Research Evaluation, volume 14, number 1, April 2005, pages 27–34, Beech Tree Publishing, 10 Watford Close, Guildford, Surrey GU1 2EP, England

Quality-only assumption

Is external research funding a valid indicator for research performance?

Grit Laudel

Although 'research income' is one of the most common indicators for assessing research quality, its validity has never been systematically investigated. The conditions under which Australian and German physicists obtain external funding were analysed in a comparative qualitative study. The study demonstrates that success in obtaining external funding is only partly related to the quality of researchers and their proposals. Therefore, the validity of a straightforward counting of external funding must be assumed to be low. A comparison of external funding with citation indicators shows ways to improve the validity of indicators based on external funding.

Grit Laudel is at the Research Evaluation and Policy Project, Research School of Social Sciences, Australian National University Canberra, ACT, 0200, Australia, email: grit.laudel@anu.edu.au

RESEARCH INCOME — funding obtained from external sources — is one of the most common quantitative indicators of research performance. It is appealing because the award of a grant is based on peer review (Gillett, 1991: 254) and because it measures current rather than past performance (Hornbostel, 2001). It seems easy to quantify by counting the number of grants or adding up the external grant money. Finally, external funding is a performance indicator that is well accepted among scientists. When Australian scientists were asked which indicators are appropriate to measure research performance in their field, external funding was one of the preferred (Tognolini *et al*, 1994).

Given these features, it is not surprising that external funding is now used in virtually all evaluations. Research funding for Australia's universities is allocated using three funding formulas in which the research income is weighted between 40% and 60%. Within the universities, research funds are distributed to departments based on a similar formula.

In Germany no performance-based funding exists at the national level, but ranking exercises proceed at a rapid pace. In these rankings of Germany's universities, external funding was used as one of the indicators, both in the form of total sums of external research funds and as external funds per researcher (Berghoff *et al*, 2003; DFG, 2003).

Research funding of British universities is based on assessments by peers (the Research Assessment Exercise, RAE). While the focus of discussions about the RAE is on the peer evaluation of publications submitted to the panel, external research funding is also one of the assessment criteria (RAE,

 0958-2029/05/010027-8 US$08.00

1999a). British university departments have to submit not only publications by research-active staff, but also the amount and sources of external research grants (RAE, 1999b). Although we don't know from the documents *how* the panels actually used the indicator, it becomes obvious that external funding was taken into account, and that panel members considered background knowledge important in interpreting the external funding data.

Furthermore, external funding is also used at the individual level as a criterion in academic job decisions; for example, in tenure track decisions in the United States. Applicants must list the grants they won in the last couple of years or during their whole research career. Usually, the source of the grant is taken into consideration; highly competitive grants from funding sources with a rigorous peer review system (e.g. grants from the National Science Foundation) are weighted higher than others.

Thus, external funding is used as a performance measure across all levels of aggregation, from national systems of performance-based funding of universities down to decisions about individual positions. The major variation in the use of the indicator is that it is either directly tied to funding via a formula or used as an indicator to inform peer review. External funding is also measured in a very straightforward manner. Although it is hardly possible to obtain a comprehensive overview of the current practices, it seems to be more common to use the amounts of funding rather than the number of grants. On aggregated levels, the size of the evaluated unit is sometimes taken into account by measuring research income per researcher. Differences between disciplines or discipline groups have commonly been considered by only comparing departments within the same discipline. The sources of funding have been taken into account by giving more weight to sources with rigorous peer review procedures. Beyond this no further qualifications have been made.

These observations indicate that regardless of its widespread use, little is known about the indicator. Its application rests on the assumption that the amount of research income (or the number of grants) depends only on the quality of a researcher or research proposals. This assumption — the quality-only assumption — implies that non-quality related conditions do not distort the quality-based distribution of funds, i.e. that quality is not only necessary but also sufficient to win a grant.

In spite of the ubiquitous application of indicators that are based on external income and thus the importance of the quality-only assumption, neither this assumption nor the validity of the indicator 'external income' have been the subject of systematic investigations. While empirical studies about the validity of the indicator are lacking, the literature raises some possible disturbing factors. Hornbostel (2001) suggests that external fund acquisition is an appropriate research performance indicator only when:

- external fund acquisition is common in the field (for example, it is not common in law sciences);
- grant proposals are reviewed by qualified peers in a competitive system (this condition might not always be given, but there is a lack of empirical data about distortions in the peer review process of grant proposals);
- there is a clear, but not disastrous competition for the limited funds;
- there is a mix of different sources; and
- an essential infrastructure is available that enables research.

However plausible these assumptions are, so far they have not been empirically tested. The same holds by and large for the Matthew effect in external funding (those who get the most external funding are likely to get even more while others are crowded out of the system). Gillett hypothesizes the possibility of a positive feedback loop 'in which those who receive grants in the past are more likely to be awarded them in the future' (Gillett, 1991: 260). The author assumes that externally funded researchers are more likely to publish than non-externally funded researchers, and that reviewers use an applicant's publication record as an assessment criterion. A similar argument is made by others who claim that grants are an input of the research process rather than an output:

> If peers consider both grants and publications to be indicative of high research productivity, then they are guilty of double counting, for some publications would not have been possible without the favourable research environment created by the grants. (Johnes and Johnes 1995: 309, similarly argues Phillimore, 1989) [1]

This hypothesis is partly supported by an empirical study at the department level in Flemish universities. The 'rich' departments, i.e. departments that acquired most external funds, became even 'richer' in the following ten years. The allocation of basic allowances, as indicated by the distribution of scientific personnel, remained stable. Thus, a situation may emerge where the basic allowances are too small for externally funded research activities (Moed *et al*, 1998). However, no conclusion about causal mechanisms is possible because the conditions of research and research funding were not investigated.

The aim of this paper is to test the causal hypothesis that is built into the 'quality-only assumption' and therefore underlies the use of external research funding as a performance indicator. This causal hypothesis can be phrased as follows: The better a researcher/a research proposal, the higher the probability of acquiring external funding. In order to test this hypothesis, I analysed the conditions which affect a researcher's success in gaining external funds. This analysis will not only lead to the rejection of the simplified causal hypothesis, but will also

indicate how the validity of the indicator can be improved, and what further research on its use is necessary.

Data and methods

In a comparative, qualitative study, I investigated the promoting and hindering conditions of external fund acquisition of German and Australian university researchers in experimental physics. This field was chosen because it requires significant resources, most of which must be available locally in the research groups' laboratories. The restriction to one field limited the variance in some intervening variables, such as institutions of external funding and resource demand.

Semi-structured interviews with 45 German physicists were conducted in 2000, and interviews with 21 Australian physicists in 2002. Scientists were selected according to the following criteria:

1. *Variation of a scientist's basic supplies*: It was assumed that the amount of basic supplies influences the ability to acquire external funding. Therefore, German professors at different levels (C3 and C4) were selected. The Australian system does not have these distinctive hierarchies; thus only scientists from varying academic positions were chosen.
2. *Variation of career stage*: Scientists of all career stages, from young scientists at the start of their career to scientists at the end of their career, were included.
3. *Variation of the subfields*: The physicists were chosen from different subfields of experimental physics in order to obtain information about the influence of epistemic conditions on the resource decisions.

The interviews included the following topics: the scientists' research trails and their origins; current research projects and their funding sources; resource needs; funding opportunities and prerequisites for obtaining funds from the different sources; support measures provided by the universities; causes of the failure to gain external funds; and adaptations of research to conditions of funding. The interviews were subjected to a qualitative content analysis.

Moreover, I analysed publication and citation data for each scientist. These data served as an additional indicator for the reputation of a scientist.[2] The whole publication oeuvre of a scientist was retrieved. Citation data comprised a ten-year time span from 1990 to 1999 for the German scientists, and from 1993 to 2002 for the Australian scientists. The Web of Science was used as a data source, therefore only journal articles were included. However, this was considered to be a good indicator since this is the dominant publication type in experimental physics. The scientists were ranked according to their reputation in their scientific communities and to the number of citations they received. The top one third of this ranking are here regarded as 'top' scientists. Against the general methodological rule, citation analyses were used on the level of individual scientists. However, this approach can be justified because citations were counted over a relatively long time-span; the indicator was used in connection with a considerable number of other indicators; and it served as a rough distinction between two groups rather than a fine-grained ranking of individual scientists.

Furthermore, I distinguished two groups of scientists according to the amount of external funding they had acquired. As a rough indicator, I used only the number of externally funded researchers (rather than the whole amount of funding) as this is usually the most expensive part of a grant. This enabled me to make comparisons over different subfields of experimental physics. I defined German scientists who had seven or more externally funded research positions as 'rich'. Australian scientists who had much less external funding sources available were defined as 'rich' if they had five or more such research positions.

Influencing conditions

The comparison of external funding obtained by 'top' and 'other' scientists (Tables 1 and 2) already calls into question the quality-only assumption. Based on this assumption one would expect that 'top' scientists have many funds while other scientists have significantly less funds. The data clearly show that the link between quality and amount of external funding is weak at best. 'Top scientists' were almost equally distributed between those who had large amounts of external funding and those who had not. Also, 'other scientists' frequently were able to get large amounts of external funding.

In the following, I will summarise the conditions I found in my empirical study that influence a researcher's success in acquiring external funds; a detailed discussion of these conditions is published elsewhere (Laudel, 2005).

In order to successfully acquire external funds, a researcher needs an appropriate funding source, enabling funds and a fundable proposal. If we look at these seemingly trivial prerequisites more closely, it turns out that they rest on a very sophisticated system

Table 1. Reputation and funding of German scientists

	amount of external funding	
	'rich'	**'other'**
'top' scientists	8	7
'other' scientists	7	23

Table 2. Reputation and funding of Australian scientists

	amount of external funding	
	'rich'	**'other'**
'top' scientists	3	4
'other' scientists	4	10

of conditions. They depend on a very complex set of cognitive, social, and institutional conditions which determine the opportunities for a researcher to actually acquire external funding. To determine the validity of external research funding as an indicator of quality, we must examine the extent that conditions for fund acquisition depend on a researcher's or research proposal's 'quality'.

Appropriate funding sources

An appropriate funding source is one that funds research on the topic of the proposal, whose eligibility criteria are met by the applicant, and whose terms of funding meet the funding needs of the project. The availability of funding sources is mainly shaped by the diversity of the funding landscape, the availability of collaborators, the epistemic room for manoeuvre, and the integration in the scientific community. A *diverse funding landscape* allows the scientist to approach different sources. The funding landscapes of Germany and Australia differed considerably in this respect. Germany has numerous funding sources accessible to experimental physicists, while Australian physicists basically relied on one single funding source — the Australian Research Council (ARC).

Funding sources vary in the amount of money they distribute, in the financial autonomy they grant, and in the topics of research that are funded. Many funding sources create programmes that offer funding for specified topics. Whether a scientist can apply for such a programme depends on their chosen research topic. Often, funding agencies are directed towards applied and industry-related research. Therefore scientists who conducted applied research could access a larger variety of funding sources. Scientists conducting basic research in Australia relied on a single funding source with a very low success rate and with the number of grants being restricted to two per scientist.

An appropriate funding source is one that funds research on the topic of the proposal, whose eligibility criteria are met by the applicant, and whose terms of funding meet the funding needs of the project

Many funding schemes are aimed at promoting collaborative research. The *availability of collaborators* (who in some cases need to be industry partners) was a crucial precondition to access these schemes. This availability is affected by:

- the size and structure of the field, i.e. how many research groups work in the field;
- by industry-relevance and existence of potential collaborators in industry;
- the quality of work, since scientists try to collaborate with the best partners available;
- the topic because collaborations must be feasible for the proposed research project.

The opportunities for scientists to adapt their proposals' contents depend on their capabilities and on their *epistemic rooms of manoeuvre*. The epistemic room of manoeuvre describes the spectre of possible research a scientist can undertake with the objects and methods available. A large epistemic room of manoeuvre means that a greater variety of projects can be designed and thus provides a larger range of opportunities to apply for external funds. For example, some scientists used universal methods, i.e. methods that were applicable in a variety of fields. They could therefore access a larger number of different funding sources and they were often welcomed as collaborators in research networks.

Availability of enabling funds

Enabling funds are necessary because external funds usually do not cover all expenses. Often the applicant's salary, basic equipment, rooms and technical support are excluded. To conduct a project requires significant funds from sources other than the external agency which receives the grant proposal. They can come from two different sources. The most important one is *recurrent funding*. The scientists' recurrent funding varied considerably: I found in both countries, Australia and Germany, relatively 'wealthy' research groups and relatively 'poor' research groups. The recurrent funding of a research group directly influenced the amount of external funding it was able to acquire. None of the research groups with poor recurrent funding were able to obtain a large number and amount of external funding while groups with large recurrent funding were often the most successful groups in acquiring external funds.

Apart from recurrent funding, scientists used *external funding from other projects* to finance basic supplies, i.e. they 'rearranged' money from other externally funded projects to supplement funds. The availability of such 'free money' is mainly influenced by the number of externally funded projects

and the restrictions posed by the funding agency on how the money can be spent.

Thus, access to enabling funds is influenced by:

- the available funds of a university, which affect what can be offered as recurrent funding to a scientist;
- the number of externally funded projects, which is influenced by quality and non-quality-related factors; and
- the quality of work. Good researchers might receive better offers of basic supplies than mediocre researchers but this is highly restricted by the funding opportunities of the research institution. The quality also has an influence on the number of externally funded projects.

Acceptability of funding proposals

The most important features of a proposal are the characteristics of the project itself, i.e. *quality and feasibility*. While a high quality is generally required by the proposed project, 'normal' projects generally have better chances than unusual ones. As peer review studies have shown, it is preferably mainstream, low-risk and disciplinary research that passes the peer review process of funding agencies (e.g. Chubin and Hackett, 1990: 69; Horrobin, 1996; Travis and Collins, 1991: 336).

It is often important for the applying scientists to demonstrate that they have either worked on the topic of the proposed project for some time or that they have at least prepared the project by some empirical work. An applicant's amount of *prior work* is increased by a continuous research trail; that is, research projects in the same area which are connected to each other. The necessary prior work can be conducted with either recurrent funding or external funding from other projects. These two sources of funding affect a scientist's opportunity to perform the prior work and to build up a track record.

A second effect of a scientist's prior work is that it contributes to their *track record* and general *reputation*, which are important factors affecting a proposal's success. It is not only the quality of the project that is assessed, but the whole of a scientist's former research. Therefore, scientists who are well-known in their community for their (good) research and who are integrated into informal networks and into decision-making processes, have an advantage in acquiring funds.

The main factor in building a reputation is of course research. Scientists needed to conduct projects that led to publications. This, in turn, depended on funding. Scientists who had small basic supplies and who were not successful in fund-raising were caught in a vicious circle of not being able to conduct the necessary prior work and maintain a sufficient publication record. This situation was particularly evident within the Australian science system with its small basic supplies and lack of diversity in funding sources.

As a result of shrinking success rates, it becomes even more important to *know how to 'play the game'*. This includes knowledge about which funding programmes are available, what the formal rules of each funding scheme are and general knowledge about how to write a grant proposal.

A general promoting condition for external funding was the *availability of time*. Time is necessary for the preparation of the grant proposal as well as for reporting activities. The amount of time a researcher had for grant applications and reporting depended on the time the researcher had to spend on non-research activities like teaching and administrative duties at the university.

Discussion

The dependence of success on quality

It was demonstrated that the quality of a proposal and the reputation of a researcher are important prerequisites for a successful acquisition of funds. However, I have also shown that a funding proposal's success depends on several factors that are not linked to quality and cannot even be controlled by scientists. An often neglected factor is a country's general investment in research. The extent to which a quality-based fund distribution can even work depends on the willingness of science policy to provide a sufficient material basis for applicants. If recurrent funding is insufficient to prepare grant applications, or if grants are so scarce that even excellent projects must be rejected, the mechanism cannot work properly. This is a fundamental difference between competitive grants and markets that are the common frame of reference in neoliberal policy. While the income from market exchange is supposed to guarantee the existence of market actors and therefore designed to cover all costs, grants only cover a part of those costs, and it is the task of science policy to make actors in science fit for the competition.

Another major factor that is not related to a proposal's or its author's quality is the research field in which a scientist is working. Fields significantly differ in the amount of grant money that is available. For example, fields that do not have any promise of applicability for industrial innovation will be funded neither by industry nor by agencies that are devoted to the promotion of industrial innovation. Basic research fields differ in the extent to which they can attract collaborators, in the amount of money that is needed for basic supplies, in the speeds with which their basic equipment needs to be replaced, in the epistemic room of manoeuvre they provide, etc. None of these factors can be changed by the researcher.

A third important factor is the availability of 'free' money to prepare project proposals, start new lines of research, etc. The amount depends partly on quality because better researchers can be assumed to obtain posts at more prestigious and wealthy

organisations, and will have more 'spare money' from other projects. However, the limitations of recurrent funding cannot be completely overcome.

A fourth main factor, which also only partly depends on quality, is the continuity of the research trail. The emphasis on track record and prior work on a topic channels the money to the previously successful researchers. Thus, for certain scientists it is difficult to get external funding, e.g. for those who are at the start of their research career or have interrupted it, and for those who attempt to begin new lines of research.

Consequences

The analysis has shown that a researcher's success in acquiring external funds depends on many factors, most of which are not or only partly related to the quality of the research proposal or of its author. This observation casts serious doubts on the validity of performance indicators based on external funding if the latter are used without further refinements. The question arises as to whether and how a valid indicator of 'external research funding' can be constructed.

To address this problem, I will draw on knowledge of another indicator of research quality which has been investigated for many years, namely citations. The introduction of citation-based indicators for evaluation purposes was accompanied by heavy controversies regarding their validity. Today citations are widely accepted as a valid indicator for measuring certain aspects of research quality. As I will show, both indicators (citations and external research funding) face similar threats to their validity. The solutions to validity problems of citation-based indicators may therefore serve as a heuristics for constructing valid indicators of 'external research funding'.

What do the indicators actually measure? Initially, the validity of citation indicators was questioned because they seemed too simple to measure such a complex phenomenon as research quality. This problem was solved by stating that citations are only partial indicators of research quality (Martin and Irvine, 1983). The current state of the art can be summarised in the statement that citations measure one important aspect of a publications quality, namely the international impact of research (Van Raan, 1996: 404). Other aspects of quality are not captured by citations and can be assessed only by peers (Moed et al, 1985: 134). A principal methodological consequence that follows from this assertion is that citation indicators should never be used alone in research assessments but always in connection with other indicators.

A similar approach to indicators based on external funding appears to be in order. External research funding is usually based on the assessment of several peers. Given the role of peer review in the system of scientific production, indicators based on external funding seem to be particularly valid for assessing research quality. However, with sinking success rates, the decisions of peers are becoming more arbitrary. It is common knowledge that today not all proposals worth funding can be funded. This in turn implies that there is quality that goes unmeasured by indicators based on external funding. Since success rates of external grant applications continue to decrease, we have more and more situations of 'disastrous competition for limited funds'. Hornbostel (2001) regards this situation as reducing the validity of indicators based on external funding.

A second limitation is that research funding signals only the success of single research proposals for a limited time. Gillett argues that decisions about grant applications have a considerably lower validity than a peer review of journal articles because they measure only research promises, whereas publications measure research output (Gillett, 1991: 254). This argument is not completely plausible. As I have stated earlier, it is usually not just the project idea which is assessed but the whole person and their earlier work. Albeit a researcher could dazzle with promises which they don't hold, they would be punished with grant failure next time if the project did not produce a scientific outcome. But even though Gillett overstates his case, it is certainly true that 'external research funding' refers only to the success of single projects and therefore cannot tell us whether a whole research programme of a scientist or research group is worthwhile to pursue. The latter decision is an important judgement made by reviewers when they evaluate research institutions (Roberts, 2003: 82–83).

Thus, both citation-based and funding-based indicators measure only certain aspects of research quality. In the case of citations, the problem is solved by using more than one quantitative indicator and additionally peer assessments. The same should be done when 'external funding' is used for assessments. When citation-based and funding-based indicators are combined in assessments, the problem of double counting, i.e. the fact that publications are not independent of funding, must be taken into account (Johnes and Johnes 1995).

Distorting factors One major objection against the use of citation-based performance indicators has always been that the indicator is distorted by many factors that are not quality-related. Examples of distorting factors are perfunctory citations, citations for 'bad' or 'incorrect work', exclusion of informal communication, the obliteration of well-known knowledge, the bias of certain types of papers (e.g. methods papers), or the arbitrary selection of references (for a list of these factors see MacRoberts and MacRoberts, 1989, 1996). For citation analyses, the problem of distorting factors can be solved by using statistical aggregations where the above-mentioned factors are treated as statistical error (Van Raan, 1998). As a rule of thumb, at least ten publications per year should be available as a basis for citation counts (Van Raan, 2000: 307–309).

For the performance indicator 'external funding', the problem of distorting factors and statistical aggregation has not yet been addressed. Some distortions seem to be negligible at the country level. Thus 'a country's investment in research' would influence the funding conditions of everybody in the same way. Factors such as 'Know-how about fundraising' can probably be handled statistically because it is likely that on higher levels of aggregation scientists who are inexperienced in grant-writing would be equally distributed.[3] Other factors — for example, 'sufficient recurrent funding' — cannot be assumed to simply 'vanish' at higher levels of aggregation. The universities in my sample differed in the amount of recurrent funding they could distribute for research, i.e. there were 'rich' and 'poor' universities. Moreover, there were also 'rich' and 'poor' university departments, which means that this factor seems to influence measurement at all levels of aggregation. Scientists at universities situated in big cities and/or with industry in the area have advantages with regard to the 'availability of collaborators'. Again, the factor distorts results even at the level of whole universities.

External research funding is used not only on aggregated levels but also on the individual level, e.g. in academic job decisions. Such a use can be justified if the indicators are interpreted by peers using their context knowledge. This might become problematic if peers have to judge the funding success of job applicants from other countries because they would usually lack the knowledge about other national science systems.

Field specificity of the indicator Citation practices differ across research fields. In order to make comparisons across research fields possible, citation indicators are normalised. International reference values, such as the average impact of journals or average citation rates of research fields are used (e.g. Schubert *et al*, 1989: 10; Moed *et al*, 1995: 399–400). However, even with all these normalisations, background knowledge of peers is still necessary for the interpretation of measurements.

The acquisition of external funding is also field-specific in many ways. Two different dimensions of field specificity can be distinguished. First, fields differ in their capital intensity, i.e. in the amount of resources that is necessary to conduct a typical research project. Since project grants mirror the field specificity, their size varies between fields. Whenever the cumulated research income is used as a performance indicator across fields, the varying size of grants will lead to significant distortions.

The second important dimension is the funding environment of fields. The interviews revealed huge differences in the availability of funding sources even within the relatively narrow field of experimental physics. Availability of funding depends on the actual topic and might therefore vary even within fields. In this context, the basic/applied character of research must also be taken into account. The more applied a researcher's work is, the more potential funding sources outside academia (industry, government, etc.) exist.

The recent practice of comparing 'equal with equal' in ranking universities by using the external funding success of whole departments assumes that the specialties with a low or high number of potential funding sources are equally represented in all universities. This assumption is not justified. There are universities (for example, technical universities in Germany) that conduct predominantly application-oriented research. These universities are clearly favoured in getting external funding.

All these factors call for a normalisation of external funding. However, the only attribute that has been used in a normalisation so far is the size of a unit. This normalisation leads to an indicator 'funds per researcher'. While this is an important normalisation, it is by no means sufficient. A normalisation by the attributes of research fields I have discussed above — i.e. capital intensity, basic/applied character and 'richness' of the funding environment — seems absolutely necessary in order to construct valid indicators. How this can be done is still an open question.

It becomes clear from this discussion that background knowledge is essential when using external funding as a quality indicator. But even then, it is questionable if this is still manageable on an aggregated level. For example, it can be doubted that RAE panel members are able to include the large amount of information for the many external grants they have to judge. Table 3 summarises the validity discussion.

Conclusions

The empirical investigation has confirmed most of the causal relationships hypothesized by Gillett (1991), Hornbostel (2001) and Moed *et al* (1998). Scientists' success in obtaining external funding depends on sufficient recurrent funding as enabling funds, a diverse funding landscape, and the scientist's track record. An examination of these and other factors that were found to affect success in

Table 3. Commonalities between the indicators 'citations' and 'external research funding'

Validity problem	Solution	
Only one aspect of research quality is assessed	Use more than one indicator	Interpretation by peers necessary
Distortion by factors that are not related to quality	Requires statistical aggregation	Interpretation by peers necessary
Is field-specific	Requires normalisation	Interpretation by peers necessary

funding demonstrated that, although the quality of researchers and their proposals were important and even necessary conditions, they are by no means sufficient conditions for success in external funding. Successful fund acquisition also depends on conditions that are only partly related to quality or have nothing to do with quality.

Thus, the current widespread simple use of external funding as performance indicators is not justified. It is even dangerous because it produces Matthew effects of funding: as stated before, those who get the most external funding are likely to get even more while others are crowded out of the system. A Matthew effect in research funding is not only unfair but also disadvantageous because it leaves part of the scientific potential unused.

The comparison with citation indicators showed ways of improving the current low validity of the indicator. While using more than one indicator and the inclusion of peer assessments can be relatively easily realised, statistical aggregation and normalisation encountered difficulties. In order to obtain a valid indicator, the following problems must be solved:

1. How can the amount of external research funding be normalised by fields and by the basic/applied character of research?
2. Which level of aggregation allows a statistical treatment of the other distorting factors?

Until these questions have been answered, extreme caution seems to be in order. The indicator should be used only at high levels of aggregation, and needs to be interpreted by scientists knowledgeable in the fields that are evaluated. A pragmatic solution for the present could be to use sufficiently long time-periods of evaluation and to distinguish only roughly between two groups: a group of scientists/research units which have no or hardly any external funding, and a group of scientists/research units which have continuous external funding.

Notes

1. It is often assumed that external funding is an input indicator of research (funds that enable research work) and therefore cannot be an output indicator. However, the widespread use of external funding as an output indicator is based on the peer evaluations that accompany grant applications.
2. For the scientist's reputation in their scientific community I extracted information from the interviews, using the following indicators: if and how often they review journal articles and/or grant proposals, their work on editorial boards, their work on committees of international conferences, invited talks at the international level, and visiting fellowships.
3. However, a situation of unequal distribution occurred after German unification where East German scientists lacked experience of grant-writing. German funding agencies reacted to this situation by temporarily setting up special rules to compensate for this disadvantage.

References

S Berghoff, G Federkeil, P Giebisch, C-D Hachmeister and D Müller-Böling (2003), *Das CHE-Forschungsranking deutscher Universitäten 2003*. Arbeitspapier 50 (Gütersloh: Centrum für Hochschulentwicklung).

D E Chubin and E J Hackett (1990), *Peerless Science: Peer Review and US Science Policy* (Albany, NY: State University of New York Press).

DFG (Deutsche Forschungsgemeinschaft) (2003), *Funding Ranking 2003. Institutions – Regions – Networks. DFG Approvals and Other Basic Data on Publicly Funded Research* (Bonn: DFG).

R Gillett (1991), 'Pitfalls in assessing research performance by grant income', *Scientometrics*, 22, pages 253–263.

S Hornbostel (2001) 'Third party funding of German universities: an indicator of research activity?', *Scientometrics*, 50, pages 523–537.

D F Horrobin (1996) 'Peer review of grant applications: a harbinger for mediocrity in clinical research?', *Lancet*, 348, pages 1293–1295.

J Johnes and G Johnes (1995). 'Research funding and performance in UK university departments of economics: a frontier analysis', *Economics of Education Review*, 14, pages 301–314.

G Laudel (2005) 'The "quality myth": promoting and hindering conditions for acquiring research funds', *Higher Education* (in press).

M H MacRoberts and B R MacRoberts (1989), 'Problems of citation analysis: a critical review', *Journal of the American Society for Information Science*, 40, pages 342–349.

M H MacRoberts and B R MacRoberts (1996), 'Problems of citation analysis', *Scientometrics*, 36, pages 435–444.

B R Martin and J Irvine (1983), 'Assessing basic research. some partial indicators of scientific progress in radio astronomy', *Research Policy*, 12, pages 61–90.

H F Moed, W J M Burger, J G Frankfort and A F J van Raan (1985), 'The use of bibliometric data for the measurement of university research performance', *Research Policy*, 14, pages 131–149.

H F Moed, R E de Bruin and T N van Leeuwen (1995), 'New bibliometric tools for the assessment of national research performance: database description, overview of indicators and first applications', *Scientometrics*, 33, pages 381–422.

H F Moed, M Luwel, J A Houben, E Spruyt and H van den Berghe (1998), 'The effects of changes in the funding structure of the Flemish universities on their research capacity, productivity and impact during the 1980s and early 1990s', *Scientometrics*, 43, pages 231–255.

A J Phillimore (1989), 'University research performance indicators in practice: the university grants committee's evaluation of British universities, 1985–86', *Research Policy*, 18, pages 255–271.

RAE (Research Assessment Exercise) (1999a), 'Assessment panels' criteria and working methods, Section III: panels' criteria and working methods (by UoA)', *RAE Circular 5/99*.

RAE (Research Assessment Exercise) (1999b), 'Guidance on submissions', *RAE Circular 2/99*.

G Roberts (2003). *Review of Research Assessment* (London: Higher Education Funding Council for England), <http://www.ra-review.ac.uk/reports/roberts.asp>

A Schubert, W Glänzel and T. Braun (1989), 'Scientometric datafiles: a comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields, 1981–1985', *Scientometrics*, 16, pages 3–478.

J Tognolini, K Adams and J Hattie (1994), 'A methodology to choose performance indicators of research attainment in universities', *Australian Journal of Education*, 38, pages 105–117.

G D L Travis and H M Collins (1991), 'New light on old boys: cognitive and institutional particularism in the peer review system', *Science, Technology, and Human Values*, 16, pages 322–341.

A F J van Raan (1996), 'Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises', *Scientometrics*, 36, pages 397–420.

A F J van Raan (1998), 'In matters of quantitative studies of science the fault of theorists is offering too little and asking too much: comments on: theories of citation?', *Scientometrics*, 43, pages 129–139.