants' articles had been published in scholarly refereed journals (95%, source: Ulrich's Periodicals Directory, 2004) in English (85%). The articles had been published between 1981 and 1996 in 532 different journals; in 32 journals, ten or more articles had appeared (Table 1). According to Thomson ISI, in the year 1998 the JCR (Journal Citation Report) impact factor of these journals (a measure of the frequency with which the "average article" in a journal has been cited in a particular year or period, revealing a journal's importance relative to others in its field) varied between 38.686 (*Cell*) and 0.114 (*Neurophysiology*).



Note.

1. Application window: 1990-1995; publication window: 1986-1994; citation window: from year of publication to the end of 2001.
2. We included only average numbers of citations per year that could be calculated with citation counts for more than ten articles (1986–1994); the years 1981 to 1985 and 1995 to 1996 could not be included.

Figure 1. (Top) Mean number of citations of articles previously published by approved and rejected B.I.F. post-doctoral applicants (*n*=397). (Bottom) Number of articles published in the year indicated.

Results

Did the committee peer review of the B.I.F. actually achieve its goal of selecting the "best" junior scientists showing the highest impact of scientific research? As shown in Figure 1 (top), the findings provide evidence that it did. The figure shows the annual mean number of citations of articles previously published by approved and rejected applicants up to the end of 2001. We determined the citation counts by using the online database SCI. For example, each of the 32 articles published in 1992 by approved applicants was cited on average 88.72 times up to the end of 2001, and each of the 171 articles published in 1992 by rejected applicants was cited on average 31.05 times. For **every** publication year, articles by the approved applicants were on average significantly more often cited than articles by the rejected applicants. The conspicuously high average citation count of articles published by approved applicants in 1987 (148.67) is due to four, highly frequently cited articles in

Cell (571 citations), *Development* (481 citations), *Nature* (311 citations), and *Proceedings of the National Academy of Sciences USA* (258 citations). Accordingly, Huber's M-estimator (Huber, 2003) – a robust alternative to the sample mean that is less sensitive to outliers – for articles published in 1987 yields a value for the centre of location of only 68.74 citations (but even this value is distinctly higher than the average citation count shown for rejected applicants in Figure 1).

While the average values in Figure 1 (top) suggest that B.I.F. committee peer review indeed selected "the better" junior scientists among the applicants, factors other than their scientific work could in principle have been responsible for the higher citation counts. Bibliometric studies have demonstrated that the following factors have a **general** influence on citation counts: number of co-authors (Beaver, 2004; Tregenza, 2002), the impact factor of the journals (Balaban, 1996; Tainer, 1991) and the size of the citation window (Daniel, 1993). In addition, in the calculation of average citation counts for research groups, a majority of the scientists enter into the statistical analysis with more than one publication, thereby violating the stochastic independence of the data (in our case, 76% of the fellowship applicants had published more than one article). By considering these factors in the statistical analysis, it becomes possible to establish the adjusted covariation between decisions made by the Board of Trustees and citation counts of articles published by the applicants.

We performed a multiple regression analysis, which reveals the factors that exert a primary influence on a certain outcome. The coefficients in the regression model, called 'partial' regression coefficients (Rabe-Hesketh & Everitt, 2004), represent the effects of each factor, controlling for all other factors in the model. Since the distribution of citation counts suggests the use of a negative binomial specification (Glänzel & Schubert, 1993), we calculated a Negative Binomial Regression Model (NBRM, Long & Freese, 2003, section 7.3). The citation counts for the applicants' articles enter into the estimation of the NBRM (model A) as a dependent variable ("complete counts"). Next, as "it would be reasonable to assume that the name order of authors listed on a given paper reflects the level of their contributions – with the greatest contributor listed first" (Lindsey, 1980, p. 148), we estimated a second NBRM (model B) using only citation counts of those articles of which the B.I.F. fellowship applicants were the first authors ("straight counts", see Lange, 2001).

Table 2. Description of the independent variables

Independent variable	Values	Mean value or percent of value '1'
Model A: complete counts of citations (*n*=1,459)		
Number of co-authors of the article	1 → 15	4.4
JCR impact factor of the journal	.075 → 42.929	5.5
Decision of the Board of Trustees	rejected (0) → approved (1)	19%
Model B: straight counts of citations (*n*=703)		
Number of co-authors of the article	1 → 13	3.6
JCR impact factor of the journal	.104 → 40.361	5.2
Decision of the Board of Trustees	rejected (0) → approved (1)	20%

Table 2 shows a description of the independent variables that were included in model A and model B. In addition to the decision of the Board of Trustees (approved or rejected), both models take the JCR impact factor of the journals that published the applicant's article and the number of co-authors of each article into account. The publication year of each article was included in the models as exposure time (Long & Freese, 2003, pp. 264-266). By using the exposure() option provided in the statistical package Stata (StataCorp., 2003), the amount of time that an article is "at risk" of being cited is considered. The violation of the assumption of independent observations by including citation counts of more than one article per applicant is considered in the models by using the cluster() option in Stata (StataCorp., 2003). This option specifies that the citation counts are independent across articles of different applicants, but are not necessarily independent within articles of the same applicant (Hosmer & Lemeshow, 2000, section 8.3; Long & Freese, 2003, pp. 74-75).

The results of model A and model B presented in Table 3 for predicting citation counts for articles published previous to the applicants' approval or rejection for a fellowship show similar results. In both models, statistically significant effects in the expected directions for factors that in bibliometric studies have been demonstrated to have a *general* influence on citation counts could be found in our model estimations: for an article, more citations are expected the greater the number of co-authors of the article and the higher the impact factor of the journal that published the article.

As to the variable "decision of the Board of Trustees", the models yield the following results: for the articles published by approved applicants, a statistically significant greater number of citations is expected than for articles published by rejected applicants. The calculation of the percent change coefficients for the Board of Trustees' decisions following the NBRM estimation (Long & Freese, 2003, p. 256) show that being an approved applicant increases the expected number of citations by 49% (model A) and 37% (model B) – holding all other variables constant. In the light of both variables indicating impact of scientific research (complete and straight counts of citations), the Board of Trustees of the B.I.F. was able to accomplish the difficult task of assessing the scientific merit of the applicants absolutely accurately and selecting the best junior scientists among fellowship applicants.

Table 3. Negative binomial regression models predicting complete and straight counts of citations of articles published previous to applicants' approval or rejection

Independent variable	Coefficient	Robust standard error	*p* value
Model A: complete counts of citations (*n*=1,459)			
Publication year of the article	(exposure)		
Number of co-authors of the article	.05	.02	.006
JCR impact factor of the journal	.11	.01	.000
Decision of the Board of Trustees (approved)	.40	.11	.000
Model B: straight counts of citations (*n*=703)			
Publication year of the article	(exposure)		
Number of co-authors of the article	.07	.03	.020
JCR impact factor of the journal	.12	.01	.000
Decision of the Board of Trustees (approved)	.32	.14	.021

Even if the findings in Figure 1 and Table 3 show that the B.I. F. Board of Trustees selected applicants with a higher impact of scientific research than rejected applicants, we still do not know whether the foundation was supporting "scientific excellence". This question can be answered only by comparing the research performance of approved and rejected applicants with international scientific reference values. Anthony F. J. van Raan of the Center for Science and Technology Studies (CWTS) in Leiden, Netherlands, recommends a worldwide reference indicator for the bibliometric evaluation of research groups: "Our most important bibliometric indicator, the 'crown indicator', is a trend analysis over a period of, say, eight years, of the number of citations to the entire oeuvre of a research group or institute, normalized to an international field-specific reference value. In this way, we are able to demonstrate whether this group or institute is performing below or above, or even far above the international level of the research field(s) concerned" (van Raan, 1999, p. 420).[2]

To calculate the "crown indicators", we used the journal sets provided by Thomson ISI (see Essential Science Indicators, ESI) corresponding to the fields "Molecular Biology & Genetics" and "Biology & Biochemistry." We selected these two sets out of the 22 journal sets provided[3] because the journals in which about 40% of the applicants' articles were published are assigned by Thomson ISI (see JCR) to the journal subject categories "Biochemistry & Molecular Biology" and "Genetics & Heredity" (see

[2] The "crown indicator" corresponds to the "relative subfield citedness" (R_w) suggested by Vinkler (Vinkler, 1986, 1997). "R_W compares … the total number of citations obtained to a quantitiy, which is independent of the authors i.e. citations attained by researchers working on the same field worldwide" (Vinkler, 1997, p. 165).

[3] Agricultural Sciences; Biology & Biochemistry; Chemistry; Clinical Medicine; Computer Science; Ecology/Environment; Economics & Business; Engineering; Geosciences; Immunology; Material Sciences; Mathematics; Microbiology; Molecular Biology & Genetics; Multidisciplinary; Neuroscience & Behavior; Pharmacology & Toxicology; Physics, Plant & Animal Science; Psychology/Psychiatry; Social Sciences, general; Space Science (for a description of the journal sets see http://www.in-cites.com/field-def.html).

Table 4).[4] Table 4 shows, e.g., that 502 of the applicants' articles were published in journals belonging to the JCR subject category "Biochemistry & Molecular Biology".

To determine the "crown indicator" for the articles by the B.I.F. applicants, we divided the mean number of citations for articles from applicants published in "Biochemistry & Molecular Biology" and "Genetics & Heredity" journals by the mean number of citations of all publications (A) in the journal set "Molecular Biology & Genetics" and (B) in the journal set "Biology & Biochemistry". The quotient allows us to determine whether the citation impact of the approved and rejected applicants is far below (indicator value < 0.5), below (indicator value 0.5 - 0.8), approximately the same as (0.8 - 1.2), above (1.2 - 1.5), or far above (> 1.5) the international (primarily the Western world) citation impact baseline for the chosen journal sets. With ratio values above 1.5, the probability of identifying very good to excellent researchers is very high (van Raan, 2004, pp. 31-32).

Table 4. Subject categories of the journals (classification according to JCR) in which applicants' articles had been published (with absolute number and relative percent of articles) previous to approval or rejection of applicants' fellowship applications

Journal subject category	Number of articles, absolute	Number of articles, in percent
Biochemistry & Molecular Biology	502	34
Cell Biology	292	20
Neurosciences	151	10
Genetics & Heredity	95	7
Multidisciplinary Sciences	91	6
Immunology	86	6
Biophysics	81	6
Oncology	61	4
Developmental Biology	59	4
Pharmacology & Pharmacy	53	4
Haematology	41	3
Physiology	39	3
Endocrinology & Metabolism	38	3
Plant Sciences	37	3
Medicine, Research & Experimental	34	2
Biology	34	2
Microbiology	32	2
Biotechnology & Applied Microbiology	31	2
Other subject categories (altogether 66 different categories, each with less than 30 articles)	422	29
Total	2,179	150

Note.
The total of the percentage is greater than 100%, since Thomson ISI normally assigns journals to more than one subject category. We calculated the percentage based on the number of articles (n=1,462) and not on the number of the assigned journal subject categories (n=2,179).

Table 5 lists the "crown indicators" of the articles classified according to journal set and year of publication (1991–1994). The average citation counts of articles published by the applicants between 1981 and 1990 are not listed in the table, since Thomson ISI did not provide baselines for articles published in those years with a citation window by the end of 2001. Furthermore, the years 1995 to 1996 are not listed, because the number of articles published by the applicants in "Biochemistry & Molecular Biology" and "Genetics & Heredity" journals is not large enough to calculate average citation counts. The "crown indicators" in Table 5 show that the articles published by approved **and**

[4] For a description of the journal subject categories see http://www.isinet.com/journals/scope/scope_scie.html.

rejected applicants were on average significantly more frequently cited than the "average" publication in both ESI journals sets: 10 of the 16 "crown indicators" are above 1.5 (between 1.54 and 10.33), and three are between 1.2 and 1.5. Only three values (0.89, 1.03 and 1.19) are in the range that van Raan (2004) denotes as "average". A comparison of approved and rejected applicants shows that the "crown indicators" of the approved applicants (six out of eight) are more frequently in the range that van Raan (2004) denotes as "very good to excellent researchers"(> 1.5) than the "crown indicators" of the rejected applicants (four out of eight).

Table 5. Average citation counts of articles published by approved and rejected applicants in "Biochemistry & Molecular Biology" and "Genetics & Heredity" journals (classification according to JCR) compared to mean citation counts of publications in the ESI journal sets "Molecular Biology & Genetics" and "Biology & Biochemistry" by publication year (1991-1994)

	Year of publication			
	1991	1992	1993	1994
(A) Baseline[1] for the journal set "Molecular Biology & Genetics"	40.16	38.22	36.83	32.63
Mean number of citations for articles by **approved** applicants from the year of publication to 2001	52.24 (*n*=62)[2]	45.29* (*n*=14)	69.40* (*n*=10)	198.67* (*n*=6)
Crown indicator (mean of citations divided by baseline)	**1.30**	**1.19**	**1.88**	**6.09**
Mean number of citations for articles by **rejected** applicants from the year of publication to 2001	35.53 (*n*=76)[2]	39.23 (*n*=73)	49.38 (*n*=63)	43.88 (*n*=59)
Crown indicator (mean of citations divided by baseline)	**0.89**	**1.03**	**1.34**	**1.35**
(B) Baseline[1] for the journal set "Biology & Biochemistry"	23.04	22.30	20.79	19.24
Mean number of citations for articles by **approved** applicants from the year of publication to 2001	52.24 (*n*=62)[2]	45.29* (*n*=14)	69.40* (*n*=10)	198.67* (*n*=6)
Crown indicator (mean of citations divided by baseline)	**2.27**	**2.03**	**3.34**	**10.33**
Mean number of citations for articles by **rejected** applicants from the year of publication to 2001	35.53 (*n*=76)[2]	39.23 (*n*=73)	49.38 (*n*=63)	43.88 (*n*=59)
Crown indicator (mean of citations divided by baseline)	**1.54**	**1.76**	**2.38**	**2.28**

Notes.

[1] Baselines are measures of cumulative citation frequencies across all papers published in a journal set: an average of 40.16 for the journal set "Molecular Biology & Genetics" in 1991 means that, on average, papers in "Molecular Biology & Genetics" journals were cited 40.16 times from 1991 to the end of 2001.

[2] n=number of articles.

* As there is a great danger that sample means can be upset completely by a few outliers when sample size is small, we calculated Huber's M-estimators (see above) in addition to the average citation counts for the years 1992, 1993, and 1994. With the values 38.67 (for 1992), 43.05 (for 1993), and 169.14 (for 1994), these robust maximum-likelihood estimators of location calculated for articles of approved applicants also lie above the baseline values for the journal sets.

Thus, when compared to international scientific reference values, the impact of scientific research is above average not only for successful B.I.F. fellowship applicants, but also for non-successful applicants. Furthermore, the values indicating excellent research performance are more frequent for approved than for rejected applicants.

Discussion

In this first comprehensive study on committee peer review for the selection of post-graduate research fellowship recipients, we analysed the committee peer review procedure used by the B.I.F. with regard to whether the foundation is achieving its goal to select the "best" junior scientists to receive fellowships. Assessing the quality of selection decisions requires a generally accepted criterion for the impact of scientific research. Citation counts are considered to be an indicator of research impact, since they measure the international impact of the work by individuals or groups of scientists on others: “A highly cited work is one that has been found to be useful by a relatively large number of people, or in a relatively large number of experiments” (Garfield, 1979, p. 363).

Our bibliometric analyses show that the committee peer review of the B.I.F. indeed achieved the foundation’s goal of selecting the “best” junior scientists with the highest impact of scientific research for fellowships. According to our model estimation, articles previously published by successful fellowship applicants are expected to receive 37% (straight counts of citations) and 49% (complete counts of citations) more citations than articles previously published by non-successful applicants. Moreover, a comparison with international scientific reference values reveals that (a) articles previously published by approved **and** rejected applicants are cited considerably more often than the “average” publication and (b) that excellent research performance is expected more for approved than for rejected applicants. This means that not only does the foundation achieve the fellowship program goal of providing financial support to scientifically excellent researchers, but it also is successfully attracting an applicant pool of highly talented junior scientists – the latter being a prerequisite for the former.

Chapman & McCauley (1994) and Mavis & Katz (2003) reported similar findings for quality ratings of graduate fellows funded by the National Science Foundation (NSF, Arlington, Virginia, USA) and for funding decisions of the March of Dimes Birth Defects Foundation (Indianapolis, USA). In addition, similar results have been reported for selection decisions in the journal peer review process. Based on mean citation rates for accepted manuscripts and rejected manuscripts that were nevertheless published elsewhere, the decisions made by the editors of the *Journal of Clinical Investigation* (Wilson, 1978), *British Medical Journal* (Lock, 1985), and *Angewandte Chemie* (Daniel, 1993, 2005) reflect a high degree of validity.

Although according to Shadish (1989) “of all the science indicators we have, only citation counts are widely available, inexpensive, intuitively plausible, perceived to be reasonably fair, and generally applicable to the scientific community and its products” (p. 394), we plan to consider further success rate factors in addition to bibliometric indicators in determining the effectiveness of the B.I.F. peer review procedure. For example, the B.I.F. has some information available on the further career paths of the fellows. However, for conducting retrospective event history analysis (Blossfeld & Rohwer, 2002; Enders & Bornmann, 2001) the B.I.F. database lacks detailed information on the different stages of the fellows’ careers (such as type of employment, start and end dates for individual periods of employment, sector of employment). As the evaluation of career course data would provide a good complement to the bibliometric analyses (see e.g. Wellcome Trust, 2001), we plan in a future study to conduct a survey of the fellows in order to gather the needed data on their career paths.

References

Balaban, A. T. (1996). How should citations to articles in high- and low-impact journals be evaluated, or what is a citation worth? *Scientometrics, 37*(3), 495-498.

Beaver, D. B. (2004). Does collaborative research have greater epistemic authority? *Scientometrics, 60*(3), 399-408.

Blossfeld, H.-P., & Rohwer, G. (2002). *Techniques of event history modeling. New approaches to causal analysis* (Vol. 2). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Boehringer Ingelheim Fonds. (1999). *A foundation in progress*. Stuttgart, Germany: Boehringer Ingelheim Fonds (B.I.F.).

Bornmann, L., & Daniel, H.-D. (2004). Reliability, fairness and predictive validity of committee peer review. Evaluation of the selection of post-graduate fellowship holders by the Boehringer Ingelheim Fonds. *B.I.F. Futura, 19*, 7-19.

Bornmann, L., & Daniel, H.-D. (2005). Selection of research fellowship recipients by committee peer review. Analysis of reliability, fairness and predictive validity of Board of Trustees' decisions. *Scientometrics, 63*(2), 297-320.

Chapman, G. B., & McCauley, C. (1994). Predictive validity of quality ratings of National Science Foundation graduate fellows. *Educational and Psychological Measurement, 54*(2), 428-438.

Cole, J. R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge. A festschrift in honor of Eugene Garfield* (pp. 281-300). Medford, New Jersey, USA: Information Today.

Daniel, H.-D. (1993). *Guardians of science. Fairness and reliability of peer review*. Weinheim: Wiley-VCH.

Daniel, H.-D. (2005). Publications as a measure of scientific advancement and of scientists' productivity. *Learned Publishing, 18*(2), 143–148.

Ehses, I. (2004). By scientists, for scientists. The Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - and how it functions. *B.I.F. Futura, 19*, 170-177.

Enders, J., & Bornmann, L. (2001). *Karriere mit Doktortitel? Ausbildung, Berufsverlauf und Berufserfolg von Promovierten*. Frankfurt am Main: Campus.

Fröhlich, H. (2001). It all depends on the individuals. Research promotion - a balanced system of control. *B.I.F. Futura, 16*, 69-77.

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool. *Scientometrics, 1*(4), 359-375.

Glänzel, W., & Schubert, A. (1993). A characterization of scientometric distributions based on harmonic means. *Scientometrics, 26*(1), 81-96.

Guston, D. H. (2003). The expanding role of peer review processes in the United States. In P. Shapira & S. Kuhlmann (Eds.), *Learning from science and technology policy evaluation. Experiences from the United States and Europe* (pp. 81-97). Cheltenham, UK: Edward Elgar.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2. ed.). Chichester, UK: John Wiley & Sons, Inc.

Huber, P. J. (2003). *Robust statistics*. Chichester, UK: John Wiley & Sons Inc.

Kostoff, R. N. (1997). The principles and practices of peer review. *Science and Engineering Ethics, 3*(1), 19-34.

Lange, L. L. (2001). Citation counts of multi-authored papers – first-named authors and further authors. *Scientometrics, 52*(3), 457-470.

Lindsey, D. (1980). Production and citation measures in the sociology of science - the problem of multiple authorship. *Social Studies of Science, 10*(2), 145-162.

Lock, S. (1985). *A difficult balance: editorial peer review in medicine*. Philadelphia, PA, USA: ISI Press.

Long, J. S., & Freese, J. (2003). *Regression models for categorical dependent variables using Stata*. College Station, Texas, USA: Stata Press, Stata Corporation.

Mavis, B., & Katz, M. (2003). Evaluation of a program supporting scholarly productivity for new investigators. *Academic Medicine, 78*(7), 757-765.

Rabe-Hesketh, S., & Everitt, B. (2004). *A handbook of statistical analyses using Stata*. Boca Raton, UK: Chapman & Hall/CRC.

Shadish, W. R. (1989). The perception and evaluation of quality in science. In B. Gholson, W. R. Shadish, R. A. Neimeyer & A. C. Houts (Eds.), *Psychology of science. Contributions to metascience* (pp. 383-426). Cambridge, UK: Cambridge University Press.

StataCorp. (2003). Stata statistical software: release 8. College Station, Texas, USA: Stata Corporation.

Tainer, J. A. (1991). Science, citation, and funding. *Science, 251*(5000), 1408.

Tregenza, T. (2002). Gender bias in the refereeing process? *Trends in Ecology & Evolution, 17*(8), 349-350.

Ulrich's Periodicals Directory. (2004). *The global source for periodicals*. New Providence, NJ, USA: R.R. Bowker.

van Raan, A. F. J. (1999). Advanced bibliometric methods for the evaluation of universities. *Scientometrics, 45*(3), 417-423.

van Raan, A. F. J. (2004). Measuring science. Capita selecta of current main issues. In H. F. Moed, W. Glänzel & U. Schmoch (Eds.), *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems* (pp. 19-50). Dordrecht: Kluwer Academic Publishers.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics, 10*(3-4), 157-177.

Vinkler, P. (1997). Relations of relative scientometric impact indicators. The relative publication strategy index. *Scientometrics, 40*(1), 163-169.

Wellcome Trust. (2001). *Review of Wellcome Trust PhD research training. Career paths of a 1988 - 1990 prize student cohort*. London, UK: Wellcome Trust.

Wilson, J. D. (1978). Peer review and publication. *Journal of Clinical Investigation, 61*(4), 1697-1701.

Ziman, J. (2000). *Real science. What it is, and what it means*. Cambridge, UK: Cambridge University Press.