Jochen Gläser[1]
Grit Laudel[1]
Sybille Hinze[2]
Linda Butler[1]

31.5. 2002

Impact of evaluation-based funding on the production of scientific knowledge: What to worry about, and how to find out

(Expertise for the German Ministry for Education and Research)

1. Introduction

In order to get 'value for money', public policy is increasingly relying on evaluation-based funding (EBF) of research. The times are gone when researchers could spend enormous sums of money by simply appealing to society's trust in science. Today, science must prove either that it contributes to solutions of societal problems or (at least) that basic research is of high quality.

The rapid growth of EBF fits into this general trend towards greater accountability of public policy. A funding practice that relies exclusively on trust in science has difficulties legitimising itself. Given the amount of public money spent on science these days, it is quite natural that the public expects science policy to show that the money is used to increase society's welfare. Additional stress on funding practices is produced by the overall perception that the sciences' need for funding is rapidly outgrowing society's financial capacity. More and more often, the decision has to be made on whom not to fund. Informed and accountable decision making on these matters requires public discussion on decision criteria.

One general effect of EBF is that it increases the pressure on researchers and research organisations to meet performance criteria. However, the question has arisen whether EBF actually promotes scientific research. There are voices of concern that point to a decreasing autonomy of researchers, to the expenses of EBF processes themselves, and to adaptation processes like over-publishing, risk avoidance, clinging to the mainstream etc. Not all these criticisms can be dismissed as complaints of 'losers' who weren't funded. Successful scientists complain too. More disturbingly, there is mounting empirical evidence on the adverse effects of EBF. For example, several investigations led to the conclusion that EBF favours mainstream, low risk research, and that the introduction of publication counts into funding formulas was followed by the emergence of a gap between quantity and quality of publications. Moreover, the last round of one of the older EBF processes, the British research assessment exercise (RAE), left even the winners with less money because more participants now know how to play the game.

Evidence of this kind poses an increasing challenge for both science policy and science studies. Science policy must admit the fact that its turn towards EBF, albeit comforting in terms of rationalisation and legitimisation of funding science, might have effects that are difficult to

[1] Research School of Social Sciences, Research Evaluation and Policy Project, Australian National University, Canberra, ACT 0200, Australia.

[2] Fraunhofer-Institut für Systemtechnik und Innovationsforschung, Breslauer Straße 48, D-76139 Karlsruhe.

detect and so far largely unknown in nature and strength. The full range of consequences of an ever more powerful political practice has yet to be explored. To date, science policy's main sources on effects of EBF practices are studies that explore researchers' opinions on these practices (see 2) or spontaneous complaints. Neither the continuation (or change) of existing EBF practices nor the introduction of new EBF mechanisms can be said to rest on solid knowledge about possible consequences of such decision making. This seems of special importance for Germany, where traditional practices of peer review have come under scrutiny (Nature 2002), and where the introduction of new EBF practices in universities and other research organisations is under way.

EBF of university research in Germany started only fairly recently. Performance oriented financing of universities and periodical assessment of research and teaching became part of the higher education law in 1998 ("Hochschulrahmengesetz" HRG, §5 and 6). Intensive discussions about how to implement EBF as well as discussions about how to perform evaluations are still continuing. Global budgeting ("Globalhaushalte"), which is a presupposition for formula-based funding because it grants the universities more autonomy in the use of financial resources, is being introduced. Models differ between the individual federal states in relation to the universities' actual degree of freedom (Federkeil and Ziegele 2001: 3). It appears EBF is being introduced both for funding universities and for resource distribution within universities, with different performance criteria being applied on both levels. However, these regulations have been introduced fairly recently or are even yet to be defined. In some of the federal states (e.g. in Niedersachsen and Baden-Württemberg) commissions or agencies were set up which should carry out or support evaluation activities. For example, in Niedersachsen a "scientific committee" ("Wissenschaftliche Kommission Niedersachsen") was founded which is since 1999 responsible for research evaluation in universities and non-university research institutions funded by the federal state Niedersachsen (Wissenschaftliche Kommission Niedersachsen 1999a). The Commission developed guidelines for the evaluation procedure (Wissenschaftliche Kommission Niedersachsen 1999b) and since 1999 about ten disciplines[3] have been evaluated. Evaluations are discipline oriented, include all relevant institutions, and are carried out in co-operation with them. Evaluations are based on "informed peer review" (ibid.: 3-4). Criteria to be considered are divided into three sub-areas: quality and relevance; effectiveness and efficiency; aspects of structural policy (ibid.: 5-6). Results derived and conclusions drawn will be commented on by the universities. The scientific committee gives recommendations for future planning to the Land and the universities. Universities will have to report after a defined period of time on how the recommendations were implemented (ibid.: 8). So far it is not clear what consequences will follow.

The introduction of EBF based on funding formulas for German Universities does not necessarily imply that similar procedures are applied within the universities. Whether or not universities introduce or have to introduce any formulas for internally distributing their resources is dealt with differently by the federal states and subsequently by universities. While no specific

[3] 1999: Chemistry; 2000: History, Biology, Civil engineering and Architecture , Electrical Engineering; 2001: Mechanical Engineering, Economics, Informatics, Vocational Training and Teacher Training; 2002: Physics, Law, Environmental Engineering.

internal resource allocation procedure for universities is fixed in the law of Baden-Württemberg and Niedersachsen, performance based resource allocation within the individual universities is required in Schleswig-Holstein. Other federal states regulated those details by "goal agreements" ("Zielvereinbarungen", e.g. Berlin, Niedersachsen, and Sachsen-Anhalt - see Federkeil and Ziegele 2001: 9). Intensive discussions concerning the development of EBF mechanisms in universities are going on at present, and knowledge about possible impacts of these mechanisms seems to be urgently needed.

The gaps in our knowledge that stifle science policy are caused by insufficient research on EBF. Science studies must face the hypothesis that there is an influential factor affecting knowledge production, which has only been investigated marginally. While ethnographic observations have revealed that factors of this kind affect knowledge production, little is known about whether and how the produced knowledge is different, how variations of EBF practice cause variations of knowledge production, how the impact of macroscopic EBF practices on knowledge production is mediated, etc. The persistent disinterest of Science and Technology Studies in these questions is due to the traditional complementarity of approaches, with institutionalist approaches giving no systematical account of what should be their dependent variable, namely knowledge production, and constructivist approaches giving no systematic account of what should be one of their independent variables, namely institutions. An instructive example of institutionalist approaches' neglect of knowledge production is the article on "the role of funding agencies in the cognitive development of science" (Braun 1998). Braun promises in the abstract to answer the following question: "by what means and processes is the distribution of funding resources transformed into scientific information and knowledge processes?" (ibid.:807). However, the question isn't answered at all, as is admitted in the conclusions: "It has been my intention to discuss some of the general features and dynamics which may influence politics and policy making in funding agencies as a major intervening variable for the cognitive development of science" (ibid.: 820). This kind of institutionalist approach has been declared as insufficient by Mayntz and Schimank (1998). Discussing institutionalist research on linkages between external expectations and science, they write:

> Important as the responsiveness of science is for the outcome of given linkages, an analysis of mediating institutions per se will rarely be able to assess it. For this, a research design would be needed that includes the performance level of the science system, something that is much easier where direct linkages, as between researchers and industry, are the subject. (ibid.: 753)

On the other side of the divide, the sociology of scientific knowledge has attracted much criticism for neglecting institutional structures, a deficiency that has been acknowledged by Knorr-Cetina (1995). Answering criticisms that laboratory studies "ignore the societal context in which laboratories operate as well as the political aspects of science", she writes:

> This critique, too, is warranted; but it would seem to be important mainly with respect to those aspects of the wider context and those relational properties of laboratories that recur in scientific work and influence the construction of knowledge. (ibid.: 162)

In spite of these criticisms and self-criticisms, and in spite of the weak knowledge base for political decisions on EBF mechanisms, no systematic effort has been undertaken to explore causes, effects and causal mechanisms that are at work in EBF.[4] The numerous studies and complaints about EBF practices may have nourished our suspicions, but they have not advanced our scientific knowledge on the subject. Therefore, a research program is needed that supports the growth of systematic knowledge on the impact of EBF practices on knowledge production in science. The aim of this article is to propose and substantiate such a research program.

We begin by reviewing the empirical evidence on the subject (2). The several studies that investigated EBF practices have paid little attention to the impact of these practices on knowledge production. Ethnographic observations of scientific practice have contributed to our understanding of adaptation processes to funding conditions, but have not provided conclusive answers to the question at hand. A review of the methodologies used in these studies leads to the conclusion that there is not much evidence we can actually rely on. We will therefore argue that there is enough material to become suspicious about effects of EBF, but no conclusive evidence. We will outline what should be investigated about EBF, and how this can be done. The research program we propose combines several methods in order to address the micro-macro link between institutional conditions of actions such as EBF and outcomes of single research processes (3).

2. Effects of EBF

In this chapter, we will review empirical findings that are related to the question of how EBF practices affect the production of scientific knowledge. Empirical investigations of EBF focussed on three areas. Not surprisingly, studies have been conducted on the peer reviewed funding of projects, which is one of the oldest and most widespread practices of EBF (2.1). A second form, the EBF of research organisations, is a more recent phenomenon. One scheme that has been applied regularly since the early 80s, and that is probably the most consequential practice of this type, is the Research Assessment Exercise (RAE) in Great Britain. For more than a decade, this EBF practice has attracted the attention of both affected scientists and analysts. Most literature on EBF of organisations refers to the RAE. However, some studies on new EBF practices in Germany have been conducted recently (2.2). A third EBF practice is formula-based funding, e.g. the allocation of a part of an organisation's funding on the basis of quantitative performance measures. This practice is relatively recent but rapidly spreading. Up to now only few studies have been published on formula-based funding (2.3).

An important question that has found only limited attention so far is how research organisations and researchers adapt to EBF practices. We will review both micro-sociological studies and

[4] The only systematic attempt to explore the relations between funding and knowledge production has been a workshop of the US National Science Foundation on ‘Funding and knowledge growth’. However, the studies presented at this workshop did not differentiate between recurrent funding and EBF. Knowledge growth was treated quantitatively rather than in terms of knowledge content, and causal mechanisms that mediate between funding and knowledge growth were not explored. The introduction to a theme section of “Social Studies of Science” made clear that the workshop led to “challenges for science policy studies” rather than answering questions (Cozzens 1986).

investigations of changing practices in research management with regard to the adaptation processes that take place on these levels (2.4). Finally, the credibility of all these empirical findings has to be assessed. By discussing the research designs and the methods applied in the studies reviewed, we will reach conclusions about the extent to which there is knowledge we can build on (2.5).

2.1 Peer reviewed funding of projects.

To apply for funding by submitting grant proposals to an agency that, in turn, submits them to peer review, has become an inseparable part of many scientists' everyday practice. While the extent to which researchers are dependent on grants varies greatly between countries and within countries, between scientific fields and between organisations, there are few researchers who neither have the opportunity nor feel the need to apply for grants to finance their research.

While several studies have investigated the peer review of project proposals, few findings on its impact on knowledge production exist. The studies on peer review of project grants were not interested in how this EBF practice affects the production of scientific knowledge. They were framed in the tradition of the 'old' Mertonian institutionalist sociology of science and thus primarily concerned with the role of norms in peer review practices. This led to an empirical interest in the procedural justice and the reliability of peer reviewed grant distribution. The studies investigated whether 'old boys' networks' affected decisions, whether the general practice was fair, if applicants' reputation influenced the outcome of decisions, etc. (Cole, Cole, and Simon 1981; Chubin and Hackett 1990; Neidhardt 1988). Owing to this focus on the 'political correctness' of peer reviewed project funding, its role in the production of scientific knowledge has escaped the attention of investigators. The blindness of the constructivist studies for this aspect of peer reviewed grant distribution is due to its supportive role. Peer review is itself a process of individual knowledge construction (Hirschauer 2002). Constructivist studies missed that process because it takes place outside laboratories and outside scientific controversies. Peer review differs from the 'normal' production of scientific knowledge through research and communication in that it links individually produced knowledge claims to the scientific community's collective knowledge production. Because of this function, peer review can be said to provide a link between the scientific community (represented by peers) and the individual researcher (Gläser 2002). Owing to this intermediary position, the content of peer review has escaped the attention of both the sociology of scientific knowledge and the institutionalist sociology of science.

In spite of their different foci, the studies on fairness and reliability of peer reviewed grant distribution nevertheless gave some hints to its impact on knowledge production. Three studies in the United States obtained figures about the fate of proposals that were not funded (Chubin and Hackett 1990: 60-65). The proportion of scientists who stopped a particular line of research when funding was initially denied varies for the three studies between one third (in a study on funding by the National Cancer Institute), 37% (in a study of funding by the National Institute of Health) and 48% (in a study on funding by the National Science Foundation). It cannot be said, of course, whether the scientific knowledge on whose production these projects were aimed was produced elsewhere or not at all. The assumption that the fate of rejected proposals is of less

importance because these are the proposals of lesser quality anyway, though comforting, is wrong: Owing to the worsening success rate of grant proposals, it can no longer be asserted that good proposals are accepted and bad proposals are not. Chubin and Hackett quoted a member of the American Cancer Society who stated that there are "many fine proposals but funds are so scarce that we must eliminate two thirds of the truly outstanding proposals" (ibid.:73).

Another significant impact of the rejection of grant proposals is that this rejection may end a scientific career. Chubin and Hackett quote a scientist who didn't continue research after her proposal was disapproved (Chubin and Hackett 1990: 64) and an observer who stated that

> [he] has witnessed many cases in which established, good quality researchers who somehow happened to have remained in research positions rather than tenured positions, suddenly found themselves in a situation where they lost simultaneously their research support, salary and position. (quoted ibid.)

But it is not only the failure of grant proposals that affects knowledge production. Respondents' answers to one of the surveys indicate that the mere existence of a system of peer reviewed grant distribution changes the behaviour of scientists who are more or less dependent on that system. Chubin and Hackett report that 60% of the respondents to one survey believe that reviewers are reluctant to support unorthodox or high risk research, while only 18% disagree with that assertion (ibid.: 69). An adaptation strategy that was also reported is that "scientists maintain a mix of more and less risky work, insuring that they will always have a stream of 'bread and butter' papers while attempting to do 'the breakthrough stuff'" (ibid.).

This perception of peer reviewed grant distribution was confirmed for Germany's "Deutsche Forschungsgemeinschaft", the most important funding agency for university research (Neidhardt 1988: 136; Nature 2002); for grant distribution in Canada (Berezin 1998) and in the United Kingdom (Horrobin 1996; Travis and Collins 1991: 336). Similarly, interdisciplinary research has been perceived as having less chance in peer reviewed grant distribution (Travis and Collins 1991; Berezin 1998). It is not important for the question of the impact on knowledge production to prove this assumption is true. The only thing that matters here is that scientists believe that it is true. According to the Thomas theorem "If people believe a situation to be real, it is real in its consequences" (Thomas and Thomas 1928), scientists who believe that risky, unorthodox, interdisciplinary proposals have less chance of success will tend not to submit them.

The assumption that unorthodox and highly innovative projects are not suitable for peer reviewed project funding is matched by the outcome of a comparative study on the impact of short-term versus long-term funding. Bourke and Butler compared the impact of short term research funded by peer reviewed grants (3 to 5 years) with that of research funded by long-term grants (for research centres, 6 to 9 years) and with research conducted by permanent research organisations. They found that the research that was funded with long-term grants or continuously had a much higher impact (measured by citations in the journal literature, Bourke and Butler 1999). While the causes of this association were beyond the scope of the study, it suggests that the time regime underlying peer reviewed funding of projects may have adverse effects on the quality of work.

A more general problem with peer reviewed project funding is that it usually lays great emphasis on the applicant's track record, i.e. the expertise regarding the project's field as demonstrated by earlier publications. This makes it difficult to start new lines of research (Mukerji 1989: 103-104).

Finally, peer reviewed grant distribution also affects knowledge production indirectly by consuming time and money that would otherwise be available for research. Quite naturally, comments on the amount of time researchers spend writing grant proposals vary greatly. Figures such as at least 50% (quoted in Chubin and Hackett 1990: 63), up to 40% (Lederman 1993) or 20 to 50% (Horrobin 1996: 1294) are mentioned beside general statements that peer review is time consuming (Hume 1994: 28, Wessely 1998: 302-302). Statements about the costs of peer reviewed grant distribution are even less precise (see 2.2). However, it becomes clear from these statements that a substantial part of resources are invested in getting resources and thus diverted from research itself.

2.2 Peer reviewed funding of organisations

The EBF of organisations occurs in two different forms: as formula-based funding that uses quantitative performance measures, and as funding based on peer review. We discuss the peer reviewed funding first. The most prominent practice of peer reviewed funding of organisations is the British research assessment exercise (RAE). The British Higher Education Founding Council (HEFCE) distributes money according to university departments' quality rating on a seven-point scale, which is assessed every five years. Assessment criteria have changed, but have always been a combination of input indicators (staff numbers, grants acquired) and output indicators (PhD students, publications).

Several studies have been conducted to assess the RAE.[5] The most comprehensive study was one conducted on behalf of HEFCE (McNay 1997). By applying a variety of methods, it investigated whether the RAE had led to changes in policy, management, staffing, or process and content of research. The last of these topics, which is relevant here, was investigated by two questionnaire surveys mailed to heads of units within universities (140 usable responses) and to academic staff (393) within 15 institutions. Respondents were asked to rate the extent of agreement/disagreement with a number of propositions (ibid.: 4).

The McNay study was conducted after the 1992 RAE. A study similar to his (but apparently independent of HEFCE) was undertaken after the following round of the RAE in 1996 (Talib 2001). A survey questionnaire was mailed to 1000 academics from "various disciplines". A total of 305 replies was received (ibid.: 33). As in the McNay study, respondents were asked to indicate their level of agreement/disagreement on a scale of 1 to 6 (ibid.: 32-33). "A number of statements were borrowed verbatim from the McNay study to enable the direct comparison of results." (ibid.:33)

---

[5] Additionally, the RAE has triggered a passionate discussion among academics, mostly fuelled by complaints about adverse effects on the RAE on various subjects. Many concerns are in accordance with the questions asked by the more systematic studies reviewed in this section. Generally, these discussions and complaints provide only anecdotal evidence that is difficult to assess. For examples of this discussion, see Velody (1999).

The biggest study on the RAE was conducted by Evaluation Associates Ltd (1999). Due to "criticism of the RAE's impact on interdisciplinarity"[6] (ibid.: 2) from the institutions assessed, a study was initiated focusing "on the assessment of interdisciplinary research by the RAE, and any disincentives that might arise from that " (ibid.: 2). A large-scale survey was carried out. In a first wave about 11,000 researchers were sent a questionnaire, non-respondents were followed-up by sending them again an "abbreviated" version of the questionnaire. In addition a shorter questionnaire was mailed to a "booster"[7] sample of 5330 researchers. In addition panel members were sent a questionnaire too, again non-respondents were sent an abbreviated version of the questionnaire. Respondents were asked about their agreement or disagreement with various statements about the RAE. The response rate was 37 per cent (ibid.: 3) among the researchers and 62 per cent among the panel members.

Another analysis of the RAE's effects was based on qualitative interviews (Henkel 2000). Henkel conducted interviews (a total of 327) with academics (from seven disciplines) and managers from 15 British universities (ibid.: 23-24).

For reasons that will be discussed later, the results of these studies are difficult to interpret. The following impacts of the RAE on knowledge production have been reported:

1) Improved quality of research

The views on whether quality of research has improved are mixed. McNay found 81.3 per cent of heads of units agreeing that research quality was better than five years ago, without explicit reference to the RAE. (McNay 1997: 43). The responses from academic staff indicated an improvement of their own "public research output" (64.0 per cent agreement, ibid.: 44). The agreement to a similar statement after the following RAE was much lower – only 38.8 percent agreed to "the RAE has had a positive effect on my research output" (Talib 2001: 36). The general comments on research quality given by staff were rather critical, with only 34.3 per cent of staff agreeing with the statement "On balance, RAE has improved the quality of research conducted in higher education" in the McNay study (1997: 46), and an even lower agreement (27.9 per cent) to a similar statement in Talib's study (Talib 2001: 37). Although the Evaluation Associates study was mainly focussed on the matter of interdisciplinarity based on the "booster sample" a "snapshot of researchers' opinions on how the RAE affects their work, through asking them to agree or disagree with a series of statements" (ibid.: 17) was implemented. According to that sidetrack of the study 52 percent agreed strongly that RAE "encouraged researchers to concentrate on the quality of their research" (ibid.: 17).

2) Interdisciplinarity

Similar to the question of quality, no clear picture can be derived from statements on interdisciplinary research. In McNay's study, 46 per cent of heads agreed that interdisciplinary

[6] In the report it was stated that interdisciplinarity as well as multidisciplinarity is covered by the use of the term "interdisciplinary research" as most researchers do not clearly distinguish between the two (Evaluation Associates Ltd 1999).

[7] The "booster" sample was created in order to "ensure adequate coverage of all subjects" (ibid.: 2).

work is hindered by the discipline base of assessment and assessment panels, but only 20 per cent of staff agreed (ibid.: 25). According to Talib (2001: 37), 60.2 per cent of researchers were of the opinion that the RAE has a negative effect on interdisciplinary research, but only 23.5 per cent reported doing “less interdisciplinary research than 6 years ago”. According to the study conducted by Evaluation Associates Ltd, 53 per cent of the researchers believe that the RAE at least slightly inhibits interdisciplinary research, and about 48 per cent of panel members share this view. This statement gains additional weight because the proportion of researchers who believe that the RAE inhibits interdisciplinary research is particularly high among those who are most heavily engaged in this kind of research (ibid.: 15). The qualitative study conducted by Henkel led to the conclusion that “both sociologists and economists felt that the RAE discriminated against interdisciplinary research and that no credit was given for publications in the journals of other disciplines.” (Henkel 2000: 141).

3) Risk-taking

According to the McNay study, 46 per cent of heads of units saw the RAE as encouraging a more conservative approach to research and a reduction in open-ended ‘blue skies’ work driven by curiosity. However, only 20 per cent of staff agreed that they have avoided more speculative or more risky research topics (McNay 1997: 25). Talib reported an agreement of 35.4 per cent to that statement (Talib 2001: 37).

4) Long-term versus short-term research

A minority of respondents to the questionnaires and some of the interviewees in Henkel’s study stated that the RAE endangers long-term research. 25 per cent of respondents in the McNay study and 31.7 per cent of respondents in the Talib study agreed that they had “avoided some research topics or projects because they would have taken a long time to complete” (Talib 2001: 37). Additionally, 45.9 per cent of respondents in the Talib study agreed that their “research topics are influenced by the RAE time scale” (ibid.) – a statement, however, that is difficult to interpret. In the "snapshot" part of the Evaluation Associates study, 11 per cent of the researchers stated that the RAE provides "little encouragement for longer-term research” while according to 63 per cent of the researchers "early research results are very important". This last statement is linked to researchers’ wish to get research results published before the next RAE exercise is performed (Evaluation Associates Ltd 1999: 17). Several of Henkel’s interviewees stated that the RAE hinders long-term research. Moreover, Henkel referred to a report from Cambridge University Press that “noted that publication proposals were shifting from the longer and more matured forms of publication to shorter-term projects, apparently to meet the time requirements of the RAE”, (Henkel 2000: 139).

5) Scientific innovation

In the McNay study, 45 per cent of heads agreed that the RAE has inhibited new work because of the time it took to establish a reputation; 25 per cent of staff avoided new research because of the time problem (MacNay 1997: 25).

6) Basic versus applied orientation

The studies provide an ambiguous picture of the RAE's impact on the basic versus applied character of research. In the MacNay study, 12 per cent of heads of units believed that the balance has shifted to favour 'pure' research, and 18 per cent of staff acknowledged the shift (McNay 1997: 25). Talib found 19.6 per cent of respondents agreeing that their research has "moved away from the applied end of the spectrum towards the more basic pure research" (Talib 2001: 36). The reverse question (about moving from basic towards applied research) was not asked in these studies. The "snapshot" part of the Evaluation Associates study provided an asymmetrical picture, with 24 per cent of the respondents stating that they were encouraged to carry out basic research, while 22 felt encouraged to focus on user oriented activities. (Evaluation Associates Ltd 1999: 17).

7) Diversity

The RAE is regarded by Henkel's interviewees as negatively affecting the diversity within social sciences (Henkel 2000:141). This is in accordance with a survey conducted by Harley and Lee (1995) who reported a strong pressure towards mainstream research in economics, and a survey conducted by Harley and Lowe (1999), which led to similar results for sociology, psychology, marketing and accounting.

8) Publication behaviour

All three quantitative studies hinted at changes occurring in academics' publication behaviour. According to both studies, many academics feel a higher pressure to publish – 42 per cent in the McNay study, 64 per cent in the Evaluation Associates study, and 78.4 per cent in the Talib study agreed (Talib 2001: 36). Furthermore, 24.5 per cent of respondents in the McNay study and 40.4 per cent in the Talib study agreed with the statement "Because of RAE time scales I have published some outputs at an earlier stage than I would prefer" (ibid.: 37). A change in publication behaviour was also observed by Henkel (2000: 139). The change of publication practices was confirmed by the observations of publishers.[8] The managing director of Cambridge University Press reported that they

> …noticed a marked increase of proposals that we would regard as 'book-making' – collections of essays or articles rather hastily and unconvincingly put together, unrevised theses and monographs on slight or ephemeral themes. (Mynott 1999: 129)

He also described that authors began to compromise on quality of their books when approaching the deadline of the RAE:

> It is common for the sort of academic works we publish to go through several drafts, often over a period of some years, and to be substantially revised, developed and improved in this process. But we found that from the beginning of 1995 many authors were unwilling to complete the sort of revisions we (and

[8] The impact of the RAE on publication behaviour was investigated in a separate study based on 72 survey responses from journal editors (Talib 2000). However, this study provided no meaningful results (see 2.5).

> our series editors and advisors) would normally insist on to ensure that the final publications were of the appropriate standard. We resisted this, but felt ourselves under considerable moral pressure since careers and departmental resourcing were clearly expected to be affected. (ibid.: 129-130)

Similar effects have been reported by a representative of Sage Publications Ltd (Walford 1999: 142-143). However, these observations provide only anecdotal evidence, and neither the scope nor the impact of the reported trends can be determined.

An interesting pattern occurs from the reported effects on knowledge production: a certain (often strongly varying) percentage of heads of units, panel members, and researchers agree that there is an effect of the RAE, while others don't. This implies, on the other hand, that between 30 and more than 80 per cent of respondents did say that there are no such effects. The results' associations with other attributes such as institutions, field, or respondents (professional) age are not analysed. Thus, it seems likely that the reported effects occur, but it is impossible to conclude under what conditions and with what strength.

Another consequence of the RAE became obvious after the last round in 2001. The number of highly ranked university departments significantly outgrew the increase in funding. As a consequence, even the departments that are ranked highest face cuts in their funding (Cohen 2002). Similar to the situation in peer reviewed project funding, the quality criterion does not work properly any longer because there are to many high quality applicants.

While it was highly successful, the university departments' adaptation to the RAE procedure was also costly. It has been estimated that the RAE expends 10% of available research funding on the process itself (Schnitzer and Kazemzadeh 1995: 58). Two observers of the RAE's development in the UK conclude that the actual expense might even outgrow the utility (Geuna and Martin 2001:30-32). Even if this does not happen, there is no doubt that a significant part of the money is not spent for research, but for the evaluation process.

Another process of EBF of research organisations has emerged in Germany, where non-university research institutes undergo regular peer reviews. The aim of the peer reviews is to assess whether further state funding of the institutes is justified. The first systematic use of this approach was the evaluation of East German state-financed non-university research after German unification (Mayntz 1994). Since then, decision making about further funding of state-financed non-university research institutes has become a common practice. Institutes of the Leibniz-Society are now being evaluated every five years, and further funding of these institutes depends on a positive evaluation (Roebbecke and Simon 2001). However, the impact of this new EBF practice on knowledge production has not yet been investigated. This is not surprising given the fact that this practice is fairly recent.

2.3 Formula-based funding

The growing demand for research evaluation cannot be fully catered for by traditional peer review that has only a finite capacity. Researchers can devote only a limited proportion of their time to peer evaluation before their own work begins to suffer. As a result, there has been an increased use of quantitative performance indicators. One more recent trend in this respect is the

application of funding formulas, where the parameters included in the formulas should reflect governmental goals or demands. Depending on their performance universities receive a budget or a part of the overall available resources from the government (Liefner 2001). Formula based funding is being used for allocating the resources for teaching as well as for research. Usually both components are treated separately. Examples of the application of formula-based research funding can be found since the early 1990s in Australia and Nordrhine-Westfalia (Jensen and Neuvians 1994, Andersen et al. 2001) since the late 90's in Baden-Württemberg (Körber-Weik and Schmidt 2000), in the Netherlands, Denmark (Schnitzer and Kazemdaeh 1995), and in Finland (Adam 2002, and personal communication of Kari Raivio, rector of the University of Helsinki). Formula-based funding has also found its way into research organisations, which have begun to internally distribute money on the basis of formulas (Schäfer and Sikora 2001; Marginson and Considine 2000; Van Leeuwen and Moed 2002: 261-264). Since it is a relatively recent phenomenon, effects of formula-based funding have not yet been subject to many empirical analyses. One of the oldest practices of formula-based funding is probably the Australian funding scheme for university research. The introduction of the Research Quantum in the early 1990s saw the first distribution of research funds to universities based on a formula encapsulating a number of performance measures (graduate student numbers or completion rates, research income, and publications). The Research Quantum was subsequently replaced by two new funding schemes – the Institutional Grants Scheme and the Research Training Scheme. Both use a formula comprised of the same three elements, though the weighting for each element varies between schemes.

No systematic investigation of this EBF practice's impact on knowledge production in Australian Universities has yet been undertaken. A study on "the impact of performance indicators on the work of university academics", that was based on questionnaires and interviews, reports an increasing pressure felt by respondents, and a changed approach to research (without further qualification) reported by 63 per cent of respondents. Comments from interviews suggest a trend to more easily publishable research and more publications of lesser quality. (Taylor 2001: 49-53) A more general study on changes in Australian universities has reached the following conclusions:

> We want to summarise some of the problems that result: first, the bias in favour of research quantity rather than research quality; second, the bias in favour of short-term research performance not long-term research capacity; third, the bias in favour of track record and conventional approaches, rather than new researchers and emergent approaches. (Marginson and Considine 2000: 171)

A recent investigation of Australia's scientific output gives raise to concerns about the continued use of formulas in their existing forms (Butler 2001, 2002). It documents a significant increase in the country's journal output, accompanied by a worrying decrease in the relative international impact of these publications as measured by citations (figure 1 and 2). The timing of this productivity increase in relation to the introduction of funding formulas suggests that there is a causal relationship.

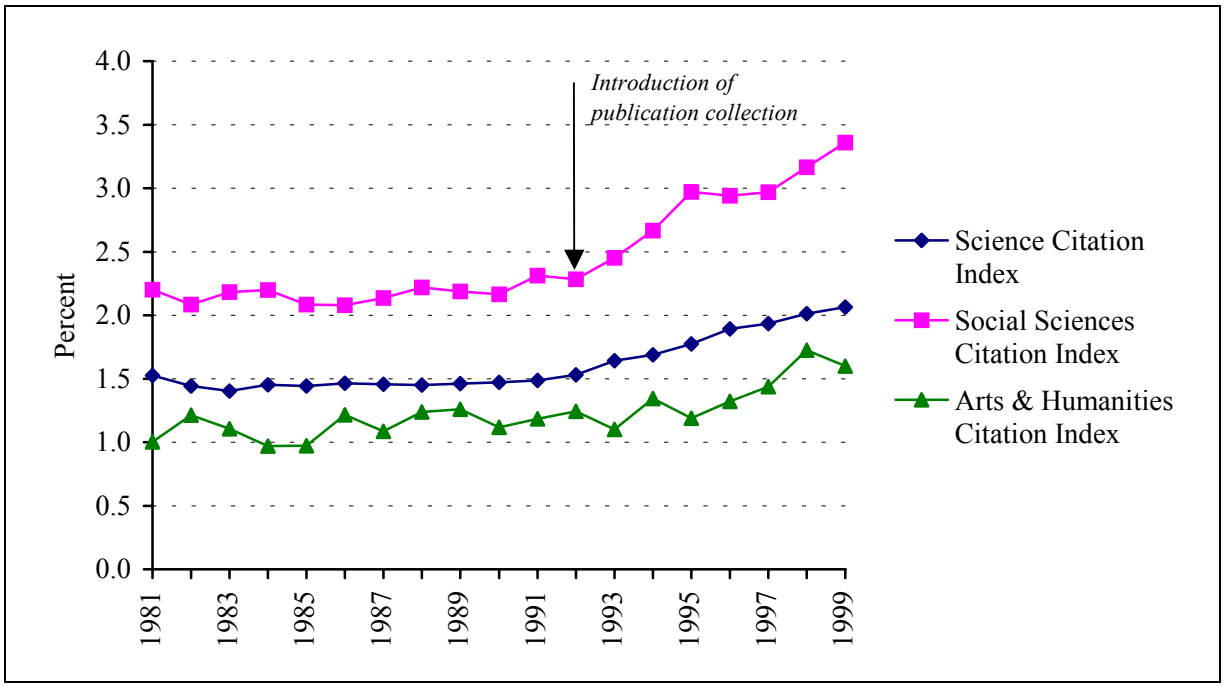

Figure 1: University share of publications in ISI's three main indices, 1981-99

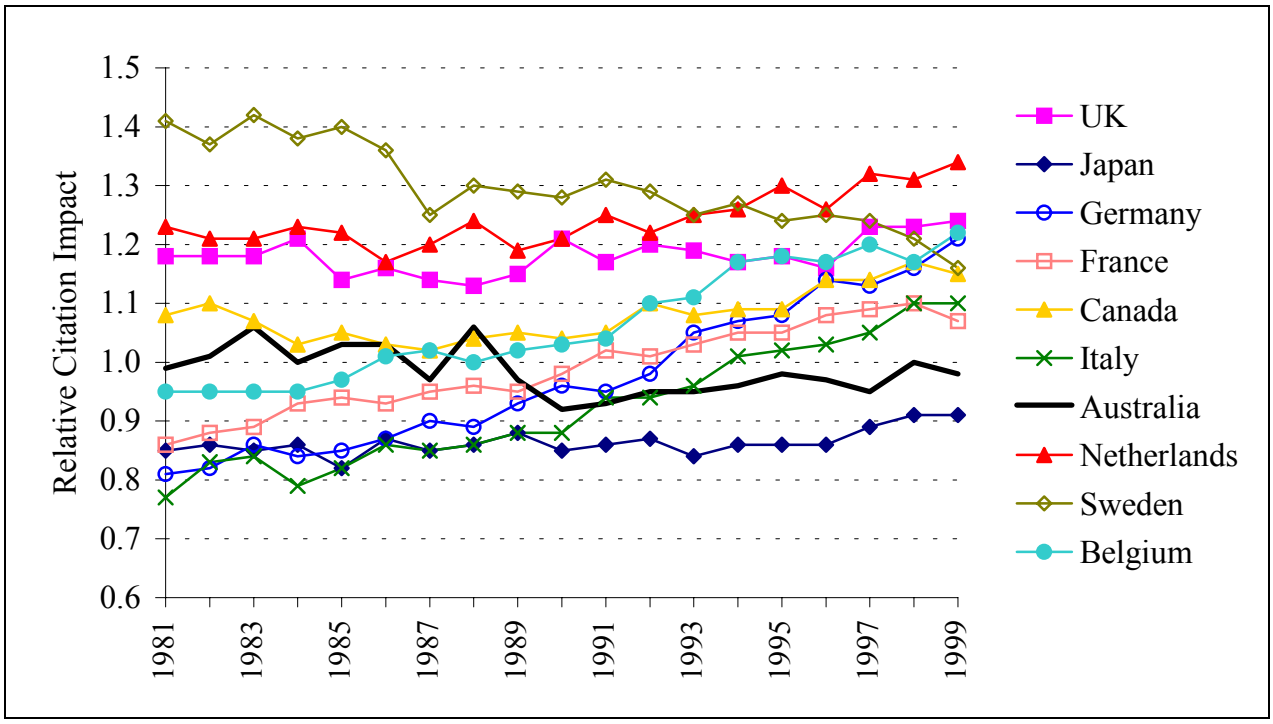

Figure 2: National trends in relative citation impact, selected OECD countries, 1981-99[9]

Figure 1 documents a significant increase in the quantitative output of Australian research after the number of publications had been linked to funding. The comparison of Australia with nine OECD countries provided by figure 2 demonstrated that the impact of this research fell behind

[9] Source: National Science Indicators database. The Relative Citation Impact (RCI) compares a country's share of citations with its share of publications. An RCI of one indicates equality of the two shares. An RCI of greater than one indicates the country's publications are attracting more citations than would be expected for the size of their output.

nearly all these countries. A comparison of two universities' output provides further support for the assumption that the coupling of increasing quantity and decreasing quality is due to the introduction on quantity-based funding formulas. The University of Western Australia introduced a formula for distributing a significant proportion of research funds, one major component of which was a publication count. The formula was much more sophisticated than the government's Research Quantum exercise, with many more publication categories and different weightings for each category in the different fields of research. Nevertheless, the formula was based on quantity, not quality. In direct contrast, the University of Queensland strategy was to instigate a strong recruitment drive targeting the brightest young researchers, including a significant number from overseas, and providing them with a strong resource base.

The relative citation impact (RCI) for the university sector declined through the 1980s and in the 1990s stabilised at just below unity (figure 3). The RCI for the University of Western Australia followed a similar, though more exaggerated trend. Its RCI declined from a peak in 1985 and declined at a faster and longer rate than for universities in general, though it has regained a measure of the lost ground in recent years.

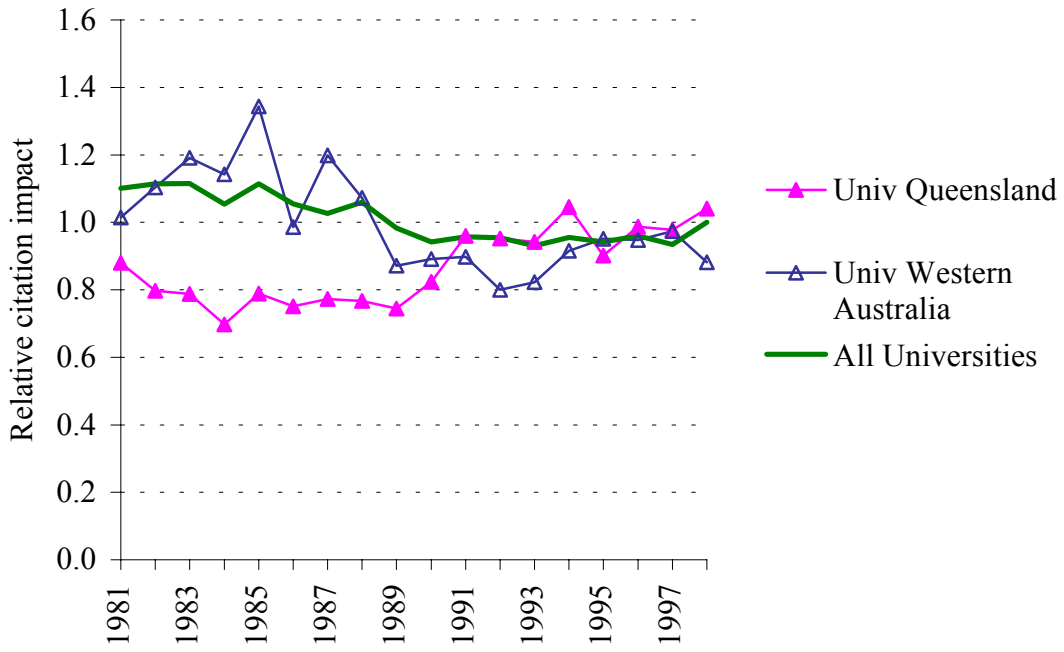

Figure 3: Comparison of two Australian Universities

In stark contrast, the University of Queensland's RCI remained steady but very low at less than 0.8 through until the end of the 1980s, followed by a significant improvement in the 1990s. It exhibits none of the decline shown for the sector as a whole or for the University of Western Australia.

A qualitative study conducted in six universities from four countries explored university professors' expectations regarding a complete change to funding based on output-indicators Liefner (2001). Liefner covered effects for teaching as well as research and distinguished between supporters, opponents and undecided. As the number of interviewed professors is rather small (n=53) results have to be interpreted with care, however supporters as well as opponents

do expect a shift from basic research towards more applied research. They also believe that research results will be transferred to the market more quickly. Opponents and undecided participants in the study expect deterioration of the quantity and in particular of the quality of basic research. With regard to publication activities, the expectations of the interviewed professors are similar to the findings mentioned above: while a growing number of publications is expected, it is feared that the quality will diminish. The latter expectation is, of course, held primarily by the opponents of output based funding (ibid.: 206-210).

2.4 Adaptation processes

Adaptation processes to EBF can be expected to occur at two levels. Research organisations that are subject to EBF might adopt their management strategies and practices in order to improve their performance in evaluations. Researchers are likely to respond threefold: to the EBF of their projects, to the EBF of their organisations as they perceive it, and to changes in their organisations' management practices (figure 4).

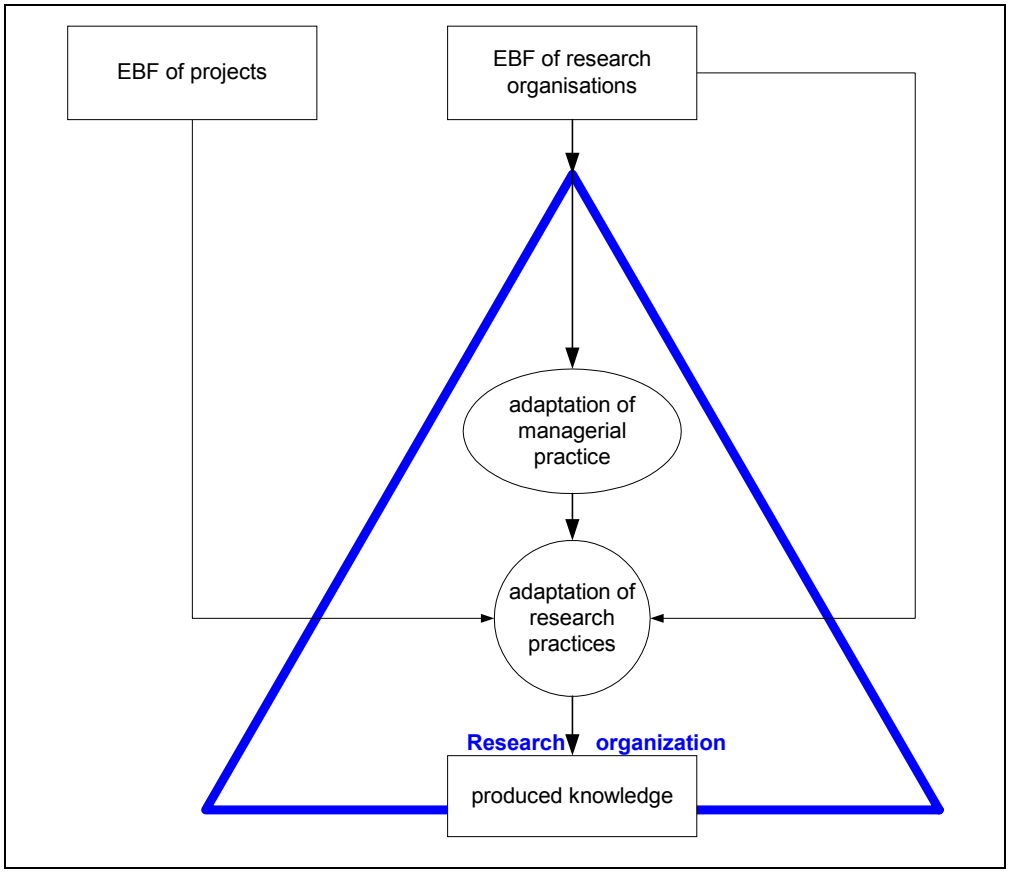

Figure 4: Adaptation processes to EBF

In the literature on EBF practices that has been reviewed in the previous sections, it has repeatedly been stated that researchers adapt to EBF practices.[10] These statements are to a certain extent supported by ethnographic studies of scientific practice. In her observation of laboratory practice, Knorr-Cetina has found that researchers' decisions on project design and research

[10] A study conducted by Luukkonen (1995) reported little or no adaptation processes to evaluations. However, this study has serious methodological weaknesses (see 2.5) that had us to question its validity.

conduct are influenced by a variety of non-epistemic factors that are often idiosyncratic to the research setting (Knorr-Cetina 1981). Among these factors are power structures of the research organisation, the available equipment, costs of energy and materials, future career opportunities, etc. From these observations it can be concluded that scientists indeed adapt the content of their research to the opportunities and constraints provided by their (scientific and non-scientific) environment. Some of these opportunities and constraints are produced by EBF practices. More direct conclusions are impossible because the impact of EBF on knowledge production is not demonstrated. For example, Knorr-Cetina reports the following observation:

> When the group's leader returned from Washington, the scientists not only changed the title of the grant proposal that had occasioned her visit, but also rewrote a substantial part of its content. (Knorr-Cetina 1981: 81)

More generally, she observed that the content of grant proposals is subject to negotiations between scientists, peer reviewers, and funding agencies (ibid.: 88-89). However, no information is given about the extent to which these negotiations and the accompanying changes in grant proposals lead to changes in the content of research projects (a possible second step in the adaptation process). Furthermore, it has not been observed whether the (possibly changed) research project leads to results that are different from the original target. Admittedly, it would be extremely difficult to identify such changes. However, without this additional information the ultimate impact of adaptation on knowledge production cannot be determined.

A more detailed study of the adaptation of funding proposals to the requirements of EBF was conducted by Myers (1990). Myers had collected all major drafts of two grant proposals (one original, one for continuation of funding) written by two biologists. He analysed the changes proposed by readers in handwriting on the drafts and the actual changes, i.e. the differences between the versions. Additionally, he conducted interviews with the authors (ibid.: 43). Contrary to the thesis of Knorr-Cetina that knowledge production is ultimately changed by adaptation to funding conditions, Myers describes the writing of both grant proposals as an adaptation of the proposals' *texts* to criticisms of reviewers and 'trusted assessors' rather than as an adaptation of the *research project* for which funding is sought. While the projects' aims and methods remain unchanged, descriptions are changed to improve the projects' fit into the fields.

> Each [author] shows, in his last version, a closer fit between his work, as he presents it, and his discipline, as he presents it. Both strategies shown in these processes of revision are attempts to deal with increased competition for health-related research funds by relating the proposed work to the consensus in the field. (ibid: 53)

Myers' study indirectly confirmed findings on the 'mainstreaming' effects of peer review by demonstrating that researchers often formulate their own aims and approaches more cautiously (as one possibility among others) and reduce the emphasis on their projects' innovations. However, according to his study, peer review changes the style of applications rather than the content of scientific work.

Another ethnographic observation was initiated with the aim to study "the effects of performance indicators on the actual conduct of science" (Law 1994: 35). Unfortunately, the observer became

distracted, and almost no information concerning this question is provided. There is only a cursory confirmation of risk avoidance in peer reviews. A reviewer stated: "Our committees won't fund *anything* risky." (ibid: 58, emphasis in original). Similarly, an ethnographic study conducted by Mukerji (1989) could have been of high relevance because the observed oceanographers were almost completely dependent on peer review grants. However, this study's focus was scientists' dependency on the state, and the question of how knowledge production was affected played almost no role.

While it has been established with some clarity that researchers indeed adapt their practices of knowledge production to the availability of resources, no ethnographic observation so far has provided sufficient data about how researchers adapt to EBF practices and how this adaptation changes the knowledge that is produced.

Adaptation processes on the organisational level have been investigated in the studies on the RAE. An overall improvement in managerial practices' support for research was reported by McNay (1997: 8, 11-12, 23). This trend seems to be the main benefit of the RAE. Beside this general trend, several adaptation processes that are consequential for the content of knowledge production were identified. Firstly, the pressure on researchers who did not meet the performance criteria of the RAE has been increased. The pressure put on organisations by the RAE was transmitted to the researchers. Secondly, hiring practices seem to have changed. The hiring of highly productive staff prior to the RAE deadlines doesn't play the role that was expected (McNay 1997: 13). However, one quarter of the unit heads reported that they lost staff specifically because of the RAE, and one quarter reported recruiting staff specifically because of the RAE (ibid.) Moreover, an increased emphasis on RAE performance criteria was reported with regard to hiring practices in economics (Harley and Lee 1997: 1450-1453) and in a study across several disciplines (Henkel 2000: 122). Thirdly, organisations seem to reshape their research profile on the basis of strategic considerations concerning the RAE. The strategies applied seem to vary: Henkel reports both a concentration on the core business, i.e. on a department's strength, and an increasing emphasis on size because of the perception that a 'critical mass' is necessary in order to get high grades in the RAE (ibid.: 126).

Australian universities responded to formula-based funding by more or less mirroring this practice within the university. Each of the 17 universities investigated by Marginson and Considine has established a formula-based system of distribution from the centre to faculties/schools (Marginson and Considine 2000: 149). Similarly to the changes triggered by the RAE, universities often try to meet external performance criteria better by applying them internally. In Germany, universities' adaptation to external EBF by developing internal ones is still under way.

2.5 'Robustness' of the findings

The review we have undertaken in the previous sections may well lead to the conclusion that there is already substantive knowledge about the impact of EBF on knowledge production. However, before we can draw such a conclusion we have to evaluate the 'robustness' of these findings by assessing the methodology and methods that were applied in obtaining them.

The examination of methodologies and methods leads to a conclusion that is at least disappointing: The empirical findings currently available, though suggestive, come from research that contains crucial conceptual and methodological weaknesses. We do not criticise here discussions of personal experiences with EBF that are sometimes reported, e.g. in Berezin 1998, Horrobin 1996 and Hume 1994, because these do not claim to be more than anecdotal evidence. Furthermore, we don't criticise microsociological studies for not systematically taking into account institutional factors because these studies have never claimed to give accounts of the impact of EBF on knowledge production. But some of the studies that claim to be systematic investigations of EBF practices are written as if there is no such thing as methodology of quantitative or qualitative methods. Others, while methodologically more aware, are severely limited by the methods they apply. We will address below our four major concerns with the studies' methodologies and methods: insufficient information about methods; poor operationalisation of research questions; neglect of the subject matter's complexity; and problems with validity and reliability.

(1) Insufficient information: Given the relevance of the claims presented, surprisingly little information is provided on how empirical evidence was obtained. Chubin and Hackett (1990) describe the sampling procedure and response rates (ibid.:219-220), but don't provide the questionnaire they applied. Talib (2001) provides only incomplete information on sampling. No information is given about respondents' disciplines, and no investigation of non-response is provided. Similarly, Evaluation Associates Ltd. (1999) does not give details on sampling procedures, nor are the three different questionnaires provided, which were used and led to responses that were aggregated afterwards. The two major qualitative studies (Henkel 2000 ; Marginson and Considine 2000) provide neither sufficient information about the questions the participants were asked, nor do they give any hint on how the data were analysed.

(2) Poor operationalisation: If one tries to investigate a pattern of causation (such as 'The RAE causes changes in academics' behaviour), it is of little use to pass this question on to respondents. However, this has been done in many of the quantitative studies. The worst example is that of Harley and Lee who even passed their central (biased) hypothesis at the beginning of their questionnaire:

> "Our central hypothesis is that the naming of mainstream journals as the journals in which publications would score highest in the RAE is having a significant effect on the recruitment and selection of non traditional economists in our institutions. We feel that the maintenance of academic diversity is at risk." (Harley and Lee 1995: 46)

In their survey, Harley and Lee passed on their questions without trying an operationalisation. For example, the questionnaire as provided by Harley and Lee (1997) contains questions about how the RAE has influenced the respondent's work (question 7) or the work in the respondent's department (question 6). Similarly, the agreement/disagreement to statements provided in the studies of McNay (1997) and Talib (2001) pass on the hypotheses on the RAE's impact on academics' behaviour. The questionnaire mailed to journal editors by Talib (2000) applies the same practice. In this case, the problem becomes more obvious because the editors are asked questions *the answers to which they can't know*.

For example, journal editors are asked if they think that more manuscripts are submitted by British academics to their journal, and if they think that the quality of manuscripts submitted by British academics has improved as a result of the RAE (ibid.: 34). Since both trends are caused by an impact of the RAE on academics who submit manuscripts, editors simply cannot know. They can of course have an opinion. But it is indicative of the poor questionnaire that the share of "unsure" answers varies between 6 per cent for a question an editor can answer; to 40 per cent for the question on quality mentioned above; to 61 per cent for another question that is difficult to answer for an editor (ibid.: 38). A similar problem lies with the study by Marginson and Considine (2000) who only surveyed university managers (ibid: 15-16). It is not clear how interviews limited to this group could lead to the highly general conclusions about the changes in knowledge production.

Even if respondents are asked questions they can answer (e.g. Luukkonen 1995), the problem remains that respondents' knowledge about adaptation processes is limited. Questionnaires and closed questions pose an ex-ante restriction to impacts we know of or can think about. This is a severe restriction even when elaborated methods are used in order to obtain hypotheses, see McNay (1997: 3-4). Another objection against questioning scientists about the impact of EBF on knowledge production is that their answers may be biased for a variety of reasons that have nothing to do with this impact, e.g. because evaluation procedures reduce scientists' autonomy. It became clear from Liefner's study that this bias severely affects responses (Liefner 2001). Furthermore, questions cannot probe deeply enough in some areas. For example, the accounts of time consumed for writing grant proposals or for peer reviewing them did not take into account that these activities always involve some component that is beneficial to the scientists' own activities. A further limitation that is shared by any method based on questions (i.e. by qualitative interviews as well) is that they are restricted to what respondents know and how they construct it in the moment they face the question. This implies that subconscious adaptation processes cannot be studied with these methods. In summary, the question-based methods restrict our knowledge to respondents' opinions about the RAE and its impact on knowledge production.

The fact that empirical studies of evaluation-based funding mechanisms have been limited to surveys and qualitative interviews is surprising because bibliometric techniques are widely regarded as the most appropriate method for observing changes in publication practices, while adaptive processes should also be investigated by ethnographic observation.

(3) The theoretical frameworks of the studies referred to above did not take into account the complexity of researchers' decisions that lead to adaptive behaviour. It is common knowledge in sociology of science that the field a scientist works in is an important source of intervening factors that may modify the impact of funding mechanisms. For example, researchers dependency on external funds, the degree of competition, the number of different funding sources accessible can all be said to vary between fields. Similarly, organisational sociology informs us that intra-organisational institutions and practices have an independent influence on members' behaviour. Finally, researchers career stages may modify the impact of EBF on research practice, as has been indicated e.g. by McNay (1997: 26). We have

several important dimensions of the research problem that have not been systematically taken into account in the studies reviewed here.

(4) Problems with validity and reliability: In section 2.2, we have already hinted at the problem of trends established by McNay (1997) and Talib (2001). We will use one example to illustrate the problem. In a non-representative sample of British academics, 24.5 per cent of the researchers agreed to the statement "Because of the RAE time scales I have published some outputs at an earlier stage than I would prefer". To this statement a researcher could agree by assigning values from 1 to 3. How should one interpret the difference between an agreement given with 1 (perhaps "strongly agree") and 2 ? Given that it has not been proven that the sample is representative, what does this finding actually say about the RAE's impact on publication behaviour of all British academics?
Qualitative studies of the subject suffer from problems with causal explanation. In order to establish trends rather than anecdotal evidence, qualitative studies must explain how a given situation leads to a certain outcome. In other words, qualitative studies must reveal the causal mechanisms that lead from a factor that is assumed to be a cause (e.g., EBF) to a change that is assumed to be the consequence (e.g., changes in knowledge production). None of the qualitative studies comes close to this kind of explanation. In the Liefner (2001) study, which is partly based on "open questions", the author himself suggested that the answers he received mainly reflect those effects that spontaneously appeared to be important to the interviewee but are not at all systematic.

If we sum up these concerns about the studies' methodologies, than two serious shortcomings must be noted. Firstly, methodologies and methods provide rather little encouragement for believing in the empirical findings reviewed in sections 2.1 to 2.3. We have anecdotal rather than systematic evidence, and questions rather than answers. Secondly, even if the results were valid we are left with little knowledge about how the production of scientific knowledge is affected by EBF practices. The empirical evidence does not do justice to the subject's complexity; and is mainly evidence that black-boxes causal mechanisms. The findings reported do not sufficiently differentiate between different fields and between different institutional environments. While the studies strongly suggest that the reported trends occur only under specific conditions, no information is given about what conditions that might be. Moreover, most of the findings are constructed out of researchers' perceptions of EBF practices. As follows from the Thomas theorem, these perceptions are important because researchers act according to their perceptions rather than according to some 'true attributes' of EBF practices. Even, if we know the perceptions that influence actions we still don't know *how* researchers react, and what the ultimate consequences for knowledge production are that follow from these reactions.

3. Outline of a research program

3.1 Refining the question

As we have concluded from our review, there is reason to be interested in how the spreading practices of EBF affect the production of scientific knowledge. We will now outline a research

program that can guide systematic investigations of that matter. To do so, we must sharpen the question and explore the theoretical contexts in which it is located.

Given the numerous aspects of knowledge production that have been mentioned in the studies reviewed, a clarification is necessary. The following considerations are aimed at a program for investigating the impact of EBF:

- on the content of knowledge, i.e. asking if different knowledge is produced because of EBF practices; and
- on communication channels, i.e. asking whether knowledge is communicated differently within the scientific community because of EBF practices.

These general questions open up several lines of research of which only a few can be briefly sketched here. A first set of questions relates to the hypothesis that researchers adapt to evaluation criteria independently of what these criteria are supposed to measure. If research quality is measured by numbers of publications, researchers will produce as many publications as possible; if it is measured by citations, they will produce as many citable items as possible and increase self-citations. All this will happen in the rhythm set by EBF procedures.

While this is likely to happen, the opportunities for researchers to do so may vary, e.g. between fields or between organisations. What happens if internal criteria of scientific communities or organisations contradict the messages sent by science policy via EBF? Researchers are likely to face a variety of conflicting signals, and how they ultimately adapt is an open question.

In the context of our study, ‘adaptation’ has probably got a slightly pejorative connotation. However, adaptation is not necessarily an adverse effect. As was stated in 2.2, many commentators on the RAE observed that the RAE did in fact increase the quality of research. From the perspective of adaptation, it could be expected to do so to the extent to which the performance criteria applied actually measure what they actually purport to measure. Since the validity of performance measures in this regard varies, it remains an empirical question under what circumstances and to what extent adaptation to performance criteria leads in the right direction.

A question that is closely related to ‘right’ or ‘wrong’ adaptation processes concerns the feedback between communication (publication) practices and the content of research processes. A ‘pure’ adaptation of publication practices would be one that leaves the methodology and results of knowledge production unchanged. While this kind of adaptation is possible, it seems highly unlikely that it occurs very often. It is much more likely that the perception of expectations concerning publication behaviour will favour the selection of research topics that make it easier to meet the perceived expectations. In fact, the application of performance measures relies on this kind of feedback. It is necessary to determine how changes in publication practices feed back to the design of research processes. This feedback can also be assumed to vary between organisations, fields, and researchers.

The change of communication practices may therefore be an effect of EBF in itself, or it may provide an additional channel of influence for changes in knowledge production caused by EBF.

This leads to the ultimate question that motivates our research program: Is different knowledge produced when knowledge production is subject to EBF?

Interestingly enough, the observer's epistemological position makes a difference here with regard to the macro-levels on which the produced knowledge is accumulated. To illustrate this point, we decide roughly between a realistic position assuming that scientific knowledge approximates the true representation of nature in the long run, and a relativistic position assuming scientific knowledge develops more or less independently of its relation to nature, with the direction dependent on social factors. The problems one has with the impact of EBF on knowledge production obviously depend on which position is taken. From a realist perspective, the scientific knowledge itself is not in danger regardless of what EBF practices are applied. Of course, knowledge production may be slowed down, may take detours, and it may take more time until a particular piece of truth comes out. However, the content of scientific knowledge is ultimately (and in the long run, one has to remember) independent of EBF practices.

While realists need not be concerned about the global development, the impact of EBF practices on local (organisational, national) knowledge production cannot be neglected. Sociologists of science should of course be interested in this specific type of environment and how it affects the construction of scientific knowledge in local settings even if they do not believe it will make a difference to the content of science in the long run. Science policy makers should be even more concerned because they must be interested in short-term rather than long-term conditions and outcomes of knowledge production. Since science is playing an increasing role in a country's economic and political strength, science policy must be concerned about the strength of the country's science base; in other words, about the quantity and quality of scientific knowledge that is produced here and now. Thus, since EBF practice could make a difference to the current content produced locally, the impact of EBF on knowledge production is a point of concern for both sociology of science and science policy even from a realist perspective.

The matter is even more serious for relativists. From their perspective, EBF affects the path knowledge production is taking, and the content of scientific knowledge we can produce and use depends (among other factors) on EBF practices. It is not only the speed of knowledge production and the question where and when certain pieces of knowledge are produced that must be of concern, but the fact that EBF has a lasting influence on what we will know. One of the crucial difficulties in investigating impacts on knowledge production – the assessment of non-knowledge – becomes especially important under these conditions.

Given that EBF is usually designed to improve research quality, there is a particular need to investigate one trend in the findings reported in our review. These findings suggest that EBF practices improve quality to the upper middle level and drive out low quality research, but suppress excellence to a certain extent. Though suggested, this hypothesis is not yet empirically founded. Doubtless it warrants careful investigation. Are there conditions under which EBF constrains rather than promotes excellence? Is there, for example, a critical level of autonomous discretion over funds, which is necessary to maintain excellence?

Another intriguing question is whether the progress of research depends on a 'requisite variety', i.e. on a certain minimal breadth of topics selected and perspectives applied. If this is the case,

maintaining this variety may be more important than promoting quality under certain circumstances. Concentration on high-quality research may then lead to gaps in scientific knowledge that suddenly may become obstacles to further progress. Are there kinds of knowledge that are systematically undersupplied by research resting solely on EBF?

Since there seems to be a worldwide trend to EBF, and since the number of different EBF mechanisms is limited, one can for the first time foresee an international homogenisation of funding practices. In connection with the previous problem, another interesting issue arises: Assuming that different funding practices favour different kinds of knowledge produced, one can imagine research problems having moved around in scientific communities and having looked for national funding conditions which made them likely to be addressed. What consequences would an international homogenisation of funding practices towards EBF have? Are important niches for knowledge production going to disappear?

These are but a few questions related to our program. They indicate the breadth of research that is needed to explore the impact of a rapidly changing funding environment for science on the various aspects of knowledge production.

3.2 A neoinstitutionalist analytical framework

EBF are grounded in systems of rules about who is eligible, how applicants are to be evaluated, and about how a decision about funding is to be achieved. EBF can therefore be said to rest on specific institutions,[11] and the question addressed here is how a specific set of institutions affects knowledge production. With this proposition, we have reached the waiting room of institutionalist sociology of science, with the 'old', 'Mertonian' institutionalist approach having been driven out, and a new one not yet having arrived. This might sound surprising because two prominent movements in science studies are explicitly institutionalist. The 'mode 2 movement' associates far-reaching cognitive changes (the dawn of a "new mode of knowledge production") with institutional changes that have been taking place over the last decades (Gibbons et al. 1994, Nowotny et al. 2001). The 'triple helix movement' postulates the emergence of new institutional arrangements that link universities, industry and government in a new way – namely, as a triple helix in which actors of every type may fulfil the functions of the other two types (Etzkowitz and Leydesdorff 1998, Leydesdorff and Etzkowitz 2001).

While increasingly prominent in science studies and science policy communities, both movements are not of much help if one looks for causal explanations. Eight years after mode 2 was announced, there is still not much empirical evidence supporting the far-reaching claims. After the authors of the 1994 book admitted that their hypothesis still awaited empirical confirmation (e.g. Gibbons et al. 1994: viii) it is rather embarrassing that the 2001 book, written

[11] Institutions are defined here rather narrowly as systems of formal and informal rules. This approach, which is taken e.g. by North (1990) and Mayntz and Scharpf (1995, see also Scharpf 1997), has several advantages. Firstly, it enables a conceptualization of institutions not only as being subject to evolutionary change, but also as formed by purposeful actions and thus as dependent factors. Secondly, the narrowed definition helps clarifying that institutions are only one of the factors influencing actions. Thirdly, the narrow concept supports the empirical search for institutions and their effects (Gläser 2001a: 701).

by three of these authors, does not provide additional empirical support. Moreover, empirical and theoretical objections have been accumulated (Weingart 1997, Shinn 1999, Gläser 2000; 2001b). The 'triple helix model', while more strongly connected to empirical studies, favours the classical focus on 'institutions, institutions, institutions', thus submitting itself to the criticism raised in the introduction. With regard to research conduct, changes in academics' values are stated but insufficiently researched (Etzkowitz 1998). This is as close to knowledge production as the model goes. Consequences of the 'triple helix dynamics' for the content or forms of scientific knowledge produced under the alleged new conditions are not investigated.

It still seems extremely difficult to marry institutionalists' and policy researchers' interest in institutions with ethnographers' interest in knowledge production. Having this combination of interest, two of us have adopted a neoinstitutionalist analytical framework – the 'actor-centred institutionalism' for purposes of sociological investigations of knowledge production (Gläser 1998, 2001a; Laudel 1999). Actor-centred institutionalism includes three important methodological considerations that are crucial for the investigation of institutional influences on knowledge production:

(1) Multi-level analysis: As has become clear from our review, institutions of EBF are designed to govern actors on several levels of aggregation, e.g. organisations, organisational subunits, networks, and individuals. Being subject to EBF as corporate actors, research organisations such as universities will in turn develop intra-organisational institutions that govern their researchers. Researchers will therefore perceive institutions on different levels of aggregation and respond to them. While institutional influences on knowledge production will be mediated ultimately in researchers' work, the influences themselves and actions mediating them will be generated on a variety of levels.

(2) Overlapping institutions: One of the important differences between 'old' and 'new' institutionalism is the latter's acknowledgement of heterogeneous institutional environments. Even a single institution may send conflicting signals to actors. But more importantly, actors always face different institutions that are maintained by different social contexts and may pose conflicting demands. This is certainly the case with EBF institutions – e.g. when researchers have to respond simultaneously to a funding formula for their department and to peer reviewed project funding. In addition to that, there are many other institutions – of the organisation, the scientific community, society as a whole – that are relevant and address expectations to the researcher. In order to investigate the impact of one of these institutions, the others have to be taken into account as intervening influences.

(3) Non-institutional influences: Another requirement that is paramount for the investigation of institutional impacts on knowledge production is the inclusion of non-institutional influences, The neoinstituionalist, non-deterministic approach to institutions also supports this enlargement of the framework. Of special importance for the investigation of knowledge production is the opportunity to include non-social factors, in our case influences stemming

from the knowledge used, the research equipment applied and from the part of nature addressed by the research.[12]

By characterising our analytical framework as neoinstitutionalist, we have already sketched the main variables that must be included in a research program on the impact of EBF on the production of scientific knowledge. Figure 5 gives an overview of these variables and of the causal relations we assume exist. In the following sections, we will describe the variables in more detail.

[12] We regard knowledge and equipment as non-social in the sense that they provide an 'objective reality' that cannot be wished away (Sismondo 1993). Only in this sense can both knowledge and instruments, albeit products of human and therefore social construction processes, be regarded as non-social. In the words of Berger and Luckmann (1967: 57): "The paradox is that man is capable of producing a world that he then experiences as something other than a human product." There are two more arguments for treating knowledge and instruments *analytically* as non-social: Firstly, they can affect the success of the production of scientific knowledge directly, i.e. without mediation by other actors. Secondly, the morphology of these factors cannot be explained by social factors alone (Gläser and Laudel 2002).

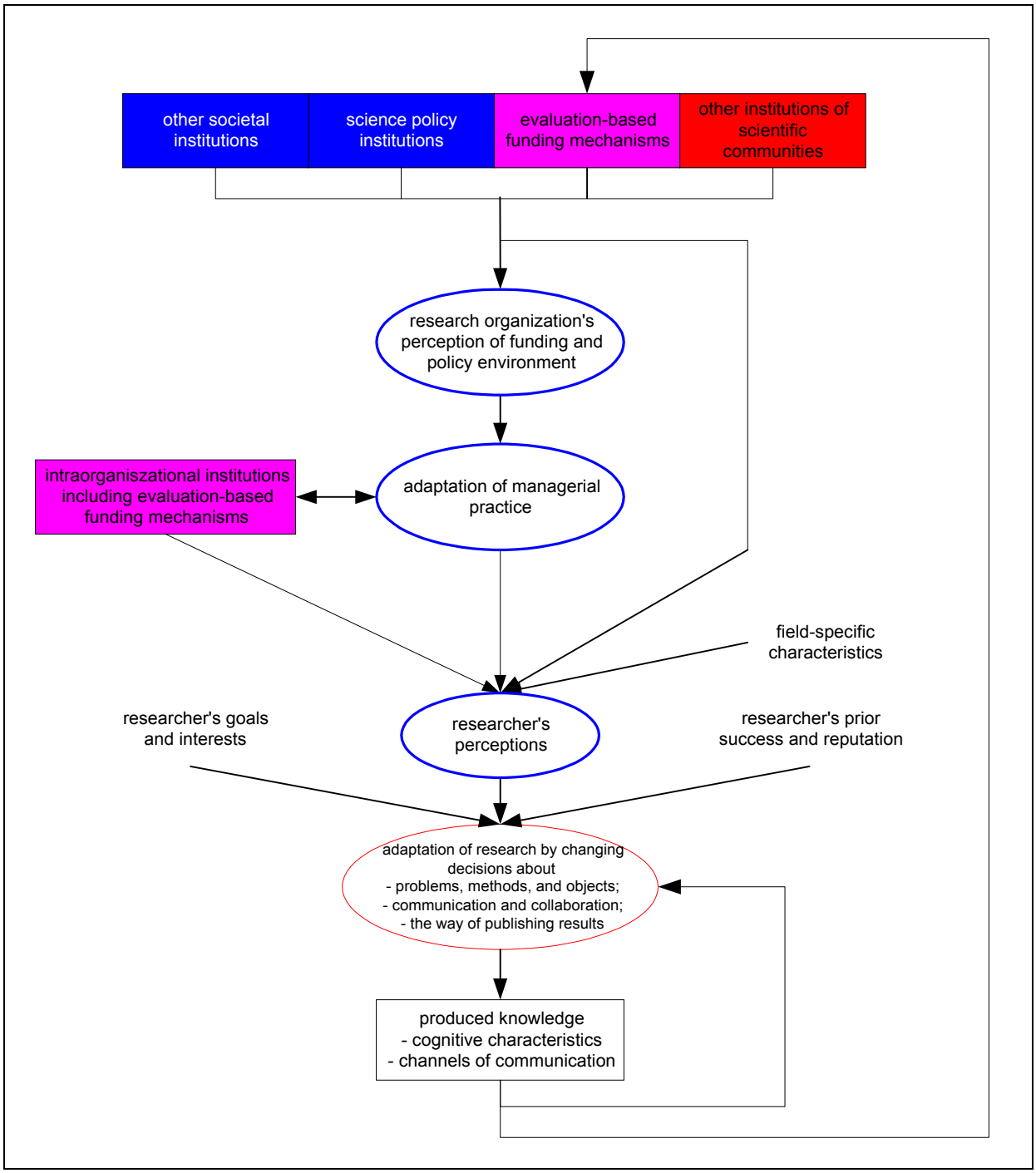

Figure 5: Variables and assumed causal relationships

3.3 Institutions

As it is shown in figure 5, we have to take into account at least five types of institutions stemming from three different sources. The institutions that are investigated as causing changes of knowledge production are EBF mechanisms. EBF has been following historical paths that are specific for each country. Due to the limited opportunities on how an evaluation of scientific research can be conducted, and how research can be funded, there are a few crucial elements that EBF practices have in common. EBF is always a specific overlap of societal and scientific institutions. This is clearly visible when scientists evaluate their colleagues in order to prepare

policy decisions on funding, as is the case with the RAE or with evaluation of research institutes in Germany. In other cases, the overlap of these two institutions may be less clearly recognisable. For example, formulas applied to the funding of universities may be seen to lie completely at the discretion of science policy. However, funding formulas are usually not established without the agreement of scientists. On the other hand, while some practices of peer reviewed project funding seem to be cases of pure self-governance because they are conducted autonomously by scientists, the money that is distributed was originally assigned to them by science policy.[13]

Since EBF is a decision process on funding of research using information on the performance of the actors to be funded, several elements must be common to all EBF practices. All EBF mechanisms are connected to a source of funding; provide funding on a defined level of aggregation; are designed to provide resources for specific purposes; are applied with a specific frequency and use a mode of evaluation. The variation of EBF practices in these dimensions can be used as a framework for analysing them.

1) Sources of EBF

Four different sources of EBF can be distinguished. The funding a researcher receives may come from the state, from the private sector (private foundations or industrial enterprises), from scientific communities, or from his or her organisation (e.g. the university). These are not the ultimate sources of money, but those the actors - a researcher (or an organisation) - may interact with while seeking funding. The EBF practices of these sources vary along different dimensions such as the amount of money that is accessible, expectations connected with the funding, etc. For example, contract research for industry does not usually emerge on the basis of competition and evaluation. While evaluations (by a company's in-house researchers) certainly play a role, a call for proposals issued by industry seems to be an exception rather than the norm.

The most important sources of EBF are the state and scientific communities. This is not because they spend most money on research (industry is on top in this respect), but because EBF is a widespread practice in science policy. While EBF is being increasingly applied within research organisations as a means of distributing funds to organisational subunits (see 2.3), in some of Germany's federal states (e.g. Schleswig-Hostein) the application of internal EBF mechanisms in universities has become a legal requirement (Federkeil and Ziegele 2001: 9).

International funding agencies set up jointly by several countries play a rapidly increasing role. The most prominent examples is certainly the EBF of the European Union. The EBF practices emerging here may include specific evaluation criteria (such as international collaboration) and a sophisticated interaction between international and national levels in terms of contributions to funding or separation of (international) evaluation and (national) funding.

The scientific community may also be a source of EBF. While the money still comes from the state, the latter delegates the decisions about who is funded and to what extent to agencies of the scientific communities (e.g. the NSF in the USA, the ARC in Australia, and the DFG in

[13] For analyses of the history and functioning of funding agencies, see Rip (1994) and Braun (1998).

Germany). There are, of course, degrees of science policy's participation and priority setting. However, areas can be still identified in which EBF lies exclusively in the hands of researchers themselves.

Another important factor is the variety of EBF that can be accessed by an actor. Research organisations and researchers may be eligible for only one EBF scheme, or they may be in a situation where they can choose from a variety of funding agencies to address. The number and diversity of funding sources depends mainly on the scientific field an actor is working in. For example, some basic research areas in natural sciences and some social sciences have almost no chances of getting money from industry or state funding that is devoted to technology transfer. Within more applied fields, a variety of private foundations exist e.g. in the area of cancer research. Depending on a field an actor is located in (and on national historical developments as well), the actor may depend on one source for EBF or may have access to a variety of sources with different EBF practices. Organisational sociology's resource dependency has taught us that the latter situation increases a actor's autonomy (Pfeffer and Salancik 1978, Bogler 1994: 173-174).

2) Scale of the EBF

EBF varies according to the scale on which research is funded. The scale of research funding spans from single activities (travel, publications) to the funding of whole organisations. the funding of single activities is of lesser interest here because its impact on knowledge production is small. It is, however, part of the overall funding of a researcher's activities.

More common is EBF of projects. This has become the most common feature of research funding. While the contribution to project and infrastructure costs may vary, many researchers' opportunities to conduct research at all depend on EBF of projects.

On a higher level of aggregation, networks of projects and temporary research organisations may be funded by means of EBF. Again, the degree of funding (and therefore the dependency on the evaluation's outcome) may vary from contributions to research costs to fully financed networks or organisations. Finally, organisations can be subject to EBF (and do so more and more often). Research projects, networks, temporary organisations and permanent organisations can be funded either partly or completely by EBF. The dependence of research on the EBF varies accordingly.

3) Share of costs that are subject to EBF

The actual contribution of an EBF practice to an actor's funding may greatly vary. There are actors who receive all or most of their funding without undergoing an evaluation, i.e. as recurrent funding (until recently, Germany's Max-Planck Society and the Australian National University's Institute for Advanced Studies were examples), while others may depend completely on EBF (e.g. the oceanographers in the U.S. described by Mukerji 1989). The introduction of EBF at German Universities started by focussing on selected positions of the budgets and limited shares thereof but it is planned that the proportion of the budget allocated through EBF will grow.

Practices of project funding vary with regard to how 'complete' the funding is. For example, project funding may or may not include the principal investigator's salary, infrastructure costs

etc. Furthermore, certain funding programmes already by definition fund only a defined share of the real costs of a research project requiring that the other part is met by the institution that applied for the project. For example, it is common practise in many EU calls for proposals that institutions have to meet a part of the total costs or have to apply for complementing funds with other funding agencies.

The important difference that has to be taken into account here is that between an EBF agency's decision on the existence (of a project, a network, or an organisation) and its decision on the size of a contribution to that existence. From the perspective of the funded actor, its dependence on an EBF agency (and therefore on the evaluation procedure included) varies with the share of its activities that is funded by that agency by means of EBF.

4) Frequency of EBF

The time characteristics of EBF practices contribute an important variation, too. How often and how regularly an actor faces an EBF practice can be assumed to have a strong influence on adaptation processes. A first distinction that must be made is that between ad hoc and regular EBF practices. Peer reviewed project funding occurs on an ad hoc basis: when a project proposal is submitted, an evaluation-based decision on its funding is made, which may include that there are defined deadlines for the submission of proposals e.g. ones or twice a year. The application of funding formulas, the British RAE or the Evaluation of research institutes of the Leibniz-Society or German Universities, are conducted regularly – from yearly in the case of formula funding to five-yearly in the case of the RAE and the evaluations of the Leibniz Society's institutes. For the German universities those periods still have to be defined in most cases.

The characterisation of frequency according to the practice of the funding agency does not necessarily coincide with the perception of researchers. While peer reviewed project funding is certainly an ad hoc procedure, researchers frequently apply for projects and thus may constantly face the same procedure. This distinction between an EBF practice and the funded actor's practice must be kept in mind because for adaptation processes it is the funded actor's perception that matters after all.

5) Modes of evaluation

The basic modes of evaluation that exist and can be applied in EBF are the classical internal mode of scientific communities (i.e. peer review), and the application of quantitative performance indicators. The latter seem to have been introduced for mainly two reasons. Firstly, they give science policy greater opportunities to develop an opinion of its own. By applying quantitative performance indicators, actors in science can be compared according to their performance more or less independent of peers' judgements, and science policy can reach conclusions seemingly independent of the scientific community to which an actor belongs. Secondly, the application of quantitative performance indicators is often cheaper and supposedly easier than peer review. In addition, the higher degree of "objectivity" of the measures are mentioned as an advantage for the use of quantitative indicators. Peer review is an instrument of limited capacity, and the demand for evaluation uttered by science policy seems to have outgrown this capacity.

The poles of the spectrum of EBF practices are represented by pure peer review (that is commonly applied in peer reviewed funding of projects) and a pure application of quantitative performance measures (e.g. in funding formulas). Between these poles, we find a variety of combinations of peer review and quantitative measures. For example, quantitative information on departments is subjected to peer review in the British RAE. Due to the limitations that are inherent to both methods (concerning the drawbacks of peer review see e.g. Horrobin 1990, for problems related to bibliometric methods see van Raan 1996) increasingly calls for a combination of both approaches (ibid.).

The five dimensions we have discussed characterise our focal institution, i.e. EBF. Other institutions that overlap EBF and whose influence on actors in science must therefore be analysed are

- *Institutions of scientific communities*. These mostly informal institutions (such as standards for empirical conduct, institutionalised goals, and preferences) exert a strong influence on researchers because they are the prime social environments in which their results are judged. Any evaluation of researchers rests on their performance as perceived by the scientific community. The extent to which criteria of evaluation-based funding mechanisms correspond to these informal institutions determines whether a researcher faces conflicting pressures when seeking funds.
- *Science policy institutions*. The practices of EBF are embedded not only in institutions of the scientific community but simultaneously in science policy institutions. The autonomy and power of actors in science and modes of interactions between these actors are defined primarily by science policy institutions. Both the relation between recurrent funding and EBF and the variety of funding sources accessible by actors in a science system are an outcome of science policy institutions.
- *General societal institutions*. Actions in science are also influenced by a society's general institutions, i.e. by institutions that affect actors of a given country both within and outside science. For example, the specific way of budgeting public money in Germany (which is now being replaced by "global budgeting") had for decades restrained the autonomy of research organisations. Other examples for general institutions that affect science are contract law and labour law, with the first affecting how funding contracts are regulated and the second affecting careers in science. Finally, society's expectations concerning research provide a universal background for all actions in science. Even when money is distributed autonomously by scientific communities, decision-making scientists know about society's expectations of science and can be assumed to be influenced by these expectations.

The overlap of these different institutions produces important aspects of a scientist's situation regarding funding of research. Variables that depend on these institutions and are likely to affect adaptive behaviour are the degree of an actor's dependency on EBF and the degree of an actor's dependency on specific EBF practices.

3.4 Scientific knowledge

Scientific knowledge is certainly the most uncomfortable variable that the sociology of science has to cope with. It is quite understandable that institutionalist approaches try to avoid it. However, without including it into our investigations we will never know what institutions ultimately do to the production of scientific knowledge.

Our research program's dependent variables cover the content of the scientific knowledge produced, the ways this knowledge is produced and formulated and the channels through which it is communicated. With regard to content, we are of course not interested in the idiosyncratic knowledge claims that are produced by single research processes. In the context of a sociological research program, it is useless to find out that adaptation to EBF practices has led a researcher (not) to produce knowledge about, say, the mating behaviour of ringtail possums. Since research leads (or is supposed to lead) to new knowledge, the outcome of every single research process consists of such idiosyncratic knowledge claims. In order to detect patterns in knowledge production, we must abstract from the content's idiosyncracies and look for attributes of that content that makes observation comparable. That does not mean that idiosyncratic changes (such as the prescription of specific research methods by a grant proposal's reviewer) are irrelevant – the task is to describe them in a way that makes these idiosyncracies comparable.

Since all research processes can be said to solve problems (theoretically derived research questions) by applying methods to certain objects, these elements can be used to characterise attributes of knowledge which are a) possibly subject to changes in knowledge production, and b) comparable between research processes. A provisional list of such attributes can be compiled drawing on the literature review provided in 2. and on methodological considerations of comparative approaches to non-social factors in the sociology of science (Gläser and Laudel 2002).

The most important element of research that is to be considered here is the research problem. The literature on EBF practices suggests that the following attributes of research problems should be included in our analyses:

- *Basic versus applied character*. It is certainly increasingly difficult to locate research problems and their solutions on a one-dimensional basic – applied scale. However, the distinction between these types of research remains important and useful in many cases. In the literature on EBF there are some slight indications of research moving to either more basic or to more applied research because of EBF practices. While the comparison of research in different fields with regard to its basic/applied character might be difficult, it should be possible to compare the work of one researcher at different points in time.
- *Disciplinary versus interdisciplinary character*. Similar to the basic versus applied character, we cannot speak of a distinction or even of one dimension here. However, the degree of heterogeneity of the knowledge applied to problem-solving is a prominent attribute in the literature on EBF.
- *Theoretical, empirical, or methodical character*. A research problem can be characterised as being either theoretical, i.e. not requiring production of empirical data in order to be solved; experimental (able to be solved by production of empirical data); or methodical (concerning

knowledge about methods rather than knowledge about nature). While the EBF literature has not addressed the question of whether one of these types is affected by EBF, our own data leads at least to concerns on this score.

- *Mainstream versus non-mainstream research*. This attribute describes how a problem is located in a field and how it relates to a field's institutionalised research goals. The knowledge that will be produced by solving the problem may be a contribution that is central to the concerns of other researchers and that may be expected to be used by others immediately, or it may refer to something that is of concern for only few researchers and is of lower priority in the field. Researchers' interests and goals may differ from what is regarded central and important by their community. However, the community may not necessarily be right.
- *Probability of success*. Research problems vary with regard to their probability of being solved. It is assumed that the greater the advance that can be achieved by solving a problem, the greater is also the risk that the problem cannot be solved given our current knowledge. The riskier problems also seem to be the ones that require the most innovative solutions. Therefore, the risk a problem poses is an important aspect of both the progress of science and a researcher's performance.
- *Time characteristics*. Research problems differ with respect to the time their solution requires. This attribute does not refer to researchers abilities in the sense that some researchers my find a solution quicker than others, but to the 'Eigentime' of research processes (Gläser and Laudel 2002), that is, to the time that is required to produce all the prerequisites for problem solving. For example, the solution of some problems may require long-term observational records, while other problems can be solved by conducting experiments within a day.

The characteristics of research objects and methods that are relevant for characterising knowledge are basically the same as those of problems in relation to basic versus applied and mainstream versus non-mainstream dimensions. However, characteristics of research objects and methods may vary independently of the problems being investigated. For example, in order to find a compromise between their interests and external expectations, researchers may try to solve mainstream problems with non-mainstream research objects, or to conduct basic research by using applied objects.

The second aspect of the scientific knowledge produced is the way it is presented to the researcher's environment. The publication practices of researchers should be investigated with regard to the channels of communication. Results can be 'salami-sliced' and published in a series of articles, or they can be synthesised in a smaller number of articles. Results can be published

- in important core journals, in peripheral journals, or in a mix of both;
- in journals of a scientist's own field or in journals belonging to other fields;
- in journal articles or in books, or in both;
- in media addressed to colleagues or in media that provide a channel to potential external applicants of the knowledge; and
- in media addressed to a national or an international audience.

In the context of this research program, the whole range of possible outputs of research must be included. It would be a highly consequential mistake to follow the lead of EBF practices by reducing the output under investigation to certain types of publications. For example, patents, works of art, reports, articles in specifically targeted media (such as general medical journals aimed at medical practitioners) must be included when the adaptation of communication behaviour is investigated.

3.5 Mediating processes and intervening factors

Institutions do not change knowledge - actors do. Thus, the process that mediates between institutions (of EBF) and knowledge is human practice. It may well be possible to find *causal relations* between changes of EBF (e.g. the replacement of publication counts by citation counts in a funding formula) and, say, researchers' publication practices. However, the *causal mechanisms* may be explored only by empirically investigating how actors respond to EBF institutions by producing different knowledge (or producing knowledge differently).

Mediating processes between EBF institutions and knowledge production occur at least at two different levels. Firstly, EBF of organisations triggers an adaptation of managerial practices in research organisations. While changes in managerial practices have been studied mainly in the more general context of the rise of new public management, several studies have recognised that research organisations respond directly to EBF practices (see 2.4). According to the literature, the following changes in managerial practices that may affect researchers and research must be examined:

- More pressure on researchers to meet EBF criteria;
- Linking internal promotion to EBF criteria;
- Adapting hiring practices to EBF criteria;
- Concentration or diversification of research potential according to EBF criteria;
- Better support for research;
- introduction of intra-organisational EBF practices.

These managerial practices themselves do not change knowledge production but affect the response of researchers to the situation. The ultimate mediation of all influences on knowledge production lies with the researchers and their research groups. It is their decision making about the elements of knowledge production that is influenced by how a situation is perceived and that subsequently shapes the knowledge that is produced. Within the numerous selections that impregnate research practice (Knorr-Cetina 1981: 33-48) there are some decisions that can be labelled strategic because they determine mid-term courses of knowledge production. These include the choice of research problems, research objects, methods, and decisions on communication and collaboration. These decisions can only be discerned analytically. Laboratory studies have shown that they are embedded in a stream of choices made by scientists in their everyday work, and that at least some of these are made without consciously considering their impact on knowledge production. Many scientists conduct more than one project at a time and shift resources from several sources between projects. All the research processes must be

included in the investigation. Only then it will be possible to consider the role that the different 'mixes' of recurrent funding and EBF play.

An important aspect of adaptation practices is their 'hardening' by fostering the emergence of informal rules in organisations and by changing researchers' attitudes to research. Continuous submission to EBF is likely to make an actor internalise the norms and values attached to it. This tendency has been described by a publisher with regard to researchers' publication practices:

> Finally, there is inherent in this situation a definite tendency for authors to view publication itself as the end-result and objective of the whole exercise, rather than the dissemination of knowledge and ideas for which publication is really only the vehicle and the initial stage. This is another manifestation of the audit culture, which characteristically confuses means with ends and process with performance. (Mynott 1999: 131)

The choice of publication channels is often one of the crucial adaptation processes. Of special interest for this program is the role publication strategies play for researchers. It is necessary to investigate whether researchers follow consistent publication practices that affect strategic decisions about their research. For example, if researchers select projects whose results are quickly and easily publishable, publication strategies feed back into knowledge production.

The adaptation processes we have discussed not only mediate the influences of EBF institutions. All factors we have discussed so far may affect researchers' strategic decisions and thus the way they adapt to EBF. As has been indicated in figure 5, institutional influences are not restricted to the actors whose behaviour the institutions shall govern. Institutions "bypass" organisational boundaries and affect researchers directly by creating "institutionally shared beliefs" (Scott 1991: 180-181). All institutions concerning research may affect researchers' mediating actions directly if the researcher perceives the institutions as relevant.

An important set of intervening factors that has not been discussed yet is non-institutional (and sometimes even non-social) *field-specific characteristics* of causation ((Gläser and Laudel 2002). For example, the cost of research varies greatly between fields and the capacity of researchers to conduct projects solely with untied recurrent funding will fluctuate accordingly. Other field-specific variables that are important in this context include the field's publication practices and citation rates, its application-orientation (and thus the extent to which funds can be obtained from industry), or simpler criteria such as its size.

Finally, an analysis of impacts of EBF on the production of scientific knowledge must take into account important *characteristics of actors* such as researchers' goals and interests, as well as their prior scientific success and reputation. These variables affect the degree that researchers are willing and able to adapt to EBF mechanisms, or the extent to which they try to, and can, avoid changing their plans while still obtaining the necessary funds.

3.6 Methodological considerations

A research program that aims to investigate institutional impacts on knowledge production comparatively on different levels of aggregation clearly poses a challenge to the sociology of science's methodology. Though scope and interests of single projects within such a framework

will be limited, the irreducible complexity underlying the research problem prevents easy solutions. We have already described the research problem as demanding both a multilevel analysis and the inclusion of numerous heterogeneous factors. Moreover, we can now characterise the issues that have to be addressed by possible respondents of empirical studies (mainly researchers and research managers) as highly sensitive and difficult to investigate. Judging from the literature and from our own interviews with researchers, the impact of EBF on the production of scientific knowledge can be said to be

- only partially known to respondents because of subconscious adaptation processes;
- only partially communicable by respondents, either because they are only partially known or because adaptation processes that are known to respondents may be difficult to describe;
- highly emotive because they are linked to respondents' self-esteem, values, stress, etc.; and
- highly esoteric because they are inseparably linked to field-specific processes.

These characteristics of our subject matter point to the importance of methodological issues. As we have discussed in 2.5, many findings of earlier studies cannot be relied upon because of the studies' poor methodology, a methodology that did not take into account the characteristics mentioned above. We think it necessary to consider the following methodological guidelines for empirical investigations.

Firstly, questionnaires mailed to respondents seem to be of little use for the matter at hand for two reasons. Since we do not know enough about our subject matter to prepare fixed-choice questions, one of the strengths of questionnaires – providing data for statistical analyses – cannot be exploited. What is more important, however, is the fact that we must probe more deeply into institutional influences, adaptation processes and knowledge production than we can possibly do with questionnaires. To investigate the impact of EBF on knowledge production, we must explore national, organisational, and scientific communities' cultures, individuals' careers, goals and interests, and the idiosyncracies of knowledge production. Questionnaires suit breadth rather than depth and therefore seem of little use for our program.

Secondly, while qualitative interviews are a much more sensible instrument, they also have limitations. Qualitative interviews still rest on asking participants and are therefore limited to what respondents remember (more precisely: reconstruct) and are able to communicate in the moment they are asked. While we should always keep in mind that this is a serious limitation, interviews can contribute to our investigation if properly used. The most important presupposition for a successful application of interviews seems to be the operationalisation of the research question. The term 'operationalisation' looks alien here because it usually refers to the translation of hypotheses into items on a questionnaire. However, qualitative research needs such translation processes too. The central task of this translation will be to disassemble the basic question "How does EBF affect your research?" into smaller questions that describe single aspects independently of each other, thereby avoiding passing the overall question to the respondent. For example, interviews on the impact of EBF on knowledge production could cover the following main themes:

- Research projects conducted during the last five to ten years (the researchers' research trail) (Chubin and Connolly 1982);
- Reasons for abandoning certain research trails and following others, especially the extent to which this behaviour is triggered by funding;
- Strategic decisions (selection of problems, objects, methods, collaborators and communication channels) that have been made in the project design phase and the reasons for making these selections;
- Sources of funding for these projects, application procedures and selection criteria applied to the researcher's proposals;
- Degree of autonomy and external pressures perceived by the researcher, and the reasons for them.

Thirdly, ethnographic observation must play an important role in research on the impact of EBF on the production of scientific knowledge. This conclusion is rather obvious after we have established that some important aspects of our research problem cannot be addressed by asking questions. The logical conclusion is to observe scientists' practices in order to investigate the aspects that cannot be communicated. However, this use of ethnographic observations has some particular issues. Ethnographic observations are usually assumed to obtain their focus from the social setting under observation rather from a theoretically guided research program (Amann and Hirschauer 1997: 19). Thus, the task will be to develop a theoretically informed ethnographic observation that combines the advantages of ethnography with the practice of looking for something the observer knows to be there but does not know what it looks like. Given the little we know about mediation of institutional influences on knowledge production, it seems impossible to apply any scheme for a systematic observation. It is exactly the combination between a 'soft' concept of method and a 'hard' concept of the empirical (ibid.:9) that is needed for this type of investigation.

Fourthly, in order to link the micro-processes of knowledge production to the macro-structures of scientific knowledge, and in order to observe changes in researchers' publication practices, it seems essential to include scientometric methods. For several reasons, scientometrics and sociology of science have moved apart. However, with a proper theoretical micro-foundation, scientometric methods provide a unique approach to important aspects of scientific practice. Similar to decisions on research problems, decisions on publication practices may be affected by EBF in a way that the researcher is not conscious of. Scientometric methods provide the opportunity to unobtrusively observe publication practices and to reveal patterns on all levels of aggregation. Moreover, since bibliometric links between publications reflect structures of knowledge, scientometric methods can be applied to discover macroscopic changes in knowledge production. These are advantages a multi-level study of knowledge production must not ignore.

4. Concluding remarks

Investigations of peer reviewed project funding were triggered by concerns of procedural justice and accountability. Most investigations of the British RAE have been triggered by scientists' perceptions of negative impacts of their work. While both concerns are legitimate, it is rather obvious that what should be of primary concern for both sociology of science and science policy has not yet been asked, let alone been answered: What does EBF mean for the outcome of the processes it intends to influence? It has turned out that this question describes a gap in our knowledge that should at least discomfort science policy and simultaneously poses interesting challenges to the sociology of science.

The research program on this subject we propose is ambitious. To find answers to the questions we have tried to formulate requires comparative analyses across organisations, countries, and fields. Therefore, we would like it to be regarded as an invitation.

References

Adam, David, 2002: The counting house. *Nature* 415: 726-729.

Amann, Klaus, and Stefan Hirschauer, 1997: Die Befremdung der eigenen Kultur. Ein Programm. Stefan Hirschauer and Klaus Amann: *Die Befremdung der eigenen Kultur: Zur ethnographischen Herausforderung soziologischer Empirie*, Frankfurt a. M.: Suhrkamp, 7-52.

Andersen, Uwe, Heiner Minssen, Beate Molsich, and Uwe Wilkesmann, 2001: *Kontextsteuerung von Hochschulen durch verschiedene Modi der Mittelzuweisung. Aktualisierte Ergebnisse der explorativen Vorstudie*. Diskussionspapier Nr. 01-1, Ruhr-Universität Bochum, Institut für Arbeitswissenschaft.

Berezin, Alexander, 1998: The Perils of Centralized Research Funding System. *Knowledge, Technology & Policy* 11: 5-26.

Berger, Peter L., and Thomas Luckmann, 1967: The social construction of reality. Harmondsworth: Penguin.

Bogler, R., 1994: The impact of past experience on people's preference: The case of university researchers' dependency on funding sources. *Higher Education* 28: 169-187.

Bourke, Paul, and Linda Butler, 1999: The Efficacy of Different Modes of Funding Research: Perspectives from Australian Data on the Biological Sciences. *Research Policy* 28: 489-499.

Braun, Dietmar, 1998: The role of funding agencies in the cognitive development of science. *Research Policy* 27: 807-821.

Butler, Linda, 2001: *Monitoring Australia's Scientific Research: Partial indicators of Australia's research performance*. Canberra: Australian Academy of Science.

Butler, Linda, 2002: Explaining Australia's increased share of ISI publications – the effects of a funding formula based on publication counts, *Research Policy* (in press).

Chubin, Daryl E., and Edward J. Hackett, 1990: *Peerless Science: Peer Review and U.S. Science Policy*. Albany, N.Y.: State University of New York Press.

Chubin, Daryl E., and Terence Connolly, 1982: Research Trails and Science Policies. Norbert Elias, Herminio Martins and Richard Whitley: *Scientific Establishments and Hierarchies*, Dordrecht: Reidel, 293-311.

Cohen, Hal, 2002: British research funding: One good cut deserves another. *Scientist* 16: 44-45.

Cole, Stephen, Jonathan R. Cole, and Gary A. Simon, 1981: Chance and Consensus in Peer Review. *Science* 214: 881-886.

Cozzens, Susan E., 1986: Theme Section 'Funding and Knowledge Growth': Editor's Introduction. *Social Studies of Science* 19: 9-21.

Etzkowitz, Henry, 1998: The norms of entrepreneurial science: cognitive effects of the new university-industry linkages. *Research Policy* 27: 823-833.

Etzkowitz, Henry, and Loet Leydesdorff, 1998: The endless transition: A "triple helix" of university-industry-government relations. *Minerva* 36: 203-208.

Evaluation Associates Ltd. 1999. Interdisciplinary Research and the Research Assessment Exercise. A report for the UK Higher Education Funding Bodies.

Federkeil, Gero, and Frank Ziegele, 2001: *Globalhaushalte an Hochschulen in Deutschland: Entwicklungsstand und Empfehlungen*. Gütersloh: CHE.

Geuna, Aldo, and Ben R. Martin, 2001. *University Research Evaluation and Funding: An International Comparison*. Brighton: Science Policy Research Unit.

Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow, 1994: *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: SAGE.

Gläser, Jochen, 2000: Limits of change: cognitive constraints on "postmodernization" and the political redirection of science. *Social Science Information* 39: 439-465.

Gläser, Jochen, 2001a: Macrostructures, careers and knowledge production: a neoinstitutionalist approach. *International Journal of Technology Management* 22: 698-715.

Gläser, Jochen, 2001b: Modus 2a und Modus 2b. Gerd Bender: *Neue Formen der Wissenserzeugung*, Frankfurt a.M.: Campus, 83-99.

Gläser, Jochen, 2002: Expertise zum Thema „Privatisierung von Wissenschaft?“

Gläser, Jochen, and Grit Laudel. 2002. Linkage Variables between nature and the social. Paper to be presented at *ISA XVth World Congress of Sociology*.

Gläser, Jochen. 1998. Kognitive Neuorientierung der ostdeutschen außeruniversitären Grundlagenforschung als Folge des Institutionentransfers. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Harley, Sandra, and Fred Lee. 1995. The Academic Labour Process and the Research Assessment Exercise: Academic Diversity and the Future of Non-Mainstream Economics in UK Universities. Pp. 48: Leicester Business School.

Harley, Sandra, and Frederic S. Lee, 1997: Research Selectivity, Managerialism, and the Academic Labor Process: The Future of Nonmainstream Economics in U.K. Universities. *Human Relations* 50: 1427-1460.

Harley, Sandra, and Peter Lowe. 1999. Academics Divided: The Research Assessment Exercise and the Academic Labour Process. Pp. 28pp in *Critical Perspectives on Accounting*. CUNY.

Henkel, Mary, 2000: *Academic Identities and Policy Change in Higher Education*. London: Jessica Kingsley.

Hirschauer, Stefan, 2002: Expertise zum Thema „Die Innenwelt des Peer Review. Qualitätszuschreibung und informelle Wissenschaftskommunikation in Fachzeitschriften“.

Horrobin, David F., 1990: The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association (JAMA)* 263: 1438-1441.

Horrobin, David F., 1996: Peer review of grant applications: A harbinger for mediocrity in clinical research? *Lancet* 348: 1293-1295.

Hume, David, 1994: Research Funding in Australia: Plugging the Brain Drain. *Search* 25: 27-31.

Jensen, Mogens Klostergaard, and Klaus Neuvians, 1994: Globalhaushalte für Hochschulen. Ein Vergleich Dänemark/Deutschland. Gütersloh: CHE.

Knorr-Cetina, Karin D., 1981: *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press.

Knorr-Cetina, Karin D., 1995: Laboratory Studies. The Cultural Approach to the Study of Science. Sheila Jasanoff, Gerald E. Markle, James C. Petersen and Trevor Pinch: *Handbook of Science and Technology Studies*, London: SAGE, 140-166.

Körber-Weik, Margot, and Diana Schmidt, 2000: Leistungsorientierte Hochschulfinanzierung in Baden-Württemberg. Reflektionen zu statistischen Adäquation im politischen Prozeß. *Jahrbücher für Nationalökonomie und Statistik* 220: 715-730.

Laudel, Grit, 1999: *Interdisziplinäre Forschungskooperation: Erfolgsbedingungen der Institution 'Sonderforschungsbereich'*. Berlin: Edition Sigma.

Law, John, 1994: *Organizing Modernity.* Oxford: Blackwell.

Lederman, Leon M., 1993: What Can We Learn from the Supercolliders Demise. *Scientist* 7: 12-12.

Leydesdorff, Loet, and Henry Etzkowitz, 2001: The Transformation Of University-industry-government Relations. *Electronic Journal of Sociology* 5: [http://www.icaap.org/iuicode?1.5.4.1].

Liefner, Ingo, 2001: Leistungsorientierte Ressourcensteuerung in Hochschulsystemen: Ein internationaler Vergleich. Berlin: Duncker&Humblot.

Luukkonen, Terttu, 1995: The Impacts of Research Field Evaluations on Research Practice. *Research Policy* 24: 349-365.

Marginson, Simon, and Mark Considine, 2000: *The Enterprise Unversity: Power, Governance and Reinvention in Australia*. Cambridge, UK: Cambridge University Press.

Mayntz, Renate, 1994: *Deutsche Forschung im Einigungsprozeß: Die Transformation der Akademie der Wissenschaften der DDR 1989 bis 1992*. Frankfurt a.M.: Campus.

Mayntz, Renate, and Fritz W. Scharpf, 1995: Der Ansatz des akteurzentrierten Institutionalismus. Renate Mayntz and Fritz W. Scharpf: *Gesellschaftliche Selbstregelung und politische Steuerung*, Frankfurt a. M.: Campus, 39-72.

Mayntz, Renate, and Uwe Schimank, 1998: Linking theory and practice: Introduction. *Research Policy* 27: 747-755.

McNay, Ian. 1997. The Impact of the 1992 Research Assessment exercise on Individual and Institutional Behaviour in English Higher Education: Summary Report and Commentary. Chelmsford: Anglia Polytechnic University.

Mukerji, Chandra, 1989: *A Fragile Power: Scientists and the State*. Princeton: Princeton University Press.

Myers, Greg, 1990: *Writing Biology: Texts and The Social Construction of Scientific Knowledge*. Madison: The University of Wisconsin Press.

Mynott, Jeremy, 1999: Publishing: the view from Cambridge University Press. *History of the Human Sciences* 12: 127-131.

Nature, 2002: Peer review reviewed: A controversial change to the peer- review process of Germany's principal funding agency is long overdue. *Nature* 417: 103-103.

Neidhardt, Friedhelm, 1988: *Selbsteuerung in der Forschungsförderung*. Opladen: Westdeutscher Verlag.

North, Douglass C., 1990: A Transaction Cost Theory of Politics. *Journal of Theoretical Politics* 2: 355-367.

Nowotny, Helga, Peter Scott, and Michael Gibbons, 2001: *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*. Cambridge, UK: Polity Press.

Pfeffer, Jeffrey, and Gerald R. Salancik, 1978: *The External Control of Organizations*. New York: Harper and Row.

Rip, Arie, 1994: The Republic of Science in the 1990s. *Higher Education* 28: 3-32.

Röbbecke, Martina, and Dagmar Simon, 2001: *Reflexive Evaluation*. Berlin, Edition Sigma.

Schäfer, Ralf, and Thomas Sikora, 2001: Positionsbestimmung mit Methode. Performance-Messung am Berliner Heinrich-Hertz-Institut für Nachrichtentechnik. *wissenschaftsmanagement* 6: 11-16.

Scharpf, Fritz W., 1997: *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*. Boulder: Westview Press.

Schnitzer, Klaus, and Foad Kazemzadeh. 1995. Formelgebundene Finanzzuweisung des Staates an die Hochschulen - Erfahrungen aus dem europäischen Ausland. Hannover: HIS.

Scott, Richard W., 1991: Unpacking Institutional Arguments. In: Walter W. Powell, and Paul J. DiMaggio: *The New Institutionalism in Organizational Analysis*. Chicago: The University of Chicago Press, 164-183.

Shinn, Terry, 1999: Change or Mutation? Reflections on the Foundations of Contemporary Science. *Social Science Information* 38: 149-176.

Sismondo, Sergio, 1993: Some Social Constructions. *Social Studies of Science* 23: 515-553.

Talib, Ameen Ali, 2000: The RAE and Publications: A View of Journal Editors. *Higher Education Review* 33: 32-46.

Talib, Ameen Ali, 2001: The Continuing Behavioural Modification of Academics since the 1992 Research Assessment Exercise. *Higher Education Review* 33: 30-46.

Taylor, Jeannette, 2001: The impact of performance indicators on the work of university academics: evidence from Australian universities. *Higher Education Quarterly* 55: 42-61.

Thomas, William I., and Dorothy Swaine Thomas, 1928: *The child in America: Behavior Problems and Programs*. New York: Knopf.

Travis, G. D. L., 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: 322-341.

Van Leeuwen, Th. N., and H. F. Moed, 2002: Development and application of journal impact measures in the Dutch science system. *Scientometrics* 53: 249-266.

van Raan, Anthony J. F., 1996: Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics* 36: 397-420.

Velody, Irving (ed.), 1999: Knowledge for what? The intellectual consequences of the Research Assessment Exercise. *History of the Human Sciences* 12: 111-113.

Walford, Leo, 1999: The Research Assessment Exercise and its effect on journal publishing. *History of the Human Sciences* 12: 142-143.

Weingart, Peter, 1997: From "Finalization" to "Mode 2": Old Wine in New Bottles? *Social Science Information* 36: 591-613.

Wessely, Simon, 1998: Peer review of grant applications: what do we know? *Lancet* 352: 301-305.

Wissenschaftliche Kommission Niedersachsen, 1999a: *Forschungsevaluation.*http://www.wk.niedersachsen.de/Themen/FE/FE.htm.

Wissenschaftliche Kommission Niedersachsen, 1999b: *Forschungsevaluation an niedersächsischen Hochschulen und Forschungseinrichtungen: Grundzüge des Verfahrens.*http://www.wk.niedersachsen.de/Materialien/FE-Konzept.pdf.