

In den Händen der Sozialforscher

Hermann Fröhlich
Managing Director



• In der Wissenschaft geht eine Pandemie um; sie trägt den Namen »Evaluation«. Auch der Boehringer Ingelheim Fonds (B.I.F.) ist längst infiziert. Um uns zu vergewissern, daß wir unserem Auftrag gerecht werden, haben wir 1999 Dr. David Evered – den früheren Generalsekretär des britischen Medical Research Council (MRC) und Berater zahlreicher europäischer Organisationen der Forschungsförderung – darum gebeten, den B.I.F. und seine Arbeit unter die Lupe zu nehmen.

Im April 2000 gelangte David Evered in einem »Strategic options appraisal for the B.I.F.« zu dem Ergebnis, daß die Stiftung über die Jahre ein eigenständiges Förderangebot entwickelt hat, welches ihr in den biomedizinischen Wissenschaften Anerkennung und Wertschätzung einträgt. Es fielen Worte wie »measurable value«, »impressive impact« oder gar »uniqueness«. Die Wissenschaftler würden – so las man – die Stiftung vermissen, wenn es sie nicht mehr gäbe.

David Evered hielt das Ansehen, das der B.I.F. in der Wissenschaft genießt, für um so bemerkenswerter, als der Fonds über vergleichsweise geringe finanzielle Mittel verfügte und trotz beachtlich gestiegenem Budget verfügt. Nach seiner Auffassung erforderten die begrenzten Ressourcen es, daß die Stiftung sich auf einige wenige Fördermaßnahmen konzentriert, die sich von den Angeboten der großen Fördereinrichtungen abheben. Nur so kann sie sicherstellen, daß sie dank eines eigenen Profils in der Förderlandschaft wahrgenommen wird. Auch das gilt sicher unverändert.

Das Kuratorium hat nicht alle Vorschläge aufgegriffen, die David Evered seinerzeit unterbreitete. Immerhin ist es seinen Anregungen insoweit gefolgt, als es den Forschungspreis für Postdoktoranden – schweren Herzens – und die Finanzierung von Methodenkursen – leichten Mutes – aufgegeben hat. Seither konzentriert der B.I.F. sich auf Doktoranden-Stipendien, Reisebeihilfen und die Internationalen Titisee Konferenzen. Einen Vorschlag übernahm die Boehringer Ingelheim Stiftung und schuf ein Stipendien-Programm für Doktoranden aus der Medizin (vgl. B.I.F. FUTURA 2002, Vol. 17, S. 3 f.).

David Evereds Einschätzung beruhte auf Gesprächen mit Kuratoren und Geschäftsstelle, Stipendiaten und ehemaligen Stipendiaten sowie Ver-

tretern der Deutschen Forschungsgemeinschaft (DFG) – mit mehr oder weniger Betroffenen also. Durch seinen Zuspruch ermutigt, wollten wir es genauer wissen: Trifft die Auswahl des B.I.F. wirklich die Besten? Fakten und Zahlen mußten her. Butter bei die Fische!

Da war es ein Glück, daß wir seit vielen Jahren mit Professor Dr. Hans-Dieter Daniel – damals am Wissenschaftlichen Zentrum für Berufs- und Hochschulforschung der Universität Kassel, inzwischen Inhaber des Lehrstuhls für Sozialpsychologie und Hochschulforschung an der ETH Zürich und Leiter der Evaluationsstelle der Universität Zürich – einen ausgewiesenen Wissenschaftsforscher kannten. Als er Anfang 2001 mit Lutz Bornmann einen Doktoranden fand, der unsere Fragen zu seinen eigenen machte, fiel der B.I.F. in die Hände der empirischen Sozialforschung.

Das Ergebnis von Lutz Bornmanns dreijährigen Bemühungen ist vor kurzem unter dem Titel »Stiftungspropheten in der Wissenschaft. Zuverlässigkeit, Fairneß und Erfolg des Peer-Review« im Waxmann Verlag erschienen (www.waxmann.com). Es brachte dem Autor ein »summa cum laude« und den Dokortitel ein. Herzlichen Glückwunsch!

Bei dem Buch handelt es sich um nicht weniger als die weltweit erste umfassende Studie zu der Frage, ob Peer-Review – die Beurteilung von Fachkollegen durch Fachkollegen – für die Auswahl von Stipendiaten taugt. Für wissenschaftliche Zeitschriften und für Forschungsgelder ist das Verfahren längst eingehend untersucht und gilt trotz mancher Schwächen als die bestmögliche aller Welten.

Die Voraussetzungen, welche der B.I.F. für eine solche Studie bot und bietet, sind günstig – nicht zuletzt deswegen, weil die Stiftung zu über 90 Prozent ihrer ehemaligen Stipendiaten Verbindung hält – zumindest weiß, wo diese sind und was sie treiben. Die Analyse berücksichtigt denn auch alle 1.954 Bewerbungen um ein Doktoranden-Stipendium, die zwischen 1985 und 2000 eingereicht wurden, und die 743 Gesuche um ein Postdoktoranden-Stipendium, die uns zwischen 1985 und 1995 erreichten.

Lutz Bornmann wertete die 2.697 Anträge aus 15 Jahren Forschungsförderung unter drei Anforderungen aus, die eine jede Evaluation prüfen sollte

und die der sozialwissenschaftliche Fachjargon als »Reliabilität«, »Fairneß« und »Validität« bezeichnet. Die Begriffe stehen für die Fragen: Ist die Auswahl des B.I.F. verlässlich? Wird sie den Antragstellern gerecht? Vor allem aber: Trifft sie wirklich die Besten?

Die Übereinstimmung der Fachleute gilt als Gradmesser dafür, wie zuverlässig ein Begutachtungsverfahren ist. Dabei zeigen Erfahrungen aus Forschungsförderung und anderen Bereichen: Je komplexer ein Sachverhalt ist, je mehr Gesichtspunkte zu berücksichtigen sind, um so weniger stimmen Fachleute in ihren Empfehlungen überein. Und was ist komplexer als der Entwurf eines Forschungsvorhabens, das in Neuland vorstößt oder doch wenigstens vorstoßen sollte?

Beim B.I.F. wirken an der Auswahl der Stipendiaten drei Instanzen mit: Ein Mitarbeiter der Geschäftsstelle führt – wo immer der Aufwand zu vertreten ist – ein Vorstellungsgespräch mit dem Antragsteller. Zu jedem Bewerber wird eine Beurteilung verfasst, in deren Mittelpunkt Lebenslauf und Person des Antragstellers stehen. Ein externer Gutachter, den die Geschäftsstelle unter fachlichen Gesichtspunkten auswählt, prüft vor allem das geplante Forschungsvorhaben und die gewählte Arbeitsgruppe. Das Kuratorium der Stiftung, sieben international angesehene Wissenschaftler, trifft die abschließende Entscheidung.

So verlässlich wie DFG und NSF

• Das Maß der Übereinstimmung, das zwischen Geschäftsstelle, Gutachtern und Kuratorium herrscht, läßt sich nicht in einer Zahl ausdrücken. Lutz Bornmann meint – und ein Blick auf Tabelle 6 auf Seite 57 seines Buches belegt –, daß sie erstaunlich hoch ausfällt. Dieses Phänomen läßt sich sicherlich nur zum Teil darauf zurückführen, daß die Beurteilungen der Geschäftsstelle Gutachten, sofern sie rechtzeitig vorliegen, zwar nicht im Hinblick auf die Beurteilung der Person, wohl aber beim abschließenden Urteil berücksichtigen. Das Kuratorium kennt wiederum sowohl Beurteilungen als auch Gutachten. Trotz dieser einschränkenden Beobachtungen bleibt das Ergebnis bemerkenswert.

Die Feststellung gilt umso mehr, bedenkt man, daß Geschäftsstelle und Gutachter bei ihren Vorschlägen freie Hand haben, während das Kuratorium bei seinen Entscheidungen an die finanziellen Spielräume gebunden ist. In jeder Sitzung müssen die Kuratoren Anträge ablehnen, die sie gerne bewilligen würden. Die Gutachter empfehlen 62 Prozent der Anträge für eine Bewilligung, die Geschäftsstelle 43 Prozent, von diesen 10 Prozent mit Nachdruck. Das Kuratorium muß die Vorschläge auf 25 Prozent zusammenstreichen.

Trotz des erheblichen Entscheidungsdrucks, der auf den Kuratoren lastet, ist auch ihre Übereinstimmung untereinander beachtlich. Die Entscheidungen fallen in zwei bis drei Durchgängen. Im ersten Durchgang erweisen sich 10 Prozent aller Anträge als klare Fälle für eine Förderung, 66 Prozent als ebenso klare Fälle für eine Ablehnung. Die Übereinstimmung der Kuratoren liegt also bei 76 Prozent. Einen Wert von 71 Prozent bis 80 Prozent stufen Sozialwissenschaftler als eine gemäßigte Übereinstimmung ein.

Zum Vergleich: Die DFG bringt es bei Anträgen für Sachbeihilfen auf 82 Prozent, die Heart and Stroke Foundation in Kanada auf 73 Prozent, die National Science Foundation (NSF) in den USA auf 68 Prozent. Was die »Reliabilität« unserer Auswahl der Stipendiaten anbelangt, befinden wir uns also in bester Gesellschaft!

Nationalität spielt keine Rolle, wohl aber Geschlecht und Ausbildung

• Wie ist es um die »Fairneß« des Auswahlverfahrens bestellt? Mit anderen Worten: Spielen bei der Auswahl Eigenschaften eine Rolle, die nicht der Wissenschaft zuzurechnen sind? Hier hat Lutz Bornmann sich zunächst auf die Doktoranden gestürzt und aus allen Anträgen ein Phantom errechnet: Im Durchschnitt ist der Bewerber beim B.I.F. männlich, deutscher Staatsbürger und Biologe. Er will das Promotionsvorhaben an einer deutschen Hochschule durchführen, erlangte den berufsqualifizierenden Abschluß mit 26 Jahren und einem Notendurchschnitt von 1,4 (beste Note: 1,0). Im Studium wechselte er mindestens einmal die Hochschule. Er legt zwei Befürwortungsschreiben vor. Gutachter und Geschäftsstelle empfehlen Förderung. Wer diese Kriterien erfüllt, dessen Chancen auf ein Stipendium stehen 50:50.

In einem zweiten Schritt hat Lutz Bornmann einen Parameter des Phantoms geändert, die anderen Variablen dabei nicht angetastet. Zunächst die gute Nachricht: Hat der Bewerber eine andere als die deutsche Staatsbürgerschaft vorzuweisen, so beeinträchtigt dieser Umstand seine Erfolgsaussichten nicht. Allerdings lassen die anderen Ergebnisse aufhören: Handelt es sich um eine Frau, dann sinken die Chancen um 17 Prozent. Bewirbt sich ein Chemiker, dann fallen sie um 25 Prozent. Will er an einem Max-Planck-Institut arbeiten, dann steigen sie um 17 Prozent.

Bei einer anderen Art der Berechnung mit diesen Variablen, die weniger anschaulich ist und die wir deswegen hier beiseite lassen, fallen die Abweichungen weniger drastisch aus. Schönrechnen lassen sie sich auch dort nicht. Zwar blieb uns über die Jahre nicht verborgen, daß Frauen und Chemiker beim B.I.F. schlechter, Doktoranden an Max-

Planck-Instituten dagegen besser abschneiden. Jetzt aber liegen Zahlen auf dem Tisch.

Fließen in unsere Entscheidungen Kriterien ein, die mit Wissenschaft nichts zu schaffen haben? Um dieser alles andere als trivialen Frage auf den Grund zu gehen, hat Lutz Bornmann seine Zahlenspiele auf die Gruppe der Postdoktoranden übertragen. Und siehe da! Dort war die Welt in Ordnung! Nicht nur bei der Nationalität, sondern auch bei Geschlecht, Fachgebiet und Arbeitsplatz bleiben statistisch signifikante Abweichungen aus. Die Beobachtung gibt uns einen Teil des angeschlagenen Selbstvertrauens zurück. Die Frage nach den Ursachen beantwortet sie nicht!

Daß »Frauen im B.I.F.« und ihre vermeintliche Benachteiligung bei der Auswahl uns seit langem umtreiben, ist in B.I.F. FUTURA 1998, Vol. 13, S. 159 f. und S. 237 f. nachzulesen. Dort mutmaßten wir, daß in der Wissenschaft wie im Leben – und allen gegenläufigen Bemühungen zum Trotz – bei beruflichen und privaten Zielen geschlechtsspezifische Unterschiede obwalten, die tief in gesellschaftlichen Erziehungshaltungen und einer entsprechenden Erziehung verwurzelt sind. Einfacher ausgedrückt: Bei der Auswahl von Promotionsvorhaben und Arbeitsgruppe spielen für Frauen persönliche Motive eine größere Rolle als für ihre männlichen Kommilitonen. Erst auf der nächsten Sprosse der Leiter, wenn der Entschluß gefallen ist, Karriere in der Wissenschaft zu machen, verliert sich der Unterschied.

Um diese Vermutung empirisch zu untermauern, hätte Lutz Bornmann bei der Beurteilung der Projekte und ihres Umfelds in fachlichen Wettstreit mit den Kuratoren treten müssen – ein hoffnungsloses Unterfangen. Und doch: Ebendiese Kriterien geben den Ausschlag, wenn die Kuratoren aus den von Gutachtern und Geschäftsstelle empfohlenen sehr guten die besten Bewerber aussuchen.

Ungleich leichter als bei den Frauen lassen sich die Abweichungen beim Fachgebiet erklären: Der Boehringer Ingelheim Fonds ist eine Stiftung für biomedizinische Grundlagenforschung, die dazu beitragen will, grundlegende Phänomene menschlichen Lebens aufzuklären. Diesem Ziel verschreiben sich weitaus mehr Biologen als Chemiker, deren Projekte beim Fonds der chemischen Industrie oft in besseren Händen sind. Leicht tun wir uns auch, wenn es darum geht zu erläutern, warum Max-Planck-Institute im Vergleich zu Hochschulen Vorteile davontragen. Max-Planck-Institute bieten in der Regel (!) institutionelle und finanzielle Rahmenbedingungen, von denen Forscher an Universitäten zumeist (!) nur träumen können. Was Wunder also, daß gerade sie den Nachwuchs anziehen, der sich besonders ehrgeizige Ziele gesteckt hat. Und auf eben diesen Nachwuchs hat es auch der

B.I.F. abgesehen. Zum Einwand »Immer dieselben?« siehe B.I.F. FUTURA 1991, Vol. 6, S. 5.

PNAS, EMBO Journal, Nature, Cell, Science...

• Reliabilität? Fairneß? Ohne Zweifel ernst zu nehmende Fragen! Und doch: Wirklich aufregend wird es erst, wenn es um die Validität unserer Auswahl geht: Scheiden wir wirklich die Spreu vom Weizen, auch wenn dabei so manches Samenkorn verloren gehen mag? Als ein erster Indikator kann gelten, daß nur zwei Prozent unserer Stipendiaten die Promotion nicht erreichen. Beim britischen Wellcome Trust beenden acht Prozent der Stipendiaten ihre Doktorarbeit nicht. Lutz Bornmann konnte sich im Rahmen seiner Dissertation mit derart schlichten Antworten schwerlich zufriedengeben. Um den Erfolg unseres Auswahlverfahrens zu erfassen, hat er sich auf diejenigen Stipendiaten konzentriert, die über die Förderung hinaus auf Dauer in der Academia verblieben sind. Den Gepflogenheiten des Fachs entsprechend hat er sich ihre Veröffentlichungen vorgenommen.

Alles in allem betrachtet Lutz Bornmann 2.039 wissenschaftliche Artikel aus der Feder von 120 früheren Stipendiaten. Sie erschienen in 508 Zeitschriften, wobei auf 36 Journale zehn und mehr Abhandlungen entfielen. An der Spitze stehen die Proceedings of the National Academy of Sciences der USA, gefolgt vom Journal of Biological Chemistry, dem EMBO Journal, Nature, Development, Cell, Science. Ein Ertrag, der unsere kühnsten Erwartungen übertrifft!

Noch schärfer treten die Konturen hervor, zählt man aus, wie häufig die Artikel von Stipendiaten im Fach zitiert wurden. Hier nahm Lutz Bornmann 1.949 Abhandlungen zur Grundlage, die zwischen 1991 und 2000 in Zeitschriften erschienen sind, die das Institute for Scientific Information (ISI, Philadelphia, USA) zu den drei Gruppen »multidisciplinary«, »molecular biology and genetics« und »biology and biochemistry« gebündelt hat. Er erfaßte, wie häufig die Artikel von Stipendiaten bis Ende 2001 zitiert wurden und verglich diese Zahlen mit der durchschnittlichen Häufigkeit mit der ein Artikel in diesen Journalen zitiert wird. So ermittelte er für jedes der zehn Jahre drei – insgesamt also 30 – sogenannte Crown-Indikatoren.

Mit Auszeichnung bestanden – kein Grund zur Rast

• Wer noch immer am Erfolg unseres Auswahlverfahrens zweifelt, der muß spätestens jetzt die Segel streichen: 21 der 30 Crown-Indikatoren, die Lutz Bornmann für den B.I.F. errechnet hat, fallen sehr gut bis herausragend aus, sieben weitere liegen über dem, nur zwei im Durchschnitt. Die Frage, die uns 1999 in die Arme von David Evered trieb und

seither nicht losgelassen hat, ist – nach den Regeln der empirischen Wissenschaftsforschung – endgültig entschieden: Ja! Wir werden unserem Auftrag und unserem Anspruch gerecht! Wir fördern in Deutschland und darüber hinaus die besten Nachwuchswissenschaftler in der Biomedizin – zwar keineswegs alle, wohl aber so viele, wie wir eben können. Wir, das sind vor allem die Kuratoren, deren Ansprüche an die Qualität von Antragsteller, Promotionsvorhaben und Arbeitsgruppe sich bewährt haben. Ihnen und ihrem selbstlosen Einsatz gebührt einmal mehr Dank!

Als wir den Kuratoren am 27. Februar 2004 die Befunde erläuterten und Lutz Bornmanns Dissertation überreichten, da widmeten wir ihnen die Exemplare mit der Bemerkung: »Nichts ist so gut, daß es nicht noch besser werden könnte.« Und in der Tat hat Lutz Bornmann selbst durch »Vorschläge für die Optimierung des Peer-Review« dafür gesorgt, daß unsere Bäume nicht in den Himmel wachsen.

Einige seiner Anregungen haben wir in Angriff genommen. Dazu zählen ein verstärkter Einsatz

elektronischer Hilfsmittel bei Antragstellung und Antragsbearbeitung und die Einführung einer Vorauswahl bei uneingeschränkter Entscheidungskompetenz der Kuratoren. Bei anderen scheint die Entwicklung uns zuvorzukommen: So gingen 2003 erstmals 60 Prozent der Stipendien an Frauen. Ob hier ein Trend vorliegt, muß die Zukunft weisen.

Wieder andere Vorschläge, die Lutz Bornmann unterbreitet hat, bewegen wir in unserem Herzen. Ob sie verwirklicht werden können, wird nicht zuletzt davon abhängen, ob der mit ihnen verbundene personelle Aufwand finanziell zu vertreten ist. Fördermittel versus Verwaltungskosten! Der Erwägung wert und hilfreich sind die Hinweise allemal. Allein ihretwegen stehen wir in der Schuld von Lutz Bornmann, Hans-Dieter Daniel und der empirischen Sozialforschung. Die Schuld abzustatten, ist uns Bedürfnis und Vergnügen zugleich!



In the hands of social researchers

For a translation of this text see pages 20–23.