IHS Political Science Series
Working Paper 48
October 1997

Evaluating Academic Research in Germany: Patterns and Policies

David F.J. Campbell
Bernhard Felderer

IRIHS Institutional Repository at IHS Vienna

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

Impressum

Author(s):
David F.J. Campbell, Bernhard Felderer

Title:
Evaluating Academic Research in Germany: Patterns and Policies

ISSN: Unspecified

1997 Institut für Höhere Studien - Institute for Advanced Studies (IHS)
Josefstädter Straße 39, A-1080 Wien
E-Mail: office@ihs.ac.at
Web: www.ihs.ac.at

All IHS Working Papers are available online: http://irihs.ihs.ac.at/view/ihs_series/
This paper is available for download without charge at:
https://irihs.ihs.ac.at/id/eprint/1022/

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Reihe Politikwissenschaft / Political Science Series **No. 48**

Evaluating Academic Research in Germany
Patterns and Policies

David F.J. Campbell
Bernhard Felderer

Evaluating Academic Research in Germany
Patterns and Policies

David F.J. Campbell
Bernhard Felderer

Reihe Politikwissenschaft / Political Science Series No. 48

October 1997

Institut für Höhere Studien
Stumpergasse 56, A -1060 Wien
Fax: +43/1/597 06 35

David F.J. Campbell
Phone: +43/1/599 91-173
e-mail: david.campbell@ihs.ac.at

Bernhard Felderer
Phone: +43/1/599 91-125
e-mail: bernhard.felderer@ihs.ac.at

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

The Political Science Series is published by the Department of Political Science of the Austrian Institute for Advanced Studies (IHS) in Vienna. The series is meant to share work in progress in a timely way before formal publication. It includes papers by the Department's teaching and research staff, visiting professors, students, visiting fellows, and invited participants in seminars, workshops, and conferences. As usual, authors bear full responsibility for the content of their contributions.

All rights are reserved.

Die **Reihe Politikwissenschaft** wird von der Abteilung Politologie des Instituts für Höhere Studien (IHS) in Wien herausgegeben. Ziel dieser Publikationsreihe ist, abteilungsinterne Arbeitspapiere einer breiteren fachinternen Öffentlichkeit und Diskussion zugänglich zu machen. Die inhaltliche Verantwortung für die veröffentlichten Beiträge liegt bei den AutorInnen. Gastbeiträge werden als solche gekennzeichnet.

Alle Rechte vorbehalten

Abstract WP 48 - Evaluating Academic Research in Germany

Patterns and Policies

David F.J. Campbell
Bernhard Felderer

In this country study of Germany the patterns and policies of academic research as well as the evaluation of academic research are analyzed, through applying the following approach: first of all, a bibliometric survey is carried out that investigates the publication output and publication efficiency of Germany's academic research within international journals; we further investigate whether the results of a bibliometric survey appear compatible with the performance of other indicators. Secondly, discourse and policies of the evaluation of Germany's university research are investigated by addressing issues such as: the current situation; the structural and cultural constraints against evaluations; the general reasons why evaluations of university research will play an increasingly important role in the future; and an overview of specific evaluation initiatives. Thirdly and finally, also the discourse and policies of the evaluation of Germany's university-related research are examined.

Preface

In the following we present an updated version of the Second Report of our currently conducted three-year evaluation study with the title *Die Evaluation der akademischen Forschung im internationalen Vergleich: Strukturen, Trends und Modelle*. The study is generously funded by the Austrian Federal Ministry of Science and Transport (*Bundesministerium für Wissenschaft und Verkehr*) and focuses on the evaluation of academic research within an international context, by comparatively analyzing structures, trends, and models. The Second Report, which we originally forwarded to the Ministry in January 1997, represents a country study that comprehensively investigates the evaluation of academic research in Germany. With the concept of academic research we address both, the *Hochschulforschung* and the *außeruniversitäre Forschung*. Using English terminology, this implies covering university research, *i.e.* research carried out by the higher education sector, as well as university-related research that includes the government sector and the private non-profit sector. Our country study is structured into three sections: first of all, we offer a bibliometric analysis of Germany's academic publication output and academic publication efficiency in international journals; secondly, we reflect discourse and policies on the evaluation of Germany's university research; thirdly, discourse and policies are reflected in reference to Germany's university-related research. Finally, in Chapter 5, we again present a detailed summary of the whole analysis.

Concerning acknowledgments, we want to express our thanks particularly to three decision makers at the Austrian Ministry of Science for their cooperation and support: *Sektionschef* Dr. Norbert Rozsenich, Dr. Edda Korsche, and Dr. Reinhard Schurawitzki. Furthermore, we want to thank all German experts who were willing to participate in those expert interviews that were conducted for the purpose of compiling the country study. The expertise of those German experts represented to us a crucial information base, which was pivotal for understanding the dynamics of the current German academic research system and of those evaluation initiatives that are discussed or already implemented. However, three comments should be added: first of all, not all experts arrived at the same conclusions – this means one must recognize the fact that a wide spectrum of partially diverging opinions exists among experts; so *pluralism* is the rule, and not the exception. Secondly, we explicitly do not claim that those experts would necessarily agree with our final analysis. Thirdly and lastly, in case of errors only we, of course, and none of the above mentioned persons or institutions are responsible.

David F.J. Campbell
Bernhard Felderer

Contents

1. The Evolution of Expenditure on R&D

Germany clearly occupies a strong position with respect to the quantitative financial input in R&D (research and experimental development). In 1993, Germany's gross domestic expenditure on R&D amounted to 2.48% of the GDP – the so-called »national R&D quota«.[1] By this Germany lay ahead of the total OECD average, 2.22%, and the average of the EU, 1.97% (see OECD, 1995b, 16). The only two larger nations, that invest more into R&D in terms of GDP share, are the United States and Japan.[2] When the analytical focus shifts from R&D expenditure to personnel, we receive a very similar picture: concerning the total R&D personnel as well as the core number of researchers in a more narrow sense, Germany again exhibits strength. If, for instance, the number of researchers is put in relation to the labor force, Germany impressively manages to outpace most OECD countries. Taking the year 1992 as an arbitrary example, only four nations – Japan, the United States, Norway, and Finland – demonstrate a better ratio than Germany (see Figure 10). So, as a first conclusion, one can set up the thesis that Germany represents a society in which investment in R&D, either in form of financial resources or as personnel, is given a high priority. The fears of the European Commission, that Europe might run, as a consequence of neglecting the funding of its R&D base, into the problem of a competitiveness gap against Japan and the United States, the two main challengers in know-how and technology, do not apply to Germany to the same extent as to other European countries and economies – although even Germany's R&D quota does not match those of Japan and of the United States (Europäische Kommission, 1995). A more urgent question for Germany seems to be, whether the R&D resources are used efficiently, and whether the structures and the whole institutional framework of the German national R&D system still operate adequately. Or, phrased differently, is there a demand for reform? So one aim will be, to summarize and to analyze the inner-German discourse among experts, which should enable us to offer a first and, more or less, preliminary assessment. In this context the issue of *evaluation and research evaluation* obviously will play a pivotal role.

However, on the other hand, one must also recognize that the German R&D expenditure has come under considerable pressure, and that this process is still in continuation. When measured in terms of expenditure as a percentage of GDP, then Germany's expenditure curve reveals an up-side-down or »wrong« V-curve (see Figure 4). During most of the 1980s, Germany's national R&D quota increased, climbing from 2.45% in 1981 to a first high of 2.88% in 1987 and a second high of 2.87% at the end of the decade, 1989. Afterwards the R&D quota again deteriorated, decreasing each year and falling to a preliminary low in 1994, with a value of 2.37% that is even lower than that of the year 1981. Such a

---

[1] The corresponding term in German would be *nationale Forschungsquote*.

[2] In 1993, respectively, the R&D expenditure of the United States equaled 2.66% of the GDP and that of Japan 2.94%. Even when the OECD adjusted data are used, that qualify the official Japanese figures on R&D as overestimated, Japan's R&D expenditure resulted in a value of 2.73% of the GDP (OECD, 1995b, 16, 78).

development should be judged as potentially problematic, particularly when long or even only mid-term effects are taken into account. What will happen, should this trend be continued in the future, and what will be the effects on the German national system of innovation?[3] And to which extent is it likely that the development of the last years also indicates the trajectory for the next years? Germany's policy and decision-makers are seriously challenged to find an adequate response.

A question, that already arises at that moment, obviously is to investigate which sector or sectors[4] are primarily responsible for such a decline in national R&D expenditure. One fact must be stated clearly: the economy – the so-called business enterprise sector – contributes the biggest share, and with an approximate volume of 51928 million DM, in the year 1994, its expenditure covered 66% of the total national financial investment for research (Figures 1 and 2). At the same time, however, it is also the economy that exhibits the most significant decrease with regard to financial input in R&D. Such a statement can be based on two empirically observable indicators. First of all, when Germany's gross domestic expenditure is analyzed according to the financial performance of individual sectors, then the reduction of commitments of the business enterprise sector appears somewhat drastic. In 1986 the R&D investment of the economy covered a share of 73.2% of the total national domestic expenditure on R&D; by 1994, however, this value had already dropped to only 66.1% (Figure 3). Secondly, when the sectoral expenditure on research is expressed as percentage values of GDP, then the economy's decline turns out to be even more dramatic, taking into account that expenditure dropped from a GDP level of 2% in 1986 to only 1.57% in 1994 (Figure 5). So one could set up the thesis, at least as a disputable ad hoc argument, that the relative decline of Germany's national R&D expenditure – when put in relation to GDP – is basically a shortfall of financial R&D resources which the business enterprise sector allocates in favor of R&D activities. This obviously triggers and feeds a discourse, whether such a behavioral pattern of German industries may harm their competitiveness at a global scale.

In our definition of academic research – that is performed by the academic cluster or academic »sector« – we are following a Continental European or German-speaking tradition which conventionally would summarize under such a concept the »Hochschulforschung« – called in English the higher education sector R&D (which we will abridge as university R&D) – and the »außeruniversitäre Forschung« that we again translated as university-related R&D.[5] This university-related R&D or »außeruniversitäre Forschung« comprises two OECD

---

[3] For further literature on this conceptual topic of *national systems of innovation* see Lundvall, 1992, and Nelson, 1993.

[4] In its statistics on R&D, the OECD usually distinguishes between four different sectors: the business enterprise sector (for short often called »industry« or »economy«), the higher education sector (which may be paraphrased with »universities«), the government sector, and the private non-profit sector (abbreviated as PNP).

[5] See again our First Report of this current research project (Felderer and Campbell, 1995a, 4–5).

standard sectors: the government and the private non-profit sector.[6,7] This clear distinction between university and university-related research in Germany – the same is true for Austria and, more or less, also for Switzerland – stems from the particular fact that in Germany, traditionally, the right to acknowledge an institution with the status of a university comes very close to a public and government-controlled monopoly. In Germany almost all universities are de facto public. Witten/Herdecke, Germany's first private university, was founded only in 1980 (Universität Witten/Herdecke 1994, 12–13).[8] So some academic institutions, which by their functional profile might be qualified as universities in the Anglo-American countries, are a part of the government or private non-profit sectors in Germany, and thus are labeled as »außeruniversitär«, that means university-related. We should also add that the reasons, why now a specific research institution is placed in the government or private non-profit sectors, often appear to be arbitrary or the consequence of a political context or of a political interest, and not so much the result of a purely scientific discourse. Therefore, in practice, there is often not such a great difference between research institutions of the government (or public) sector and institutions that are located in the private non-profit sector. Not too seldomly they carry out similar research tasks.

In 1994 no less than 34% of Germany's domestic expenditure on R&D was performed by the academic sectors (higher education, government, and private non-profit). In absolute figures this equaled a financial sum of 26710 million DM (see Figures 1 and 2). This clearly indicates that academic R&D is of a great importance for the German national R&D system. Even more interesting, however, is the mid or long-term evolution of Germany's R&D expenditure. Expressed as percentage values of the total domestic research expenditure during the period 1986–1994, university research increased from 13.6% to 18.8% and university-related research from 13.2% to 15.2%. Therefore, combined, this implies that academic research went up from 26.8% to 34% (Figure 3). This increase is not only a percentage increase, but also an increase in »real terms«. When put in relation to GDP, then the university-related R&D expenditure stayed constant (0.36% of the GDP in 1986 and 1994); university R&D, however, expanded from 0.37% (1986) to 0.45% (1994). So we can conclude that aggregated academic R&D expenditure went up, during the years 1986–1994, from 0.73% to 0.81% of GDP. Since there was a real growth of GDP during all those years, with the only exception of 1993 (see Figure 6)[9], this clearly underlines the phenomenon of a structural increase of financial resources for academic research.

---

[6] See the *Bundesbericht Forschung 1993,* issued by the German authorities, that offers exactly such a definition for the »außeruniversitäre Forschung« (BMFT, 1993, 61–63).

[7] By this, admittedly, we employ a much broader concept of academic research then the one which was proposed by John Irvine, Ben R. Martin, and Phoebe A. Isard (see Irvine *et al*., 1991, 1–17).

[8] The exact legal term in German would be »Universität in freier Trägerschaft« (Universität Witten/Herdecke, 1994, 6–7).

[9] Concerning those basic economic data, which underpin Figure 6, see the *Main Economic Indicators* reports that are published by OECD (1992d, 1994e, and 1996b).

This comprehensive description of the evolution of the financial base of R&D in Germany, during the 1980s and the first half of the 1990s, allows us to draw the following conclusions that already indicate some crucial issues:

(1) When put in relation to GDP, Germany's domestic R&D expenditure is in a process of steady decline. After a peak in the years 1987 and 1989, the current expenditure level has dropped below that of the early 1980s. This obviously provokes two questions. Firstly, does this already indicate a long-term trend for the future? And secondly, will it turn out that new research foci are only possible when other research concerns will be abandoned, implying that we are entering zero-sum scenarios; in short: must one research institution be shut down, as a prerequisite for opening up a new research center?[10]

(2) A sector-sensitive analysis clearly reveals that this drop of German national R&D funding can be attributed primarily to reductions of R&D expenditure by the business enterprise sector. So if someone seeks a causal relationship, Germany's decline in R&D financing is not so much the case of a general decline across all (or most) sectors, but, first of all, a reduction in the R&D outlays of the economy. Other sectors stabilized their R&D funding base.

(3) Approximately one third of Germany's total domestic R&D expenditure is performed by the academic cluster, that means by university and university-related research.[11] This gives a clear impression of the weight of academic research for the German national innovation system and decisively justifies approaches that aim at evaluations and the implementation of evaluation models for university and university-related R&D. A long-lasting neglection of such issues would seriously undermine the »self-reflexivity«, that means the domestic expertise of German society in reference to its own research base. Or to phrase it slightly differently: *What does Germany know about Germany's research?* And which overview has the German society about research that is conducted in the context of its universities? Concerns like this gain even more ground, when a temporal dimension is included that reveals an underlying dynamical trend: while industry R&D expenditure declines, academic R&D expenditure expands. This indicates a growing importance of the academic cluster for German society and for the international competitiveness of its industries.

(4) In Germany university and university-related research are of almost equal weight. To illustrate such a statement, one can refer to the year 1986, in which 13.6% of

---

[10] Such a »steady state« hypothesis was expressed by Wilhelm Krull and Ekkehard Winter in their preface to a seminar reader that documented a seminar on research foresight, which was organized by Max Planck Society in Munich in November 1995 (Krull and Winter, 1996).

[11] By this three OECD standard sectors are covered, *i.e.,* higher education, government, and private non-profit.

Germany's domestic expenditure on R&D was performed by universities and 13.2% by university-related sectors (see Figure 3). Such an empirical pattern certainly can be explained by several factors, ranging from historical developmental trajectories of a national innovation system[12] to structural conditions, that, for instance, create incentives for the government to allocate resources preferably to university-related sectors.[13] Although this duality equilibrium has shifted, during the last years, gradually in favor of university R&D, which is demonstrated by the performance share of Germany's gross domestic expenditure in the year 1994: the university-related sectors performed 15.2% and the university sector 18.8% (see again Figure 3). Interpreted as a mid-term trend over the years 1986–1994 and calculated as a percentage value of GDP, we can conclude that the R&D expenditure of the university-related sectors stayed quite constant; 0.36% in 1986, and 0.36% in 1994. University R&D expenditure, however, again rose from 0.37% (1986) up to 0.45% in 1994 (see Figure 5). This underlines that within the academic cluster university research gained importance. And this also might indicate that the German public authorities take the German Science Council's (*Wissenschaftsrat*) recommendations more seriously, not to neglect university research too much (Wissenschaftsrat, 1988, 66–89). Such an empirical trend clearly reinforces the demand and call for systematic evaluations of Germany's academic R&D and, in particular, also of Germany's university research.

Our following analysis of the evaluation of academic research in Germany will be structured into three sections:

– In Chapters 2.1 and 2.2 we investigate patterns of publication output and again comment shortly on Germany's publication efficiency. We will refer to only one form of publication output, that consists of articles which are published in international journals as covered by SCI and SSCI. Our interest will be to put Germany's publication profile into relation with other OECD countries.

– In Chapter 3, our main section, our attention focuses on university (or higher education sector) R&D – the so-called »Hochschulforschung«. We will analyze the contemporary discourse and different opinions on how university research should be (or should not be) evaluated, what the practical experiences are, and, beyond that, which strategic scenarios should be developed for the future. In Chapter 3.1 we present an overview of the contemporary general trends in Germany's higher education sector. In Chapter 3.2.1 the current situation, concerning the evaluation of university research, is reviewed. In the Chapters 3.2.2 and 3.2.3 we discuss the

[12] For a generic overview on the evolution of Germany's university-related research (*außeruniversitäre Forschung*) see Hohn and Schimank (1990).

[13] An argument, sometimes raised in that context, is that the reluctance and »immunity« of universities against influences from the outside also deterred potential (public) funders.

structural and cultural constraints that, up until now, prevented a more comprehensive application of evaluations of university research; in Chapter 3.2.4, however, we summarize those arguments that stress why also in Germany the evaluation of university research will become more important in the future. In Chapter 3.2.5, finally, we give an overview of those evaluation initiatives of university research which are currently carried out in Germany.

– In Chapter 4, we summarize those evaluation procedures that focus on Germany's university-related research cluster – called in German the »außeruniversitäre Forschung«. A particular emphasis will be placed on the current evaluation exercise of the »Blue List« institutes, which also involves the development of a more generic masterplan which, in principle, also could be applied to other university-related institutes (in other Central European countries).

2. The Evaluation of Germany's Academic Publication Output and Publication Efficiency

2.1 Theoretical Basis for Bibliometric Analyses

It is widely accepted among experts that publications are one of the most important output indicators for academic research. Even though academic research activities resemble diversified processes which must fulfill several functions, the publication of research results – or at least of major segments of such a research – certainly is an intrinsic goal built into the machinery of academic research. Obviously academic research also can be used more commercially. Patents, transfer knowledge in the natural sciences and in engineering for basically all industries, or application-oriented recommendations for public (and private) agencies in the social sciences would be such examples. Still this does not undermine our thesis, by no means, that *publications might not be a sufficient, but certainly they qualify as a necessary goal* for the whole process of academic research. This would come close to something like a Theory of Complementary Relationships. So, no matter what the academic research communities actually or primarily do, they must make sure that their activities are »also« reflected in their publication patterns. Publications could be interpreted as a level of »conscious self-reflexivity« of the whole scientific system, or, to phrase it in simple terms, *as a mirror,* in which academic research is expressed adequately enough – this legitimates speaking of processes of correspondence between academic research and academic publications, although everybody would admit that this relationship certainly is complicated and diffuse, not one-dimensional and that during certain periods in specific fields (or disciplines) also biases might occur. To give a drastic example: probably no engineering department could survive, in the long run, within a university environment and certainly could not defend something like a scientific competence, without demonstrating a certain publication profile; and the development of application-oriented know-how in engineering does not prevent the issuing of interesting publications.

For a first assessment of Germany's academic publication output – thus taking *bibliometrics* seriously – we referred to the number of articles which are published in international journals, or to put it in more explicit terms, in journals that are covered by SCI (Science Citation Index) and SSCI (Social Sciences Citation Index). Those two indices total about 7000 journals. Evaluated by their contents those indices cover science (e.g., natural sciences and engineering) and the social sciences (including business and economics) in general; the humanities, however, are only represented to a lesser extent. For the quantitative analysis, the following methodology was applied:[14]

[14] For a more detailed description of the methodology, which was applied, see again the First Report of our Evaluation Study (Felderer and Campbell, 1995, 17–22).

- Only four document types were taken into account – articles (research articles, proceedings in journals), letters, notes, and reviews (review articles) – which we will always refer to as articles in the following to simplify our terminology.

- Each article only counts as one, no matter by how many authors (or in which SCI or SSCI journal) it was published. Through such a convention no article is being weighted.

- The country assignment is based on the first »corporate address«, that means the first *vocational address* that is affiliated to the article; usually, but not always, this will coincide with the vocational address of the first author.

Such a methodological approach, that first of all aims at a *quantification of scientific research results* and, secondly, wants to measure publication efficiency by referring to international journals which are covered by SCI and SSCI and, as a consequence, use mainly the English language as a means of communication, almost »traditionally« provokes severe criticism by the German academic communities – or, to be more precise, by certain (sometimes prominent) members of those communities. This criticism is often channeled into following patterns of argument:

(1) To which extent, now considered as a principle question, is it possible to »measure« scientific research output at all? What should be the measurable units at stake? And how, in particular, can the performance of pure or pre-applicated basic research be assessed adequately?; in the German context often the phrase of something like an »application-remote basic research« or *anwendungsferne Grundlagenforschung* is brought into discussion.[15]

(2) What does quantity tell us about quality? Is there any meaningful relationship between both, or should quantity or quality be regarded as two totally independent dimensions which have nothing in common? So consequently one can phrase the question, what just the plain number of articles and, beyond that, the number of citations of those articles really should stand for? Arguments against the use of citations emphasize the potential of biases, which might be the result of so-called »citation cartels«, and would even go so far to propose that the most frequently cited

[15] For instance Uwe Schimank uses this term of an *anwendungsferne Grundlangenforschung,* when analyzing the contemporary patterns of German university research (see Schimank, 1995, 334–336). In Chapter 3.1 we will discuss some of Schimank's theses in more detail.

articles are not necessarily the ones with the highest quality, but those which express a certain fashionable or stylish trend or which contain extremely absurd information.[16]

(3) Whereas the Anglo-American countries can look back on a very long and well established tradition of publishing in journals, the German academic communities developed a different publication culture. In Germany the book as a means of communicating ideas and information has a much higher value. This ultimately implies that a comparative frequency analysis of journal articles will discriminate against the output performance of German academics; but not, because they are not industrious, but because a »wrong« output indicator has been chosen. The term wrong in such a context means that a different quantitative indicator, namely the number of book publications, would lead to different results and conclusions since the publication focus of German scholars and researchers concentrates clearly on books. So for Germany the appropriate publication indicator would be to concentrate on books – by this books represent the crucial arena, where the publication competition of German scientists and researchers takes place. The key importance of the Habilitation[17] for an academic career within the German university system is sometimes understood and seen as an additional manifestation of Germany's book-writing academic culture.

(4) A perhaps extreme, but nevertheless mentioned opinion stresses the following line of argument: given that the assumption is correct that the publication of articles is becoming increasingly important for scientific communication at a global level, then this should not, under no circumstances, be regarded as the intrinsic outcome of something which might be labeled as the rationale of scientific progress. In reality this only reflects that the Anglo-American academic culture has become world dominant and by this is inclined to impose its primary mode of academic publishing as a new standard that decides how communication will take place in the sciences. But should Germany's academics subdue themselves to such a »foreign« cultural hegemony?

---

[16] To give an example, one expert asserted that an article, which contained a very obscure theory of AIDS, was frequently cited as a negative example for flaws or bad research work. However, this expert could not name a proper or verifiable source for this statement of his.

[17] In Germany the *Habilitation,* which is translated into the English sometimes as »higher doctorate« (see Irvine *et al.,* 1991, 52), represents for domestic academic career paths an institutional threshold for becoming eligible to apply for a professorship. A publication profile is part of such a *Habilitation* process; despite differing expectations across various disciplines or universities, the standard procedure would be that the *Habilitation* applicant would have to write a »thick book» – this is particularly true in the social sciences and humanities. Only if an academic became a professor in a country other than Germany, he or she could bypass the domestic *Habilitation* requirement by entering the German science system from an international point of departure. Outside of Germany, Austria, and German-speaking Switzerland, the *Habilitation* is largely unknown. So international experts are often inclined to paraphrase – and perhaps criticize – the *Habilitation* as a »German invention« (for a more comprehensive summary on the *Habilitation* and its function for the German universities, with a particular emphasis on the humanities, see Brenner, 1993).

Why should the academic communities of Germany (and perhaps those of other countries) not have the right to continue their indigenous traditions of publishing?

(5) This Anglo-American academic cultural hegemony additionally selects the validity of a specific content of scientific research. Therefore, research topics which are of interest for the Anglo-American countries will be favored in those international journals, whereas research questions, that are of a prime national or regional interest for Germany, are, at the same time, systematically suppressed. Particularly in the social sciences and humanities this implies an overall discrimination against the performance of the German academic research system and will therefore lead to biased conclusions.

(6) There are *different types of knowledge representation*. So one argument asserts that books enable and develop a more comprehensive view and world picture, whereas information, that is stored in articles, covers only smaller sections of reality. Therefore, there is an implicit danger that an emphasis on articles might lead to a scenario in which our knowledge and know-how structures become increasingly fragmented and diffuse. Taking into account that the quantity of information is growing and that there is also a need for interdisciplinary linkages, then an article-induced knowledge fragmentation would turn out to be even more disastrous. Such a perception ultimately leads to the conclusion that the necessity to keep an overview on information creation demands that book publications should continuously be given a top priority.[18]

Such arguments obviously represent a severe criticism against bibliometrics in general, that means quantitative publication analyses, and, in addition, against a heavily article-based bibliometric comparison of Germany with the international academic community. However, this criticism is not unanimous since, at the same time, there are many experts who express opposite views and who emphasize the usefulness of bibliometric analyses: bibliometrics *per se* and bibliometrics that particularly aims at journals and articles. There are powerful arguments that underpin and legitimate bibliometrics as an appropriate tool to understand how scientific information is being created and how knowledge-based innovation takes place. The practical and empirical examples are manifold (for a summary see Weingart, 1995, and van Raan, 1995).[19] From such a conceptual point of departure, obviously, it makes sense to conduct a bibliometric, that means journal-oriented international comparison of article output and to attribute to such an exercise the quality of

[18] At this point an analyst also might insert a cultural notion, arguing that the compilation of books fits better into the German »way of academic life» than the writing of articles.

[19] Anthony van Raan put forward the notion that SCI is the *best known, most loved* but also *most hated* data base in the world: »Die meisten von ihnen werden wissen, daß der Science Citation Index (SCI) die einzige Quelle für diese Art von Information ist. Dieser Index ist die bekannteste, meistgeliebte und bestgehaßte Datenbank der Welt« (van Raan, 1995, 89).

an evaluation. Within such a line of argument it is, at least in principle, justifiable to offer a first assessment of Germany's academic publication performance that is based on international journal markets, as represented by SCI and SSCI.[20]

Now what are the arguments that are put in place to qualify bibliometrics as a useful tool? The first objection addressing a SCI and SSCI-based output evaluation of German academic research, which proposes that scientific research output cannot be measured at all, is very generic and appears to be a fundamental falsification launch against bibliometrics *per se*. Therefore, we want to discuss this hypothesis somewhat later in Chapter 3.2.4, where we will attempt to demonstrate the weaknesses of such a radical viewpoint that could be paraphrased as »anti-bibliometrics«. The first argument against journal-oriented bibliometrics, which we would like to discuss at this point in more detail, is the premise that within the world of academic research and inquiry *different types of knowledge representation* exist. We believe that such a statement should be taken seriously. At the same time, however, it is also important to realize *the whole spectrum* of typology of knowledge representation, which ranges, classically spoken, from publications – books and journals – over electronic and computer-based retrieval systems or data bases[21] to other modes (for instance scientific films). So clearly journals and journal-based articles are only one element of information supply; but the same is also true for books, which can not claim a monopoly for knowledge representation and whose value, consequently, should not be underestimated but also not overestimated. Therefore, at this stage of the debate, we want to emphasize the following arguments:

(1) Searching for indicators which could express the viability and competitiveness of a national academic research system, one of these indicators could be defined as the capability of a system to create and/or to access information across a wide spectrum of various means. In practice this would imply that an academic research system, which is labeled as viable, would engage itself in very different types of knowledge representation.

(2) As a hypothesis for discussion we would like further to propose that it is wrong or at least misleading to believe that the relationship between those clearly distinct modes of knowledge representation resembles something like a zero-sum game; a zero-sum mechanism, in that context, would imply that one mode aims at displacing other modes. Alternatively we are convinced that the strength of an academic research

---

[20] For further current examples on how bibliometrics is being used for the purpose of evaluating German institutions and disciplines, not only in a national but also in a comparative context, see the following references: Daniel, 1988a; Daniel, 1988b; Daniel and Fisch, 1988; Finkenstaedt and Fries, 1988; Lehrl *et al.,* 1988; Rau and Hummel, 1988; Winterhager *et al.,* 1988; Daniel, 1989; Münzinger and Daniel, 1992; Herbertz and Müller-Hill, 1993. For a very interesting analysis of peer review systems of international journals, exemplified for *Angewandte Chemie,* see also Daniel, 1993.

[21] The impact of INTERNET on the practical daily work of scientists is an excellent example of how electronic means or networks influence the course of academic research.

system must be seen in its capability to combine those various information-creating and information-representing means in parallel to the core academic research process. Such a comprehensive approach ultimately implies that no matter how well developed some of those modes of knowledge representation are, the overall output performance and competitiveness of a national academic research system might be seriously weakened, when only one key mode did not mature – whatever the reasons are.

(3) Now referring back to the most »classical« way how academic scientists and researchers express their research results, that is writing books and writing articles, an analyst must arrive at the following conclusion, when the above said is taken seriously: books and articles are two distinct and crucial modes of knowledge representation, which are not caught in a displacement conflict dilemma, but which, in an ideal situation, mutually reinforce the output performance and thus competitiveness of a national academic research system. Therefore, even when we assume that there is a consensus among experts that the German academic publication culture emphasizes books or »thick« books, our line of argument would indicate the following conclusion: an impressive book publication record of German academics would have to be regarded as a plus; however, such a book record could not be interpreted as an adequate compensation for a possible lack of journal articles. Thus a comparative bibliometric analysis of Germany's presence in those journal markets – as represented by SCI and SSCI – can be legitimized by the argument that those international journals clearly represent a key mode of knowledge representation. Or to put it simply and as seen from a systemic perspective: *publishing many books is no excuse for not publishing articles*. In practice many scientists anyway understand and use articles as a means to promote recent book publications; so the article behaves like the summary or abstract of a lengthy book; however, with the advantage that the article is disseminated widely, since top international journals guarantee a prominent visibility of their articles.[22]

Besides certain information storage limitations, when compared with books, articles in international journals also have their advantages and strengths, which again justifies interpreting them as a distinct and crucial mode (or type) of knowledge representation. We would like to highlight and summarize some of the most prominent features:[23]

– In a rapidly changing world the up-to-date status or half-life period of empirical information is coming under severe pressure. Because of their reduced size (between

[22] Should someone be interested in an example, so compare Schmidt, 1982, with Schmidt, 1983; Manfred G. Schmidt is a leading German scholar in the discipline of political science.

[23] See also our arguments on the importance of articles in our First Report of the Evaluation Study (Felderer and Campbell, 1995a, 9–10).

ten and thirty pages), articles can be written – and often also published – much faster than books, no matter if they are »thick« or »slim«. Therefore, under normal conditions articles will perform a *temporal lead of competitiveness* when compared with books, and books, again, permanently suffer under those temporal restrictions.

– Critics often propose that it is mainly »fragmented knowledge« which is presented in articles. We can refer to three crucial arguments that take much of the heat off such a hypothesis. First of all, the limited page-space of an article could also be interpreted as a challenge, since by this the author is forced to focus his analysis. Unlimited page-space, such as that of books, quite often favors information redundancy. *Therefore, the limited page resources of an article, after all, encourage innovativeness, in the sense that this leads to »few-page« but »content-heavy« publications;* a different circumscription for articles. Secondly, the temporal resources of scientists and researchers are even more constrained. Consequently, *the shorter the publication, the higher the likeliness that it will be read*. So the chances that an article – or the abstract of an article – is being read, are dramatically higher than an in-depth coverage of a book. Within the scientific discourse, however, only that information survives which is retrieved (read) and, most importantly, which is *cited*. Thirdly, because authors write and publish not just one, but normally several scientific articles (and books), this, as a final consequence, helps to craft a comprehensive world view. Using metaphorical references, one could say that *individual articles could be interpreted as individual chapters of an imaginary or »virtual book« that is being written an rewritten permanently*.

– Currently information, that is stored in articles, reveals a higher degree of transparency and international visibility than book-bound information. The reason for this is that retrieval systems or data bases, which deal with articles and the abstracts of articles, have progressed further than indices which refer to books. Since this is also widely known among scientists and researchers, the consequences should not be surprising: scientists are increasingly inclined to publish their research results in international journals, because such a publication behavior coincides with the interest of scientists to place their published output strategically. The growing publication share of scientists from non-English speaking Western European countries within those international journal markets is a strong argument that empirically underpins such theoretical considerations (see again Felderer and Campbell, 1995a, 29–30). Speaking in more comprehensive terms, this appears to be an excellent example for *how closely linked structure and content are*. In other words: since international journals guarantee a pervasive visibility, which again is a prerequisite for a frequent citation coverage, they ultimately attract high quality input, that means qualified article contributions.

– Many scientists, anyway, opt for so-called *publication-cluster strategies*.[24] At the core *center* of such a cluster they would place perhaps one or two voluminous books, where research results are presented and discussed in great length; whereas at the cluster *periphery* those same scientists would produce and publish a series of articles which again reveal a strong referential linkage to the »center books«.

In parallel to those core arguments that convincingly demonstrate the conceptual strengths of articles in international journals, which exactly can be derived from the premises and intrinsic criteria that operate academic research systems, and the usefulness of systematic analyses which focus on such a publication output, we additionally want to assess two main objections that are frequently thrown into discussion (and which we already presented earlier in this chapter):

(1) One standard statement is that quality cannot be quantified. So in that line of argument the mere fact that an individual researcher can demonstrate a long list of publications does not allow any clues as to the quality of his or her writings; and from the individual researcher often an analogy is drawn for the national academic system at aggregated level. Now despite our recognition that a measurement of quality clearly resembles a major challenge with no simple solution, we are, at the same time, also convinced that a total condemnation of such attempts is even less justifiable and comes close to a »naive oversimplification«, which is primarily ideologically based, but offers only a weak scientific justification. The following arguments should support our position.

– Truly, from a quantitatively impressive publication record one can not automatically conclude a high-quality profile of that individual researcher. On the other hand, the opposite relationship is even less likely: that means, if a researcher produces only very few publications, this certainly does not, under no circumstances, imply that they already are of an outstanding quality. To put it simply: a genius in the sciences is normally not discovered on the basis that he or she publishes nothing.

– A systematic survey of the history of thought in the natural and social sciences, and in the humanities, probably would come up with the following conclusion: at least in the majority of cases the most famous and influential scientists and researchers were those persons, who published good quality and who published a lot. Such a proposition could be reinforced also from a different perspective. When focusing more specifically on the quantity of citations, and when we take, for instance, the discipline of political science as an example, then recent studies seem to demonstrate that those scientists and publications, who or which are cited the most, are by tendency

[24] We invented this term because we believe that it describes accurately, in close contact to reality, the publication strategies of scientists.

also those who express a competitive quality (see Klingemann, 1988, and Goodin and Klingemann, 1996).[25] This would falsify or at least substantially weaken the hypothesis that citation frequencies are primarily a question of fashion or styles.

– It is not legitimate – at least not on the basis of scientific premises – to draw simple analogies between the level of the individual researcher and the aggregated level of a national academic research system. Of course, focusing on a specific scientist, we must admit that there is no automatic correlation between publication quantity and quality. This means that when a scientist can demonstrate a long publication record, this does not prove that his publications had a great influence on the scientific discourse. At the aggregated level of the national research system, however, somewhat different processes operate. To us it appears plausible to assume a normal distribution effect, with regard to the quality of publications, for instance article contributions to international journals. Such an assumption implies that most publications perform an average or median quality, whereas the high-quality as well as the low-quality publications would be the exception. So this would reveal an inverted U-curve of quality, with a long and thin tail on the left and right side (the low and high qualities), and a peak of »good standard quality« in the middle. Furthermore we believe that such a normal distribution of quality applies to all national academic research systems, so that, under specific circumstances, the medium quality of publications of different countries would fall into a comparable spectrum. Three such specific circumstances could be mentioned: first of all, countries with a similar level of socioeconomic, industrial, and educational level (e.g., the advanced OECD countries); secondly, publications placed into the same frame of reference, for instance articles in international journals that are covered by SCI and SSCI; and thirdly, one should not forget that scientific communication and the resulting scientific methods and standards are becoming more and more globalized. To deny a comparable median quality of research publications in »comparable countries« would imply the danger of falling back into nationalistic thinking and chauvinistic behavior. And why, anyway, should scientists and researchers from one Western European country produce a much higher publication quality than those of a different Western European country (when, for instance, the GDP per capita, the R&D investment, and the quantities of academic degrees are comparable)? Do we really want to believe, that, as a freely invented example, German scientists in general publish a better quality than French scientists (or the other way around)? And how is it possible to operationalize and then

[25] Although Klingemann (1988, 201) would note, in that context, in a contribution with the title *Zitierhäufigkeit als Qualitätsindikator* [Citation Frequency as an Indicator for Quality]: »Jeder Versuch, den wissenschaftlichen Rang von Fachbereichen zu bestimmen, löst mit hoher Wahrscheinlichkeit Kontroversen aus. Die Ergebnisse von Forschungen, die Urteile solcher Art begründen, werden von der Öffentlichkeit und der Profession in der Regel mit hohem Interesse zur Kenntnis genommen. Da das Meßproblem jedoch komplexer Natur und nicht so einfach zu lösen ist, wird die Gültigkeit des Urteils zumindest von den Institutionen bezweifelt werden, die einen der hinteren Ränge belegen.«

to verify (or falsify) such a hypothesis scientifically? So this finally leads to the proposition that in the overall assessment, and under the assumption of a similar quality unit per publication unit per publication medium, a national academic research system probably demonstrates some form of quality supremacy at least in those areas, where it also produces a larger quantitative output. This hypothesis we want to offer for discussion.

– We consider it as extremely important to emphasize that the number of articles in international journals, which are covered by SCI and SSCI, do not only represent a »meaningless quantitative output«. Contrarily, we are much more inclined to invent and use in that context the concept of a *quantified quality* (or a massified quality), since those journals normally rely on a peer-review system. This means that articles, which are forwarded by their authors for the purpose of publication to such a journal, will be evaluated, primarily on the basis of quality of their content, by reviewers. This procedure is in principle very similar to the peer evaluations of research proposals of academics, who apply for earmarked funding at public agencies; such as DFG (*Deutsche Forschungsgemeinschaft*) in Germany, the Research Councils in the UK, TEKES in Finland, NWO in the Netherlands or FWF in Austria. Therefore, peer reviewed journals guarantee that a standard quality threshold has been implemented which must be passed by each forwarded article, before that article actually can be published. So there is an internal quality security check built into the system. This probably also explains why, for instance, in the context of the evaluation of Dutch university research, as asserted by Anthony van Raan, in the majority of cases traditional peer-reviews and bibliometric indicators would reveal similar results (van Raan, 1995, 93).[26]

– Such observations could be used to develop the hypothesis that even if there is no simple or one-dimensional linkage, there might be, in the »long perspective«, some connections or interactions between quantity and quality – particularly, when systems with similar features are compared, which certainly holds true for the national academic research systems of developed industrial countries during the 1980s and 1990s. The hypothesis, which we would like to emphasize for the course of debate, would be: *in the sciences it cannot be ruled out that, under specific conditions, quantity and quality are two dimensions that communicate (or are correlated) with each other*.

(2) Additional to the question whether quality can be quantified, which we have now discussed in great length, German critics of bibliometric methods often like to refer to a »cultural« argument by emphasizing that in those international journals covered by SCI and SSCI an Anglo-American hegemony is manifest. The core structure of such a

[26] For additional reflections on bibliometrics see also Peter Weingart (1995).

cultural argument would be as follows: this hegemony of the English-speaking countries consequently implies that scholars from non-English speaking regions are discriminated in their ability to successfully forward article manuscripts to international journals, mainly because of two reasons; firstly, regional knowledge outside the sphere of Anglo-American countries is not of a great interest for the English-speaking academic communities – particularly for the social sciences and humanities such effects would be visible, since in those disciplines information and knowledge structures are more regionally dependent than in the natural sciences. Secondly, English-speaking academic communities have their own conceptual and theoretical traditions, which they value higher than the scientific development paths of other countries or world regions; for instance German or other Continental European approaches. Although we admit that some truth surely is attached to such propositions, we also emphasize not to overvalue their influence. The following arguments again clearly »relativate« and limit the impact of a possible cultural factor of Anglo-American academia:

– Patterns of hegemony and center-periphery cleavages always existed during the course of human history. So they represent a fact to which systems, nations, and individuals have to adapt. Regarding economic leadership (see Maddison, 1986, 29–42) or scientific supremacy (European Commission, 1994a, 7–58), always some nations demonstrated saliency. On the other hand, those hierarchical saliency structures were never static, but dynamic and often changed over time. So when a country is placed, at a given point of time, at the »periphery«, then this should be interpreted by that country as a challenge to improve its positioning. Therefore, speaking theoretically, since Germany represents an advanced industrial society, why should Germany's academic communities not be in a position to seriously challenge the Anglo-American dominance in those international journals? This also would imply that Germany should perhaps re-think or re-assess critically some of the cultural traditions that underpin Germany's academia, such as a high emphasis on books written in German or the *Habilitation* system – particularly the *Habilitation* is something which many German experts would evaluate critically in the context of a confident conversation.[27] Speaking in more generic terms, the crucial argument would be: »periphery« or »non-center« countries – or, as in our case, national academic research systems – must seek to develop and implement strategies, by which their peripheral status may be overcome. It does not appear legitimate to use such a one-time peripheral or non-center location as a perpetual argument which would explain, in a deterministic fashion, why an improvement is not possible. When the countries compared demonstrate similar socioeconomic attributes (for instance the group of OECD member countries), then such a proposition becomes even more valid.

[27] In Chapter 3.2.5 we will discuss this so-called *Habilitation* issue in more detail.

- In all of the sciences clearly English is the most important language, and we should be prepared to expect that this dominance or hegemony will even increase in the future. From that observation, of course, one could deduce that English-speaking academic communities of English-speaking countries are favored in their capability to forward manuscripts and publications to international journals and international publishers. Certainly those communities possess some »lingual advantages« over research communities of non-English speaking countries. On the other hand, we also emphasize not to overvalue such a factor. First of all, there were always dominant languages in human history, but this did not prevent a broad evolution of knowledge and technology. In the medieval period and during early modern times surely Latin had for sciences, philosophy and the whole academic and intellectual life a very similar function to contemporary English. This, however, did not prevent a diffusion of academic activities across all of Europe. Secondly, when a country is non-English speaking, then it should develop strategies how to improve the English-speaking skills of its academic research communities. That could be understood as a part of an overall strategy of a »peripheral« or »non-center« national academic research system to improve its performance (see again the preceeding paragraph). And thirdly, it should be expected, at least in a not so far future, that individual members of academic research communities can read, speak, and write in English – particularly of those academic research communities that are embedded in the context of a socioeconomically advanced country (*e.g.,* most of Western Europe). So academic communities must prove their readiness to use English as a means and by this to overcome one of the crucial thresholds against participation in the global dialogue of world-wide sciences. Therefore, during a confident »four-eye« or face-to-face conversation some leading experts are willing to criticize the reluctance of their countrymen against publishing in English: of course, varying across disciplines, such a criticism would be blamed against members of academic research communities in Germany and in France, nevertheless to a significantly lesser extent against research communities in Switzerland, the Netherlands, Scandinavia, and Finland. But for the »younger and new generation« of academic researchers across all of Continental and Northern Europe the trend can be observed that the willingness to use English continuously increases, at least in principle and partially as a means of communication in the context of scientific discourse.

- Now, when evaluating Germany's academic research performance on a bibliometric »hard facts«-basis, that means by counting and analyzing article frequencies in international journals which are covered by SCI and SSCI, we will use as the *primary and »hard« frame of reference* those countries which are part of non-English speaking Western Europe (Continental and Northern Europe). By this the so-called argument of »Anglo-American academic cultural hegemony« in international journals, which tries to make plausible why because of that German academics are systematically

discriminated, looses most, if not all of its substance. Since all non-English speaking Western European countries have started at a very similar point of departure regarding the English language, a German supremacy or deficiency could not be explained or justified by the English-language factor.[28] In our opinion this probably represents the most important bottom-line argument in favor of a journal-based bibliometric comparison of German academic publishing.

2.2 Empirical Bibliometric Analysis of Germany's Academic Publication Output and Publication Efficiency

In Figure 7 – for the OECD countries[29] and Israel – the aggregated quantitative sum for all articles in SCI and SSCI journals is documented for the year 1993.[30] After the United States, the UK, and Japan, Germany already ranks at the fourth position. By this it is clearly demonstrated that Germany represents one of the most important article-producing OECD countries. So, consequently, the German science and research system publishes quite an impressive quantitative output of articles in international journals and as a result occupies a salient position. Articles, »made in Germany«, are a source of research information which must be taken seriously – and it is taken seriously by the global sciences community.

However, when the article publication output is put in relation to population, then the overall picture – in regard to Germany – changes again. In Figure 8 a weighted ratio is presented between articles and population, that means the number of articles per a population of 100,000. Whereas most of the English-speaking and many of the smaller Continental European countries demonstrate a top ranking, Germany suffers from being pushed down to a lower-ranking position. Referring to our country sample, fourteen countries place better than Germany and only eight reveal a subordinate ranking. This does not seem to be very good news for Germany. Obviously one could raise the question whether unification of the two German states, in the year 1990, caused some output-biasing effects? For the purpose of investigating such a question, we also compare in Figure 8 the population/publication ratio of West Germany in 1990, with that of unified Germany two years afterwards (1992). Interestingly, there is almost no difference in ranking. While for West Germany we can count 46.86 articles per 100,000 inhabitants, this ratio decreases only marginally to 45.79 for unified Germany. This demonstrates that unification has not necessarily constrained the article-writing performance of the German science system. And, on the other hand, we can also mention that the East German science system performed perhaps better at publishing

---

[28] See again in our First Report the arguments for having invented the concept of *non-English speaking Western Europe* (Felderer and Campbell, 1995a, 13–14, 33–42). With *non-English speaking Western Europe* we cover the following fifteen countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland.

[29] Excluding Iceland, Mexico, and Turkey.

[30] Regarding additional methodological information, see again Chapter 2.1.

articles in international journals than one would expect. Although, thinking in theoretical terms, we are inclined to believe that this East German performance was restricted primarily to science (*e.g.,* the natural sciences), since in the social sciences and humanities the communist regime imposed tight ideological limits on the scientists.

Now how should Figure 8 – in which article output is related to population – be correctly understood? First of all, again starting from a *theoretical point of departure,* Figure 8 *may not* be interpreted as an efficiency indicator since efficiency is always defined as a relationship of input versus output. Therefore, speaking in strict terms, Figure 8 is an output indicator because population surely does not qualify to be taken as a research input measure. Nobody would claim the possibility of deducing from population quantities the amount of research input. Secondly, now thinking in *empirical categories,* scenarios are conceivable which somehow relativate the validity of the above statement which is at least correctly formulated in theory. The crucial point seems to be the following: when a national research system performs and publishes articles at a very high efficiency level, then such a country – given that the research input meets comparable international standards – again will produce a favorable ratio between publication output and population. This forces us to propose the following comprehensive interpretation of Figure 8: primarily, and at first hand, any population/publication ratio represents an output indicator; secondly, however, this output indicator also contains some »hidden empirical« information on efficiency. So this duality or ambiguity must always be kept in mind, when population/publication ratios are discussed during the course of a professional debate.

In addition, critics, of course, could also claim that Figure 8 does not offer too much new information and could base their accusation on the following line of argument: only by throwing a quick look at Figure 8, one sees that countries (or national research systems) with a high input in R&D[31] are in general also those which express a high population/publication ratio. So it should not be surprising that within that ratio ranking, countries such as Switzerland, Sweden, and the United States occupy top positions whereas countries like Spain, Greece, and Portugal are pushed downwards to bottom positions. This first guess and *ad hoc* impression can also be reinforced by applying simple statistical methods. When the population/publication ratio (see Figure 8) is correlated with the gross domestic expenditure on R&D (as a percentage of GDP) and the number of researchers per thousand labor force – on basis of the year 1992, whenever possible –, then we receive the following results:[32]

---

[31] This could mean, as an example, that a high percentage of GDP is used for gross domestic expenditure on R&D or an impressive number of researchers (in full-time equivalents) when compared with the total labor force (see OECD, 1995b, and 1996a).

[32] See Figures 9 and 10, and again Figure 8, which contain the data basis for our correlation procedure.

- In our first procedural scenario, we excluded Japan as a case country (since we regard Japan in reference to its publication pattern as non-typical, when compared with other advanced OECD countries). Under that condition the correlation of the population/publication ratio to the R&D expenditure (0.6950**) is highly significant, and to the number of researchers (0.6230*) is significant.[33]

- When Japan is included into the country sample, the population/publication ratio still correlates significantly with R&D expenditure (0.5758*) but only non-significantly with the quantity of researchers (0.4022). This again underlines the fact that Japan behaves in many ways atypically when put in comparison to other OECD countries.

Therefore, summarizing the previous in reference to Figure 8, we agree with those critics who propose that in general a high population/publication ratio correlates very closely with the input that a society is willing to invest into R&D or, speaking more structurally oriented, into its national research system. So in this regard our empirical observation is a consensus. On the other hand, however, the relational pattern which is revealed by Figure 8 still contains substantially crucial information, which must be taken seriously by all potential critics. First of all, and this is perhaps the most important message, Figure 8 states clearly: *sufficient research output demands an adequate input in research*. In slightly other words: a properly operating national sciences system depends on resources that society must be willing to invest. Without such a general commitment, the development and – thinking in long-term categories – evolution of a national research system might be endangered. Policy makers, and particularly public policy makers, must be aware of the importance of a solid resource base for R&D. Secondly, the mutually reinforcing correlation between the population/publication ratio, on the one hand, and R&D expenditure and the number of researchers, on the other hand, is only a general statistical message which does not rule out the possibility of exceptions. Now leaving aside Japan, clearly Germany and France represent such deviant cases. Both European countries demonstrate an impressive input in R&D (see Figures 9 and 10), at the same time, however, their population/publication ratio ranks comparatively weak – that of Germany even weaker than in the case of France (see Figure 8). This already could indicate that Germany's academic research system is facing some performance problems, when bibliometrics is taken as indicator and is applied to those international journals that are covered by SCI and SSCI.

Now taking non-English speaking Western Europe as geographical frame of reference for our analysis[34] – which explicitly implies to exclude the Anglo-American countries –, in Figure 12 a relative ranking of the strength of different German disciplines across the whole

---

[33] For our correlation analysis we used a one-tailed Pearson correlation. The term »significant correlation« is applied at a significance level of 99% and the term »highly significant correlation« at a level of 99.9% (see SPSS, 1990, B15–B19, and SPSS, 1992, 283–293).

[34] All together our conceptual definition of *non-English speaking Western Europe* covers fifteen countries (Felderer and Campbell, 1995, 13–14, 33–42).

spectrum of sciences is presented. As mode of calculation the following formula was applied: for all of science and the social sciences and *per* discipline the number of German articles is expressed as a percentage value of the overall publication output of non-English speaking Western Europe. Two different temporal ranges are distinguished in Figure 12; first of all only the year 1993 and, secondly, the average for the five-year duration period 1989–1993. The following observations can be drawn and put forward for discussion (compare with Figure 12):

(1) At the aggregated average level, Germany's article output performance in science and the social sciences behaves quite similarly. In science, during 1993, Germany achieved a share of 25.28% and in the social sciences a share of 27.46%.

(2) In the »hard« science disciplines, Germany performs the best in chemistry (30.76%), followed by engineering (30.28%), physics (29.77%), mathematics (25.52%), and life sciences (22.47%).[35] What makes the case of chemistry so interesting, is that in that discipline the US and English-speaking dominance is the weakest.[36] So for some reasons, which are out of reach of our current interpretational potential, Continental Europe and Germany in particular developed a strong publication profile in chemistry.

(3) The two strongest German social science disciplines are law (48.27%) and political science (47%), then followed by business (33.81%), psychology (33.71%), sociology (29.5%), economics (20.4%), and interdisciplinary research (15.31%).[37] In almost all social science disciplines the average five-year value for 1989–1993 and the one year-value for 1993 are very similar, with the only major exception of business whose output sharply declined in 1993 (23.48%) – this demonstrates that annual changes are more of a gradual character and not that dramatic. Interestingly enough, Germany performs the weakest in interdisciplinary research. This of course could induce very different interpretations: since our category of interdisciplinary research is of an »experimental« character (Felderer and Campbell, 1995a, 169), critics can refer to that for claiming that no meaningful interpretation appears legitimate. Other observers, however, could come up with a very different and alternative hypothesis which would propose that the German university system in general – and also functionally disaggregated to the level of teaching and research – is much too tightly structured within an organizational framework that values the old paradigm of traditional disciplinarity as most important. If this is really the case, then the German university system is facing a big problem since interdisciplinarity or »transdisciplinarity« is regarded by many key analysts as a crucial challenge and demand for how modern

[35] Figures in parentheses refer to the average percentage value for the whole period 1989–1993.
[36] Compare with Figures 44 and 45 in our First Report (Felderer and Campbell, 1995a, 118–119).
[37] Figures in parentheses again refer to the five-year average of 1989–1993.

science and future-oriented research should be carried out (see Gibbons *et al.*, 1994, and Campbell, 1995, 402).[38]

(4) With the exception of only four disciplines – chemistry, physics, mathematics, and interdisciplinary research – we must state as an overall empirical phenomenon that in general, that means for the majority of disciplines (in total eight), the average percentage value for the five-year period 1989–1993 is salient to 1993 (see again Figure 12). This tendency again is reinforced by the fact that at the aggregated level for all of science and for all of the social sciences this pattern again is reproduced, implying a slight decline of the 1993 one-year value when compared with the five-year average. Now how should this phenomenon be correctly interpreted? First of all we must state that in absolute numbers the total output of German articles in aggregated science expanded in the period 1980–1990 as well 1980–1993.[39] This expansion in absolute figures can also be observed for the aggregated social sciences over the period 1989–1993 (compare with Felderer and Campbell, 1995a, 152, 158). Therefore, an adequate interpretation must seek for a different explanation; however, at this point we can already point at the core of the problem or at least the dilemma for the German situation: *despite the fact that Germany's article output is expanding, the crucial point seems to be that other national academic research systems are expanding even faster than Germany*. So the correct question to ask would be: Which countries are the winners? Two hypotheses could be tested empirically: the first would be to compare Germany with all of the OECD countries, that means including the English-speaking nations. The second hypothesis to investigate, would be to juxtapose Germany with the overall performance of non-English speaking Western Europe (all fifteen countries, including Germany).

– **Comparing Germany with the OECD:** In Figure 11 Germany's article output is expressed as a percentage value of the total output of our country sample – twenty-three countries, covering more or less all of the OECD. Focusing on the long-term evolution only of science, one must state a steady decline of Germany's share which dropped from 7.56% (1980) down to 6.89% (1990). In 1991 this value again rose up to 7.76%, however, due to unification, and stabilized at 7.2% in 1993.[40] This gives the impression that when concentrating our attention on »former« West Germany and relating it to the OECD country cluster, then an empirical analysis verifies a relative decline of Germany's article output in science. It takes all of Germany, i.e., unified Germany, to achieve in 1993 a similar percentage value to that of »old« West Germany in 1980. In the social sciences, the situation appears somewhat different.

---

[38] Already in our First Report we shortly discussed that issue of interdisciplinarity in the German context (see Felderer and Campbell, 1995a, 34, and also 36–37).

[39] Beginning in 1991, the data series refer to unified Germany.

[40] Just to remind us: the data series for Germany refers up until 1990 only to West Germany, and beginning with 1991 to Unified Germany (see Felderer and Campbell, 1995a, 18).

Here we can observe a continuos expansion of Germany's article share. Placing at 3.23% in 1989, this value increased to 3.59% in 1991, and even further expanded to 3.95% in 1993. This clearly implies that article expansion in the social sciences is not an arbitrary effect of unification, but seems to be caused by other structural conditions; perhaps the German social scientists are becoming increasingly aware of the importance of international journals. However, since – in absolute terms – the number of articles in the social sciences (SSCI) is overwhelmingly outnumber by articles in science (SCI)[41], Germany's relative output decline in science cannot be compensated by a relative expansion in the social sciences.

– **Comparing non-English Speaking Western Europe with the OECD:** While in Figure 11 Germany's article output was put in relation to the overall OECD performance, in Figure 13 the same methodological approach is applied by expressing the total article output of non-English speaking Western Europe (including Germany) as a percentage value of the total OECD output[42]. In science, during the period 1980–1987, the article output of non-English speaking Western Europe stayed quite constant – in relative terms –, however, after 1987, Europe's share expanded continuously, climbing from 27.08% (1987) up to 30.08% (1993). The same trend is also observable for the social sciences, this means a stable structural expansion from 11.47% (1989) to 14.4% (1993). This ultimately leads to the conclusion that while Germany's academic research system was only able to stabilize its percentage output share in science, all of non-English speaking Western Europe was in that respect much more successful, since it managed a significant expansion of its article output share – when put in contrast to the whole OECD – in science as well as in the social sciences. Only in the social sciences can Germany demonstrate a similar performance increase. This further implies that the significant expansion of non-English speaking Western Europe must be primarily explained as a consequence of the performance of European national academic research systems, other than Germany. So Germany contributed only relatively little – or less than a theorist would (reasonably) expect – to the article growth of non-English speaking Western Europe.[43]

Now after having discussed Germany's article output in great detail, we finally want to summarize our thoughts and offer several hypotheses for a comprehensive assessment and evaluation of Germany's academic research that is based on bibliometrics:

---

[41] When the total article number in science and the social sciences is aggregated for all twenty-three countries of our sample, then, in 1993, science articles covered a share of 87.71% and social science articles were pushed to a margin of only 12.29% (Felderer and Campbell, 1995a, 76–77).

[42] Excluding Iceland, Mexico, and Turkey.

[43] At this point it appears helpful to remember that we are talking about article output and article growth in relative terms. Again thinking in absolute numbers, one should not forget that Germany still represents the single most important article-producing nation of non-English speaking Western Europe (see Figure 7).

(1) Already in the previous chapter (Chapter 2.1) we discussed the theoretical basis and fundamentals for bibliometric analyses that focus on articles which are published in international journals with an incorporated peer-review system – the two currently most important data bases for the purpose of bibliometric analysis in science and social sciences are SCI and SSCI. Particularly in the German context much criticism is raised against such a methodological approach, by formulating arguments such as: (a) research quality *per se* cannot be measured and cannot be quantified; (b) there are different types of knowledge representation, and article-based knowledge is biased towards a fragmented understanding of our world; (c) and the theory of the specific and historic growth of academic cultures, proposing that while Anglo-American scholars prefer to write articles for journals, German (and perhaps also other Continental European) scholars developed an inclination for publishing books – the German tradition of *Habilitation* can be used as an additional reference to reinforce such a statement. Notwithstanding that such criticism must be taken very seriously, we agree with other leading experts who emphasize the value of bibliometric analyses; particularly, when the specific strengths of such an approach are used sensitively and adequately. The following arguments we consider as key arguments that speak in favor of a bibliometric approach which focuses on articles in international journals:[44],[45]

– The decision of inclusion or exclusion (or acceptance or rejection) of forwarded article manuscripts by those journals, which are listed in SCI and SSCI, is generally based on a peer-review system. This comes close to something like a »built-in quality threshold« that guarantees that all published articles reveal a minimum and comparable basic standard of quality. Therefore, it is legitimate to interpret the number of articles in such journals as an indicator for or approximation of *quantified quality,* what again demonstrates the possibility of measuring quality – and falsifies the assertion that quality *per se* is beyond the scope of measurement.

– At the level of individual authors, obviously, one can imagine that a great variance of quality exists with regard to different articles.[46] However, at the aggregated level of those different national academic research systems probably some effects of a *normal distribution of quality* will come into effect. If this is the case (which we and other experts believe), then this consequently implies that the variance between the aggregated median quality of articles of different national academic research systems

[44] In our case this means: journals that are covered by the two data bases SCI and SSCI.

[45] See again Chapter 2.1 for a much more detailed development of our arguments.

[46] Although such a statement appears very plausible, we still face the problem of how to operationalize it methodologically? So it is much easier to propose that different articles reveal a different quality than to develop a scale (or something else) that actually has the capability to measure quality objectively – in practice, the academic community probably could never agree unanimously on one such standard scale that could represent the whole »range of quality«.

is smaller – particularly when countries reveal a similar economic performance and similar socioeconomic standards, which is the case for most OECD nations – than the median publication quality of authors at the level of the individual researcher. Therefore, the frequency counting of articles makes more sense or appears more valid when conclusions should be drawn about the publication quality of different national research systems; at least more valid than a frequency-counting based comparison and assessment of the performance of individual researchers.

– Certainly, there are very *different types of knowledge representation*. Regarding the publication markets more directly, the two perhaps most classical but also most diverse examples would be books on the one hand (no matter if they are »thick« or »thin«), and, on the other hand, articles in journals. Each of those different publication media has, of course, its very distinct profile, which implies that it can refer to a set of functional strengths and weaknesses (or advantages and disadvantages). Especially this relative view is important, since it means that every observer must realize that no particular publication medium offers only advantages or only disadvantages – when, for instance, considering the competition between books and articles. For us the crucial point is the following: *the overall performance and competitiveness of a national academic research system is defined by the extent to which it is in a position to develop very different types of knowledge representation and its capability to operate along a wide spectrum of diverse publication means*.

– Books, obviously, demonstrate certain strengths. Articles, on the other hand, also have their advantages – particularly when they are published in international and reviewed journals –, which make them a powerful tool for storing and presenting information. In that respect the following characteristics of articles should be valued highly: first of all, the limited page space of an article forces the author to focus or to concentrate his arguments. One could say that those constraints on page resources encourage innovativeness and information efficiency. An important side-effect of this is that the chances that a »short« article is read and cited are probably higher than in the case of »thick« books; also the time for reading has developed into an ultimately scarce resource of scientists. Secondly, since a short (»few-page«) article can be written faster than a long (»many-page«) book, articles often demonstrate a temporal lead of competitiveness when compared with books (although the tedious review process of article manuscripts can often create substantial time lags). Thirdly, the data bases and retrieval systems which administrate articles have developed contemporarily a higher degree of sophistication than the computer-based documentation of books – for instance, article-oriented data bases contain information on the vocational address of authors and mostly also provide article abstracts. This, by itself, has encouraged a trend in which scientists are inclined to publish articles in such journals, since this guarantees a higher visibility of their research results (so this

would be an excellent example for how certain »structural advantages« can lead to an increase of the quality of the »content«). Fourthly, and finally, we also must keep in mind that scientists write not just one, but normally a whole series of articles. This prevents an exaggerated fragmentation of knowledge and helps crafting a more holistic image or a more comprehensive scientific understanding of the world.[47] In that context we also invented the term of *publication-cluster strategies,* which are applied by many scientists who publish. This means that many authors would place at the center of such a self-created publication cluster one or two (perhaps »thick«) books, which are reinforced in the periphery by a series of articles that demonstrate a strong referential linkage with those so-called »center« books.

– In reference to the *cultural argument* which claims an Anglo-American or English-speaking dominance in those international journals (as represented by SCI and SSCI), two arguments appear particularly crucial to us: firstly, patterns of cultural hegemony always existed in human history, also in sciences – at this point one could recall the pivotal role of Latin during the medieval and early modern period. So, consequently, when a country or a national academic research system does not belong to the *cultural core* of global sciences, then it must develop strategies of how to cope adequately with such a situation. In that respect Germany (or any other country) cannot demand an exceptional treatment. And, secondly, as a proper frame of reference – for the purpose of comparison – we always use for Germany the geographical concept of non-English speaking Western Europe: all together fifteen countries, covering all of Northern and Western Continental Europe. Since those countries are all defined by the same cultural criterion that they are *not English-speaking,* the »cultural point of departure« for those academic research communities appears quite similar.

(2) After discussing and summarizing theoretical key premises for a bibliometric analysis of article output in international journals, our main concern is the empirical assessment of the research performance of Germany's national academic research system. This ultimately leads to the conclusion that – when the absolute number of articles is used as reference[48] – within the OECD country framework *Germany represents one of the most important and salient article-producing nations* (see again Figure 7). Only the United States, UK, and Japan lie ahead of Germany, while Germany can outpace any Continental Western European country. In comparison to France, as an example, Germany managed, in the year 1993, an aggregated total article output for science and the social sciences which was higher by a factor of 28%. So within the context of global sciences Germany occupies a strong position,

---

[47] In that respect we wrote in Chapter 2.1 that »individual articles could be interpreted as individual chapters of an imaginary or ›virtual book‹ that is being written and rewritten permanently«.

[48] That means, articles in SCI and SSCI journals.

which again implies that at the level of international knowledge exchange and dialogue the German national research system developed into a »collective actor« which cannot be ignored by the international scientific research community. German academic research is a key element for the ongoing process of the global evolution of sciences.

(3) If, however, not the absolute number of articles but different dynamical or »output/output« relations are taken into account, then the empirical assessment of Germany's article output performance leads to a more critical result. Such a conclusion we can base on several key observations:

– When article output is put in relation to population for the purpose of calculating a population/publication ratio, then this ratio has a significant positive correlation with gross domestic expenditure on R&D and the number of researchers per thousand labor force (see Figures 8, 9, and 10). Only three countries deviate from this pattern substantially: Germany, France, and Japan. On the one hand they can be characterized as nations that practically devote a major investment of resources into R&D. On the other hand they only achieve a below-average population/publication ratio with regard to articles. So a critical observer could raise the provoking question whether those three countries represent »worst-case scenarios«?

– When Germany's aggregated article output in science is expressed as a percentage of the total OECD[49] output, then one must state a gradual relative decline over the period 1980–1990. Only because of the geographical enlargement as a consequence of unification, could unified Germany of the early 1990s achieve a similar percentage value like »old« West Germany in the years 1980 and 1981 (see Figure 11).[50] At the same time the fifteen-country cluster of non-English speaking Western Europe – including Germany as one country case – expanded, during the late 1980s and early 1990s, within the overall OECD context in aggregated science as well as in the aggregated social sciences faster and more successful than Germany (see Figure 13). This, ultimately, implies that some of the non-English speaking Western European countries managed a growth rate of their article output that clearly outpaces Germany's growth rate.

– If on a discipline-by-discipline basis Germany's article output is calculated as a percentage value of the total output of non-English speaking Western Europe, then one general trend can be diagnosed: with the only exception of four disciplines, the average value for the five-year period 1989–1993 is higher than the value for 1993 (see

[49] Again without Iceland, Mexico, and Turkey.

[50] In the social sciences Germany's national academic research system performed, during the years 1989–1993, a more successful expansion – in relative terms – than in science.

Figure 12). This additionally supports our previously mentioned observation that the aggregated article output of non-English speaking Western Europe expands faster than that of Germany.

(4) After having investigated those dynamical output/output interactions, the next »logical« step – in an analytical sense – would be to focus the attention on input/output relationships, that means to ask questions in reference to the *issue of efficiency*. In its most basic sense efficiency or the degree of efficiency is always defined by the amount (or quality) of output that is generated by a specific input. So a simple but very clear-cut model[51] for assessing the efficiency of academic research would be to compare the publication output of articles in science and the social sciences[52] with the input of resources into academic R&D.[53] We decided to use two different means of academic research input: monetary resources and personnel. The empirical results of such a survey, however, lead to some very critical questions concerning Germany's academic research performance. The following two data sets support such a conclusion:[54]

– In Figure 8a the interrelation between monetary input and publication output is expressed, by offering and displaying the following ratio: the number of articles *per* one million $ (in PPP) which are invested into academic R&D (mainly for the year 1992). By occupying only the seventeenth position out of a comparative ranking of 21 OECD countries, Germany clearly places only in the bottom third.

– In Figure 8b the focus concentrates on the relationship of personnel and publication output through investigating the ratio of: the quantity of articles *per* one »person year« measured in full-time equivalents (the reference year is 1991). In reference to that indicator, Germany's academic research system can demonstrate a somewhat better efficiency performance. However, by positioning at rank 11 – again out of a sample of 21 OECD member countries –, Germany lies only within a medium and not a top efficiency range.

– Therefore, under the premise that efficiency of academic research is modeled by referring the article output in journals (covered by SCI and SSCI) to the input in

---

[51] The term »clear-cut» in that context implies that every scientific observer knows exactly what this model is addressing.

[52] That means in journals, which are covered by the two data bases SCI and SSCI.

[53] Under the comprehensive »umbrella« term *academic research* we summarize three OECD standard sectors: higher education, government, and private non-profit. In Chapter 1 and in our First Report we broadly developed and discussed the reasons and delivered the legitimation, why we decided to employ that particular concept of academic R&D (see Felderer and Campbell, 1995a, 4–5).

[54] See again our First Report for a more fundamental description of the particular model that we developed for measuring the efficiency of academic research (Felderer and Campbell, 1995a, 53–56). For further information on the validity and methodology of OECD data on R&D, see OECD (1994d).

academic research – monetary and personnel resources –, an empirical evaluation of Germany's academic research performance will arrive at a considerably critical conclusion: *Germany's academic research system exhibits some efficiency problems that should be taken seriously*. Taking this into account and simultaneously keeping in mind that in Germany considerable resources are invested into R&D[55], then we set up the hypothesis for discussion that the weak article output of Germany's academic research system – when the population / publication ratio is taken as the methodological basis (see again Figure 8) – indicates primarily an efficiency problem and not so much an input problem. This additionally reinforces our original interpretation of the population/publication ratio, by emphasizing that in strictly theoretical terms it is only an output measure but empirically speaking it also contains some information about efficiency. In contrast to the – exaggerately so phrased – »worst case scenario« of Germany (a low article output, despite a comparatively generous R&D input), there are, on the other hand, so-called »positive« counterexamples of other non-English speaking Western European countries: Switzerland, Sweden, and to a somewhat lesser extent also the Netherlands can be characterized as countries with a substantial investment into R&D, an above-average article producing efficiency and an above-average article output (on basis of the population/publication ratio).[56] So in that respect this defines a pressure or a necessity for the German academic research system and the German policy makers to analyze and to evaluate those small-sized Western European countries carefully and perhaps to derive from that comparison some new policies in science and research which should be applied in the German context.

(5) For any analyst the challenge now seems to be to decide how those German article output deficiencies should be correctly interpreted? From a theoretical point of view two different interpretative strategies are possible. The first hypothesis would be to assert that those article publication weaknesses resemble only an isolated and not representative indicator, from which no general conclusions should be drawn on the performance of Germany's academic research performance – such a position, in its final and ultimate consequences, probably would imply denying that empirical bibliometric analyses are a very useful tool in understanding scientific dynamics. We clearly represent the opposite view – also shared by a large group of experts – which emphasizes that bibliometric indicators are a very helpful device that should always be seen in a larger context: this, of course, implies that when bibliometric indicators point at some deficiencies or problems then there is a high likeliness involved that also other indicators in other areas are facing some serious constraints. Therefore, the core content of our hypothesis – that clearly emphasizes the necessity of a

[55] This obviously becomes clear when Germany's R&D input is compared with that of other OECD countries (see Figures 9, 9a, and 10).
[56] Compare again in sequence Figures 8, 8a, 8b, 9, 9a, and 10.

comprehensive or overall (holistic) concept and understanding of academic research – can be stated as follows: *Germany's academic research system in particular and German R&D in general exhibit certain deficiencies and competitiveness problems when the international science and technology markets are taken as frame of reference.* Within that definition the key term is expressed by the word »international«. For us those SCI and SSCI journals clearly represent a global market for the sciences, in which scientific knowledge is exchanged at an international level. And there are also other examples which, in our view, convincingly support our hypothesis:

– Compared with their Western European (and American) colleagues, German academic scholars were in the privileged position of a very generous national research funding system. This impression was commonly shared by the Germans as well as their neighbors. For example, the rejection rate of research proposals at DFG (German Research Society), the most important intermediary public agency for earmarked research funding (*Drittmittelfinanzierung*) in German universities, was significantly lower than in other Western European countries. According to some experts, this had lead to a situation in which German academic scholars oriented themselves, in their search for research funds, primarily towards domestic German agencies and institutions. Research funds at the level of the – previously so called – EC were not targeted that systematically or, to phrase it more directly, were completely ignored by some German scholars. Therefore, academic researchers of other Western European countries – from the UK and the Netherlands, to give an example – developed much more determined and goal-oriented strategies, already during an earlier period, driven by an interest to access those EC funds successfully. Now when, as a consequence of unification and other economic problems (see Figure 6), the German monetary R&D resources came under pressure, the policy makers advised the German academics to apply more determinedly for EC (or EU) funds. The German scholars, however, were confronted with two serious challenges. First of all, the capability and capacity of successfully applying for EU research funds demands an in-depth expertise and know-how. Such an expertise can only be developed on the basis of practical experience. Secondly, the German academics must compete against researchers of other Western European countries with a salient tradition in building contact networks at the EU level (Felderer and Campbell, 1994a, 52–53).

– The German national system of research and innovation, in its broadest sense[57], can be characterized as having developed a strong orientation towards Europe and, in economic terminology, towards the EU domestic market; at the same time, however, Germany exhibits certain problems of competitiveness at the international and global R&D and technology markets. In its 1991 report on SCIENCE AND TECHNOLOGY POLICY

[57] This means looking at the academic and industrial research and technology performance comprehensively.

the OECD arrives at the conclusion that »... Germany is a technology-diffusion centre of European rather than world scope; its export/import ratio remains in excess of unity with its EC partners, whereas it has been deteriorating for the Asian zone and is below unity (though advancing appreciably) for the United States« (OECD, 1992b, 132). In a more current study, the OECD emphasizes that with regard to high-tech trade Germany is doing best in medium-high and medium-low technology, but concerning high technology Germany clearly reveals trade deficits and competitiveness problems. In some of the most dynamic technology markets the German industry – or national R&D system all together – suffers from a loss of market shares (OECD, 1994a, 201–202). The internationally renowned IMD, that is located in Lausanne, Switzerland, also concludes a decline of the current competitiveness of Germany's economy. Compared with OECD and other industrial countries, Germany ranked at position six in 1995; in 1996, however, Germany was pushed down to the rank position ten (IMD, 1996, 18–29).[58]

– Another key term for evaluating the international competitiveness of R&D of a given national research system is the technology balance of payments (also abbreviated as TBP) that covers the exports and imports of technology. Those technology products and services, which are taken into account, easily can be understood and interpreted as an outcome of R&D activities.[59] In Figure 19 we express to which extent the technology imports are covered by technology exports of several OECD nations at the beginning of the 1990s.[60] This means that while the United States, in 1993, exported more than four times as many technology products and services than they imported, for example in Austria the exports only covered a ratio of 28% of the imports. Out of a ranking of sixteen OECD countries, Germany rates only at position ten and by this lies below the average (see again Figure 19). What counts perhaps even more is the fact that with a value of 69% (for the year 1993) Germany imports more technology than it exports.[61] Therefore, speaking in the context of a general perspective, this not very favorable (although, of course, also not completely unfavorable) technology balance of payments for Germany could be used as a reference indicator for setting up the hypothesis that in its technology or technology-driven R&D activities Germany is not oriented enough towards the international markets. It would be an interesting question to investigate to which extent such an assessment is also shared by German policy makers and German academic scholars? Now beyond that short-term or mid-term tactical policy question, we can point at some interesting interrelations

---

[58] For a detailed overview on Germany's technological performance, see BMBF, 1996a.

[59] The OECD offers the following definition for TBP: »It consists of money paid or received for the use of patents, licences, trademarks, designs, know-how and closely related technical services (including technical assistance) and for industrial R&D carried out abroad, etc.» (OECD, 1996a, 64).

[60] The mathematical formula is to divide the exports (receipts) by the imports (payments).

[61] The value 100% would indicate a perfect equilibrium between exports and imports, and any value higher than 100% would imply an export surplus.

between the technology balance of payments and other science output indicators that might reveal an underlying pattern which is much broader or more fundamental than many experts might expect. When the technology export surpluses (or shortcomings) are compared with article publication output in reference to population and the efficiency of article output[62], then it appears that some interaction takes place between those different indicators.[63] This, further, could indicate that there might be – using a statistical notion – a factor or a dimension according to which the R&D activities of a nation can be characterized as more »domestically« or more »internationally« oriented (see also Felderer, 1995a).

(6) When referring to the previous and when arriving, however preliminary, at a more comprehensive assessment of Germany's academic or total national research performance, then we are inclined to set up for discussion the following hypothesis: *one of Germany's main problems seems to be that Germany's academic research (and perhaps also national R&D) is biased towards the »domestic pole« and is not enough internationally or outwardly oriented. Germany's academic life can be characterized to have cultivated a very strong* **domestic discourse** *which could be criticized for not demonstrating enough openness for reflecting what is happening internationally.* In that respect Germany's academic research system seems to feature some similarities with France – and perhaps also with Japan. Those three countries can be characterized as »larger medium-sized« nations, when the global context is taken as a frame of reference, so they developed different academic research strategies than the small-sized Western European countries. Whereas countries such as Switzerland, the Netherlands, and Sweden, partially out of a scarcity of national R&D funds and a lack of national publication means, developed much earlier a willingness to publish in English and in international journals (Felderer and Campbell, 1995a, 62–63), the German academic research system expresses the following weaknesses:[64]

– Currently, when a provoking thought should be thrown into discussion, Germany's academic research system – at least partially – may be paraphrased as having developed a *self-referentially closed national domestic discourse* in the humanities, social sciences, and to a certain extent also in science. Of course we know that such a statement certainly is an exaggeration, but sometimes also exaggerations are helpful in pinpointing problems; and a problem definition is a prerequisite for developing strategies that aim at improvements. So when the famous German scholar Niklas Luhmann, who became an influential thinker on modern systems theory, uses

---

[62] For that purpose, compare again Figure 19 with Figures 8, 8a, and 8b.

[63] Within that setting of indicators, Japan again would represent an exception.

[64] Much of this criticism probably also applies to French academic research. At least it would be interesting and intellectually fruitful to discuss such a hypothesis.

self-referentiality as a key term[65], then we could go a step further an ask ourselves the question, whether not the whole German university and academic research system could be described as a system biased for self-referentiality? When an interaction between thoughts and the socio-economic environment is accepted, then the theoretical or abstract models of German scholars might contain much »hidden« or indirect information about German society. Characteristics such as a reluctance against the use of English, the inclination for books and the tradition of the *Habilitation*-system would underpin such an interpretation or impression of a partially closed national discourse which an outside observer might have who should assess Germany's universities.

– German academia still puts a very strong emphasis on the German language and on publishing in German. On the one hand, such a behavior appears reasonable and can further be legitimated by the reference that all nations try to encourage in intellectual life their own native language (or languages). On the other hand, it is a given fact and condition that English has become the most important lingual means for communication and the exchange of ideas in global sciences. Therefore, each non-English speaking national academic research system must take this into account and accept – as a consensus and compromise for the advancement of sciences – the necessity to publish in English. So this should provoke a demand for Germany's academic scholars to publish more in English.

– Beyond the demand for an increased use of the English language, German academic scholars increasingly must become aware of the importance of those international journals – in practice, this means journals which are documented in well-known retrieval systems, such as SCI and SSCI. This, by no means, implies that German scholars should not write books. It does, however, imply that the publication of many books is not necessarily an excuse for not writing articles.

– Peculiarities of the German academic university system, such as the necessity of an individual academic to pass the *Habilitation* threshold for becoming eligible to apply for a professorship position, perhaps should be more carefully – and perhaps also more critically – evaluated by the German policy makers and also the German academic research community themselves. The pivotal question to ask and to respond to would be: Currently, what is the scientifically based legitimation for an institution like the *Habilitation?* Since nothing can claim the attribute of a permanent paradigm in the sciences, this is also true for historically grown institutions that frame

[65] As an example, in reference to cognitive systems Niklas Luhmann applies the following wording: »Wir erkennen die Realität, weil wir aus ihr ausgesperrt sind – wie aus dem Paradies. Oder um es nochmals paradox zu formulieren: die kognitiven Systeme operieren als umweltoffene Systeme, weil und soweit sie selbstereferentiell geschlossen operieren. Offenheit beruht auf Geschlossenheit« (Luhmann, 1988, 294). For a detailed introduction into the thought of Niklas Luhmann, see Helga Gripp-Hagelstange (1995).

the sciences. So the »isolated« argument of a tradition *per se* certainly is not a sufficient justification for the *Habilitation*.[66]

– In our empirical bibliometric analysis we observed that in the area of »interdisciplinary research« Germany's academic article output performance expresses a certain weakness (see again Figure 12). Now despite the possibility of a statistical coincidence or a statistical artifact – which must be taken seriously – this might indicate a problem for Germany's academic research system. So we want to offer the following challenging question for discussion: To which extent is Germany's university system too tightly structured according to the traditional logic of disciplines or *Fachrichtungen* (or *Fachbereiche*)? And to which extent is there a lack of institutional flexibility manifest in German universities? It will not be easy to find a correct and fair answer, although a careful investigation certainly appears necessary.

– Perhaps all national academic research systems – also those of the English-speaking or Anglo-American countries – can be described as being biased towards a strong domestic discourse and by tendencies favoring a self-referential national closure. Since, however, English evolved as the world-dominant language and above all the Anglo-American countries – particularly the United States – are in many ways salient, this has lead to the perhaps »paradox« situation that the domestic sciences discourse of the United States is, at the same time, also a part of the core of the international discourse of sciences. Perhaps this appears to many as unjustified, but the real world never was truly egalitarian, implying that there were always center-periphery relations. So the German academic scholars, like all those of the rest of the world, must accept this as a fact – but it is up to them to change this situation by supporting the build-up of a Western European expertise in sciences that can challenge the United States more effectively.

[66] See Chapter 3.2.5 in which we further discuss in more detail the *Habilitation* issue.

3. The Evaluation of Germany's University Research: Discourse and Policy

3.1 Current Trends in Germany's Higher Education Sector

The German higher education sector already impresses by its size, since it represents in quantitative figures by far the largest Western European tertiary educational system.[67] In 1994, there were no less than 325 institutions of higher education; among those were 88 »classical« universities (*Universitäten*), one comprehensive university (*Gesamthochschule*), 136 *Fachhochschulen*[68], 6 colleges of education (*Pädagogische Hochschulen*), 17 colleges of theology (*Theologische Hochschulen*), 46 colleges of art and music (*Kunsthochschulen*), and 31 colleges of public administration (*Verwaltungsfachhochschulen*). Again in 1994, the German higher education sector counted 1.7 million students, of which 1.3 million or 74.1% were enrolled in a university (including comprehensive universities, colleges of education, and colleges of theology) and only about 24.4% in Fachhochschulen (BMBF, 1995a, 138–141, and 1995b, 72–75). This clearly demonstrates that the majority of students still favors participation at a classical university while non-university tertiary education (for instance at the Fachhochschulen) does not appear that attractive to students – although at the level of newly enrolled students, Fachhochschulen gain in popularity: in 1994, 63.2% of the new entrants enrolled for universities and already 35.7% for Fachhochschulen. This might indicate that the shorter curriculum (on average only four years)[69] and the more labor market-oriented educational mission of the Fachhochschulen seems to meet increasingly the interests of younger students.

In international comparison, Germany's tertiary educational performance is impressive. In that respect the OECD uses two different key ratings: the one is university education as a »percentage of the population 25 to 64 years of age that has attained a specific highest level of education« (OECD, 1995a, 20); the other is non-university tertiary education (with the same methodological reference). Now when the year 1992 is taken as a basis, then

---

[67] For a comprehensive survey and analysis of the German national R&D and innovation system, that also covers the higher education sector, see the report *Bundesbericht Forschung* which was issued in 1996 by the Federal Ministry of Education and Research (BMBF); on page 23, the report shows a diagram in which the national funding of R&D is represented (BMBF, 1996b). Concerning a more historically oriented description of the evolution of the West German research system after 1945, see the analysis by Massow (1986a, 1986b).

[68] The OECD (1995a, 278) defines the Fachhochschulen as follows: »*Fachhochschulen* (trade and technical schools) are attended by students after completion of vocational training and practical occupational experience. These schools provide advanced vocational training (leading for example to masters' or technicians' qualifications». The German authorities (BMBF) provide the following description for Fachhochschulen: »Their mission is to provide highly practice-related training for occupations which require the application of scientific findings and methods or artistic skills. They offer study courses above all for engineers and in the field of economics, social studies, agriculture and design. The study courses are shorter than university study courses. In some German Laender, Fachhochschule graduates have direct access to doctoral studies« (BMBF, 1995b, 70).

[69] See OECD (1995a, 277).

Germany achieved for university education a value of 12%. In Western Europe only Denmark (13%) and the Netherlands (21%) maintain a higher ranking. Outside of Europe, the only two OECD countries with a higher ranking are Canada (15%) and the United States (24%).[70] Although one must add, speaking in favor of Germany, that while in Canada, the US and all English-speaking countries the Bachelor degree represents the first diploma, in Germany – and some other Continental European countries – this first diploma level is the Master degree. So when the Master degree is taken as reference, then Germany achieved in 1992 a value of 13% and by this ranked ahead of Canada (4.8%) and the United States (9.1%).[71] Other critics, of course, could emphasize that the non-existence of Bachelor degrees in German tertiary education should not be interpreted as an excuse for Germany, but actually marks a problem, this means the lack of short study programs. When we again shift our attention from university education back to non-university tertiary education (see at the beginning of this paragraph), then we can state that Germany achieved, in 1992, a value of 10% and by this clearly placed above the OECD average of 8% (OECD, 1995a, 20).

This impressive – in international comparison – diploma output of Germany's higher education sector is a consequence of a *dramatic expansion or »massification« of tertiary education*. The following figures support such an observation. Focusing on the territory of former West Germany – the so-called *Alte Länder*[72] – the number of students increased from 291.1 thousand in 1960 to approximately 1.7 million in 1994. Expressed as a percentage value of the whole population aged 19–26, this implies an increase from 4.3 to 28.8%. Looking at the newly enrolled students (*Studienanfänger*), we receive a very similar picture again for former West Germany. Their numbers grew from 79.4 (1960) to 229.3 thousand (1994). Again displayed as a percentage value of the population, now aged 19–21, this represents an expansion from 7.9 to 34%. For the *Neue Länder*, the territory of former East Germany, similar trends of an increase of the number of students are perceivable, although the percentage level is still significantly lower than in the *Alte Länder* (BMBF, 1995a, 140–141, and BMBF, 1995b, 74–75). When the forecasts of the KMK (*Kultusministerkonferenz*)[73] are analyzed, then those expansion (or »massification«) trends will continue during the next two decades. First of all, it is expected that the number of students for all of Germany will increase from 1.8 million in 1992 to no less than 2.2 million by the year 2010. Only afterwards should this number drop to 2.2 million by 2015. And also

---

[70] For Japan the corresponding value is missing (OECD, 1995a, 20).

[71] Those figures have the following methodological basis: »ratio of public and private university education graduates to population at theoretical age of graduation by type of degree» (OECD 1995a, 218).

[72] The territory of unified Germany is structured into sixteen *Bundesländer,* abbreviated and commonly referred to as *Länder,* this means »Federal States« (or provinces) that can claim a certain autonomy *vis-à-vis* the central Federal Government. Since it has become usual to use the German term *Länder* (or *Laender*) also in English texts, we will follow this practice. Adding a linguistic comment, the German word »Länder» is the plural form of the English »land«.

[73] The *Standing Conference of Ministers of Education* of the German *Länder,* abbreviated in German as KMK, is a public body which aims at coordinating public policies in higher education as well as in research (see also Block and Krull, 1990, 428).

the number of newly enrolled students is expected to increase from 290 thousand, in 1992, to 360 thousand by the year 2010 (BMBF, 1995a, 149–150, and BMBF, 1995b, 79–80). This clearly refutes earlier predictions from the late 1970s which estimated that the number of newly enrolled students would already decrease after 1985. Therefore, it is a realistic scenario to expect that in the future between 35 and 40% of those aged 19–21 will enroll at a higher education institution (mostly the »classical universities«) (see also Lange, 1994, 337). For the territory of former West Germany the HRK (*Hochschulrektorenkonferenz*)[74] estimates that, in the year 1993, 37% of those aged approximately 20 were also eligible for university entrance; and the overwhelming majority of those who are eligible also actually enroll at a university or another institution of the higher education sector (HRK, 1996a, 3–4).

Such a massive expansion of tertiary education undoubtedly promises a series of advantages for an advanced industrial or »post-industrial« society, such as Germany. On the other hand, however, this expansion of university education also implies several new problems with which Germany's university system must struggle. The following issues are being strongly debated in the German discourse on the efficiency of the academic system:

– As Josef Lange has phrased it, there is a fundamental difference if 3% or 30% of those aged twenty enroll at a university.[75] Such a dramatic increase demands that the university curriculum must reflect this new situation and must be adapted. Therefore, the old – perhaps stereotype – notion of educating a small scientific elite certainly does not apply anymore. So the German Science Council (*Wissenschaftsrat*) emphasizes the importance of a vocationally oriented curriculum (*berufsbefähigendes Studium*) that enables or supports a professional activity after completion of the university. In addition, the Science Council underscores the necessity of a scientifically oriented post-graduate training for those few students who plan a scientific academic career.[76] Between those two very different programs – vocational and scientific post-graduate education – universities should always draw a clear distinction (Wissenschaftsrat, 1993a, 3, 36–37).[77]

---

[74] The HRK is an interest group or conference in which the rectors and presidents of universities and other German higher education institutions are represented. As of August 1994, the HRK had 236 member institutions of the higher education sector (*Mitgliedshochschulen*) (for a further summary information on structure and history of the HRK, see HRK, 1994).

[75] »Die Hochschulen, insbesondere die Universitäten, müssen die Konsequenzen aus dem säkularen Strukturwandel in der Bildungsbeteiligung ziehen: 30% eines Altersjahrganges fordern und erwarten zu Recht eine andere Ausbildung als 3 oder 5%« (Lange, 1994, 337).

[76] »Nur ein kleiner Teil der Studenten ist darüber hinaus an Wissenschaft und Erkenntnisfortschritt durch Forschung interessiert und nach Qualifikation hierzu befähigt« (Wissenschaftsrat, 1993a, 36).

[77] »Universitäten müssen in Lehrangebot und Organisation des Studiums stärker zwischen dem auf Wissenschaft gegründeten berufsbefähigenden Studium und der nachfolgenden Ausbildung des wissenschaftlichen Nachwuchses für Wissenschaft, Wirtschaft und Gesellschaft durch aktive Beteiligung der Graduierten an der Forschung unterscheiden« (Wissenschaftsrat, 1993a, 36).

– One main problem of Germany's tertiary educational expansion appears to be that the number of students, who study longer than the regular or *normal time,* is expanding faster than those who are still on schedule. This becomes obvious when the year 1990 is compared with that of 1977. During that period the increase of students in »normal time« (*Regelstudienzeit*) expanded by a factor of 48%; the number of students beyond normal time, however, grew by a value of 106% – and, during the same period, the number of graduates (*Absolventen*) only increased by 20% (see Figure 14). When the average duration of a study program for the first degree is analyzed, then a similar trend can be observed. Based on the number of university semesters (*Hochschulsemester*), the average time for a student to complete his or her first-degree study program at a university lasted, in 1977, 6.1 years; in 1992 this value had already increased to 7.1 years – this means a student must invest currently a whole year longer to achieve the same degree output when compared with students two decades earlier (BMBF, 1995a, 276; BMBF, 1995b, 116). So the basic problem seems to be that the increase of students has lead to some educational performance deficiencies. The German Science Council arrives at the same conclusion by naming and pointing at the following problems: first of all, the empirical duration of first-degree university programs at German universities is long, probably too long. Secondly, this implies that the age of graduates and their entry into professional life is high when compared with other countries. Thirdly, the number of student drop-outs has also gone up. And, finally, the university study curricula do not appear properly prepared for this new situation (Wissenschaftsrat, 1993a, 11–12).

– Therefore, in its recommendations for a general reform of the university-program curricula the Science Council emphasizes the necessity that the first-degree programs should be designed in a way that they can be completed by students in eight, nine, or ten semesters at the most (this means four to five years).[78] And if students cannot meet those criteria, then this should not necessarily be blamed only on the students since it might also indicate a curriculum mismanagement of the universities. In addition, the Science Council put forward the following recommendations: firstly, universities should significantly increase their offer and supply of part-time study programs (*Teilzeitstudiengänge*), so that students have a realistic opportunity to work in parallel to their studies (Wissenschaftsrat, 1993a, 40); secondly, particularly the expansion of the *Fachhochschulen* should be given a high priority in the future, since their curriculum seems to meet those new demands on tertiary education in a suitable way (Wissenschaftsrat, 1993a, 31, 34–36); thirdly, for the purpose of supporting more decisively those students who work on their dissertation and, at the same time, plan and appear able for a professional career in

[78] »Das berufsbefähigende Studium mit dem Abschluß Diplom/Magister/Staatsexamen soll so konzipiert werden, daß es von den Studierenden in einer Planstudienzeit von acht bis neun Semestern, in begründeten Ausnahmefällen in zehn Semestern, abgeschlossen werden kann« (Wissenschaftsrat, 1993a, 36).

> the sciences, the implementation of so-called »Graduate Schools« or »Graduate Colleges« (*Graduiertenkollegs*) should be emphasized (Wissenschaftsrat, 1993a, 42–46); finally, a university curriculum should support the *general ability* of a graduate to be successful in his post-university professional career; but, on the other hand, a university curriculum should not be mistaken as a too narrow training for specific jobs (in German one would phrase it as the difference between *Berufsbefähigung* and *Berufsfertigkeit*) – in that context one must also realize that the life-cycles or »half-life« periods (*Halbwertszeiten*) of information validity and information applicability are becoming shorter and shorter (Lange, 1994, 337).

Perhaps, as the most dramatic current development, it should be noted that this massive performance and output expansion of Germany's higher education sector was not accompanied by a similar expansion of the resources. The following figures document this growing gap between resources and achievement, between *input and output,* or between *supply and demand*[79]: during the period 1977–1990 the numbers of students beyond »normal time« grew by 106%; new students entrants by 73%; students in »normal time« by 48%; and the number of graduates by 20%. During the same period, however, the input or resource increase was much lower: earmarked funding (*Drittmittelfinanzierung*) by DFG expanded by 18%; the total expenditure for universities by 12%; planned positions for students (*räumliche Studienplätze*) by 11%; and the size of university staff (*Personalstellen*) by 7% (see Figure 14; see also Wissenschaftsrat, 1993a, 10; and Lange, 1994, 343). This achievement increase of Germany's higher education sector may also be demonstrated by a different performance indicator, which is the number of article publications in international journals. Now if the SCI is taken as reference base – covering the journals in science – then we can observe for Germany 22954 published articles in 1980, and 33795 in the year 1993 (Felderer and Campbell, 1995a, 152). This is an increase by 47% which lies clearly above the already mentioned resource increases for the higher education sector (see again Figure 14).[80] When the total expenditure for Germany's higher education sector is expressed as a percentage value of the GNP (*Bruttosozialprodukt*), then according to an analysis of the German Science Council the expenditure declined from 1.32%, in 1975, to 0.93% in 1992 (HRK, 1992a, 17; Lange, 1994, 342; HRK, 1996a, 5–6; see also Wissenschaftsrat, 1988, 244). In international comparison, Germany's higher education sector also appears somewhat »underfinanced«. Two indicators appear crucial in that context. First of all, when the total public expenditure for tertiary education, in the year 1992, is expressed as a percentage of GDP, then we receive for the territory of former West Germany – the so-called *Alte Länder* – a value of 1%; of the OECD countries, only Belgium, France, Italy, Spain, and Japan achieve a lower score. And, secondly, when the public expenditure per student in

---

[79] The Science Council would phrase this development as *Auseinanderentwicklung von Nachfrage und Ressourcen* (Wissenschaftsrat, 1993a, 10).

[80] Although one would have to add that those SCI journals do not only have article contributions from authors with a vocational address at a university, but include also such authors who work at institutions in the university-related research cluster (*außeruniversitäre Forschungseinrichtungen*).

tertiary education is calculated in US dollars (using PPPs), we receive for former West Germany's territory, for 1992, a value of 6550 $. Of the OECD countries, only for Austria, France, Italy, Spain, and New Zealand is the public per-student expenditure lower (see OECD, 1995a, 75, 92).

Now before continuing our analysis it would be helpful to mention the mode of how universities – or other institutions of the higher education sector – are being financed. The funding base of a Continental European university – which are mostly not private in an Anglo-American sense – typically exists as **two different categories** or types. The **first category** is the so-called »basic funding« (*Grundfinanzierung* or *Grundmittelfinanzierung*) or »institutional funding« (*institutionelle Finanzierung* or *institutionelle Förderung*) which is technically and methodologically called by the OECD as GUF (General University Funds). GUF are public funds or public »block grants« that are annually transferred to the higher education institutes, which use this money for research, teaching, administration, and other activities (such as, e.g., health care). Conventionally it is only *ex-post,* that means afterwards possible – by employing questionnaires or other means – to know empirically (and by this with a certain accuracy) for which function which share of the monetary resources was used. In a proscriptive or forecast mode, that means *ex-ante,* only estimations, based on previous experience in the past, can be applied to predict the current or future use of GUF by higher education institutions. So in a certain and perhaps extreme interpretation, GUF-funded university activities – for research, teaching, *etc*. – always behave like a »black box« (or several black boxes) and it is difficult to impose transparency into those internal and »hidden« processes of university life. *Therefore, evaluations are regarded by many as a key tool to access and visualize the inner dynamics of university activity*. The **second main financing category** of Continental European universities is the so-called »earmarked funding«, which in German is called *Drittmittelfinanzierung*. In OECD terms this would include »Direct Government Funds« and contributions from the following sectors, all with a project or program orientation: business enterprise, higher education, private non-profit, and funds from abroad (not included, obviously, are the General University Funds). The German Science Council offers the following list of sources for *Drittmittelfinanzierung* – earmarked funding – of German universities: (1) DFG[81] and other primarily public research-funding agencies; (2) Federal Ministries; (3) Ministries of the *Länder* (provinces); (4) institutions supporting and promoting the post-graduate training of young scientists (*wissenschaftliche Nachwuchsförderung*); (5) international organizations; (6) foundations (e.g., the *Volkswagen*-Foundation); and (7) the business enterprise sector and its interest groups (*Verbände*) (Wissenschaftsrat, 1993b, 13). Assessed by its functional profile, earmarked-funded or *drittmittelfinanzierte* university activities – at least in

[81] DFG is the most important public or semi-public intermediary research-funding agency in Germany for university and university-related R&D. Thus it fulfills a function comparable to the Research Councils in the UK (see also Atkinson *et al*., 1990, and Felderer and Campbell, 1994a).

the German context – are almost completely directed towards R&D (Wissenschaftsrat, 1993b, 8).

So to conclude with a simplified message, one can state that the funding base of a German university consists of public GUF for teaching and research, and public and private earmarked funds mainly for research. This then implies that university R&D has two funding components; public GUF and public/private earmarked money (for a further going discussion see OECD, 1989a, 44–46; see also OECD, 1994d, and OECD, 1995c). When the specific weight of those two components is assessed, one obviously must keep in mind that GUF – the »basic« funding – represents by far the most important monetary source for university research. In 1991 no less than 71.4% of higher education sector R&D was funded by R&D. Public earmarked funds contributed 21.6% and private earmarked funds contributed only 7% to the financial base of university R&D (see Figures 17 and 18). This implies that in Germany more than two thirds of higher education sector R&D activities depend financially on public basic transfer funds – so-called GUF – while the funding share of earmarked funds is less than a third.[82]

Now when the growth of those two different university funding components – GUF for teaching and research and earmarked funds for research – is analyzed during the recent years, then one arrives at the following key conclusion: in the period 1980–1990 the earmarked funds grew significantly faster than GUF. While the earmarked funds almost doubled, general GUF only increased by 37% (see Figure 15). This widening gap between GUF and earmarked funds can also be demonstrated by using a growth index. When, for GUF as well for the earmarked funds, the index is set at 100 in the year 1980, then this value climbed up to 189.2 for earmarked funding by 1990; but for GUF only an index value of 137.3 was achieved (see Figure 16).

How should this differing growth between GUF and earmarked funds be properly interpreted and what are possible implications? In that context we would like to indicate three issues which we regard as particularly interesting:

(1) First of all, potential funders of university research are increasingly inclined to use the means of earmarked-funding mechanisms for allocating monetary resources to the higher education sector. Since the public or the *Länder* Ministries of Education are the prime funders (for GUF)[83] this then suggests the following hypothesis for discussion: either public policy makers somehow distrust the universities as to whether this basic transfer money (GUF) is being used really efficiently; and/or those public policy

[82] For more data for an international comparison of OECD countries see OECD, 1995c; see also Felderer and Campbell, 1994a, 266–267.

[83] In Germany the *Länder* governments, and not the central Federal Government, are responsible for the GUF-funding of the universities, depending on their geographical location (BMFT, 1993, 15; HRK, 1996a, 2; see also OECD, 1995c, 146).

makers believe that earmarked-funded university research promises a higher quality and efficiency. *Furthermore, this again can be interpreted as an indication that key policy makers – and key experts – consider a structural reform of the university complex as crucially important.*

(2) While it is conceivable and workable to evaluate earmarked-funded university research *ex-ante* (*e.g.,* project proposals) and *ex-post* (*e.g.,* project reports), we face, in the case of GUF-funded research, a completely different situation; because of structural reasons it is more or less only *ex-post* possible to analyze and thus to evaluate GUF-funded university R&D. Referring back to our previously used metaphor of a black box when speaking of GUF, this, in reference to a temporal dimension, implies that only *afterwards* transparency or »light« can be shed into the university-internal processes of R&D. Since more than two thirds of Germany's university R&D are GUF-funded (in 1991 exactly 71.43%)[84] this leads to the following conclusion: *as a structural consequence of Germany's university-research funding structure one can argue decisively that particularly in the case of Germany there is a strong need for systematic ex-post evaluations of university research*. Referring to the issue of policy and the question and challenge of the proper means, policy makers – either of the public or policy experts who themselves are members of the academic community – should become increasingly aware of the usefulness of the tool of ex-post evaluations, particularly when a reform of the universities is being discussed.

(3) As a third and somehow alternative interpretation of this »relative« stagnation of GUF, *i.e.* the basic funding of universities, it could also be understood – at least by some analysts – as an indication that the funding base *per se* of the universities is coming under considerable pressure. Universities are facing serious financial constraints.

Indeed, in contemporary German policy discourse there is a debate whether the German universities are *underfinanced* and, if this is the case, then to what extent – so two key terms in German that are often used during discussions in that context are *Unterfinanzierung* and *Finanzierungsdefizit*.[85] The HRK which, of course, is an interest-representing body and by this something like a »functional lobbyist« for the universities, emphasizes that the German higher education sector is underfinanced and postulates a major financial gap. The HRK estimates the annual financial deficit or, to phrase it from a different conceptual perspective, the annual additional financial demand for the German higher education sector – the German universities – to be between 6 and 9 billion DM. Based on the price indices of 1993, the annual deficit amounts to 6 billion DM. However, when the relationship of input resources and output performance (for example, the number

[84] See again Figure 18.

[85] »In international vergleichenden Statistiken wird sichtbar, daß der Hochschulbereich in der Bundesrepublik Deutschland im Vergleich mit anderen hochindustriealisierten Ländern unterfinanziert ist« (HRK, 1996a, 18).

of students), based on the year 1977, is taken as a methodological reference, then the annual financial gap approaches the figure of 9 billion DM, of which 6.8 billion would have to be invested into the *Alte Länder* and 2.2 billion into the *Neue Länder* (HRK, 1996a, 22–23; see also HRK, 1994a, 56). Josef Lange, general secretary of the HRK, also quotes the figure of 9 billion DM as an appropriate estimation for the currently existing financial gap (Lange, 1994, 344). The German Science Council, a more university-independent institution that certainly cannot be classified as an interest-driven university lobbyist, arrives in principle at a judgment very similar to that of HRK.[86] Also the Science Council concludes that the universities are underfinanced (»unterfinanziert«), particularly that the basic funding is insufficient. Since competition for earmarked funds increases, this could imply that the research competitiveness of universities versus university-related research institutions (*außeruniversitäre Forschung*) might deteriorate. Although the Science Council does not quantify the additional financial demand of the universities in exact numbers, it recommends that universities should be given a higher priority concerning public expenditure (Wissenschaftsrat, 1993a, 11–12, 63; see also an earlier report of the Science Council: Wissenschaftsrat, 1988, 35–89, 233–269).[87]

In his new book *Hochschulforschung im Schatten der Lehre*[88] Uwe Schimank from the Max Planck Institute for *Gesellschaftsforschung* (MPIFG), in Cologne, presents an in-depth analysis of the consequences of this financial underinvestment on academic university life. Schimank's main thesis is that the relative decline of basic funding of universities puts replacement pressures on university research. Since the number of students increases continously, this produces the following *vicious circle:* the teaching load and teaching burden on professors and other academic university staff obviously cannot be reduced. So more and more resources are being re-allocated, that means withdrawn from research and invested into teaching (and administration). Concerning available time, the academic university staff developed strategies to safeguard sufficient research time. On the one hand, the growing teaching burden demands an expanding input of time. On the other hand, however, through the application of different strategies the academic university staff could avoid a redistribution of research time in favor of teaching. Speaking in empirical terms the most important »strategy« – or consequence, to phrase it more honestly – was that the individual members of the academic university staff had to increase their personal working time; and only this general acceptance to work overtime prevented a significant decline in the »temporal base« of university research. So where university R&D really is suffering is

---

[86] For a quick survey on the Science Council – in German *Wissenschaftsrat* – see Block and Krull, 1990; see also Krull, 1994, 207–209.

[87] »Die Hochschulen sind unterfinanziert. Die Realisierung des vorgeschlagenen Konzepts für die künftige Hochschulpolitik erfordert einen höheren Stellenwert für die Hochschulen in der staatlichen Finanzpolitik. Gelingt es nicht, in den kommenden Jahren den Hochschulen wieder einen angemessenen Stellenwert in den öffentlichen Haushalten zu geben, wird die schleichende Auszehrung der Grundausstattung der Hochschulen weitergehen« (Wissenschaftsrat, 1993a, 63).

[88] The title *Hochschulforschung im Schatten der Lehre* could be translated into English as: University Research in the Shadow of Teaching.

during the process of allocation of monetary resources. An expansive share of capital and operational expenditure (*Sachausgaben*) is absorbed by teaching. This ranges from small-scale implications for daily university life – for example, the production of xerox copies or the ordering of library books in favor of students' interests – to larger issues, such as questions of long-term investment; *e.g.,* the conflict of interest between research equipment or the implementation of new facilities that support teaching. Therefore, coming down to those basics of monetary resource allocation and assessed empirically then – according to the analysis of Schimank – university research is seriously exposed to replacement pressures of teaching activities. Regarding the practical consequences of those replacement pressures in daily university life, Schimank presents, based on interviews with professors and other university staff, many interesting examples with a partially obscure connotation (Schimank, 1995, 301–336; see also Schimank, 1992a, and Schimank, 1992b).

The current German university can be characterized as being guided by the ideal of simultaneously delivering research and teaching. Uwe Schimank also discusses the possibility that in the future there might be a more clear-cut distinction between so-called Research Universities (Forschungshochschulen) that primarily perform R&D and the »Normal« Universities (»Normale« Hochschulen) that emphasize teaching. Schimank arrives at the conclusion that the current status quo should be continued, which means a combination of research and teaching within the context of a university institution. He justifies his recommendation by stressing that teaching offers a crucial »piggy-back legitimation« (»Huckepack«-Legitimation) for some areas of basic research or »application-remote« basic research – anwendungsferne Grundlagenforschung – that otherwise might not be performed anymore (Schimank, 1995, 323–336). Schimank reinforces his statement also empirically by referring to the fact that a majority of West German professors (at least of those who where sampled during two different surveys) prefers the status quo. Although one should add that a clearer distinction between research and teaching universities – or, as a variation of that, a distinction between research and teaching professorships – has significantly gained popularity among professors when the mid-1970s are compared with the early 1990s (Schimank, 1995, 323–325).[89]

Coming down to a summarizing conclusion, we again would like to indicate in more detail the following arguments that should help explain those phenomena of an underfinancing of universities – *Unterfinanzierung der Universitäten* – and of the underfunding of university research:

---

[89] In an earlier article, also authored by Uwe Schimank, the interaction patterns between funding mechanisms of different actors are analyzed. Schimank develops the scenario of cyclical interactions for which he uses the expression of a »cyclical auto-dynamic of technology transfer« (Schimank, 1988; see particularly the diagram on page 336 of that article).

(1) Universities and university research cannot claim such a high priority on the public expenditure agenda as perhaps desired by the universities. Speaking in simple terms, this would imply that public policy makers give preference to a flow and distribution of public monetary resources in which the benefits for universities are clearly limited and non-expansively restricted. However, this could also indicate a certain failure by the universities themselves, in the sense that they were not that successful in convincing the public and the politicians of their importance and value for society and the national economic performance.[90]

(2) Public expenditure in Germany is facing serious constraints. There are massive pressures to cut back government spending. The tremendous costs of German unification – that are still continuing (Campbell, 1994, 887) – and the balanced-budget demands as a consequence of the EU plans to implement a single European currency do not allow much scope to increase public expenditure, since any increase would have to be compensated with financial cut-backs in other areas; an almost impossible task. Therefore, in the current situation it would be enormously difficult for universities to persuade a ministry – either at the federal or *Länder* level – to increase the public financial base for institutions of the higher education sector.

(3) Among public policy makers – and experts in general – there is some dissatisfaction with the structures and performance of Germany's contemporary university system. And this dissatisfaction explains to a large extent also the reluctance of the public to increase the public funding base for university activities, such as university research. The traditional resistance of parts of Germany's academic university community against systematic evaluations that would impose consequences on universities depending on evaluation results, has had the negative effect – as some experts would put it – that the »distrust gap« between the public and the universities could not be bridged. To illustrate this assessment one could cite Wilhelm Krull who asserts that certain analysts are willing to phrase the West German mass university as the Achilles heel (*Achillesferse*) of Germany's national R&D system.[91] Or, as Hans-Uwe Erichsen sees it, government and universities are two actors that partially observe each other with suspicion (Erichsen, 1995a, 25).

---

[90] In Chapter 2.2, in which we presented an empirical bibliometrics-based analysis of Germany's academic publication output and efficiency, we set up the hypothesis that Germany's academic discourse has developed – at least partially – a bias of self-referentially closure.

[91] »... wird die westdeutsche Massenuniversität seit geraumer Zeit als ein schwerkranker Patient oder gar als ›Achillesferse der Bundesrepublik Deutschland‹ bezeichnet« (Krull, 1994, 206).

3.2 Evaluation Policy of University Research in Germany

3.2.1 Some General Comments on the Current Situation

Is it possible to evaluate university research properly? And, beyond that, is there even a demand or necessity for evaluating university research? Those questions certainly belong to the most controversially discussed issues in the current discourse of experts, academics, politicians, and decision makers in general, concerning the performance, quality, competitiveness, and efficiency of Germany's contemporary university system and also, more focused, of Germany's university research.[92] Among several experts there is a widespread perception that already the evaluation of university research *per se* represents a concept or a mode of thinking which is exposed to a lot of criticism – and is even rejected – by a large segment of Germany's academic community. This means that the »anti-evaluation« attitudes of parts of German academic university communities appear to be the product of a distinct historical tradition and thus are deeply rooted in academic culture. Therefore, many experts express the consensus that currently, as a statement about the *status-quo* situation, *something like a consequent evaluation system is still not extant and is still not applied for Germany's university research*. In other words, it is an extensively debated issue how such an evaluation system should be designed and what the adequate means for implementation are. To illustrate such a conclusion, Hans-Dieter Daniel, for instance, claims that universities up until now failed in their duty to be transparent to the public and to inform the public of their performance; and such a transparency insufficiency certainly is not helpful in overcoming the current problem of underfunding of university research (Daniel, 1995, 205, 208).[93] When put in international comparison – either with Western European countries or non-European nations abroad, for instance the United States, Canada, Australia, or New Zealand –, then often the criticism is raised that there is a fundamental time lag between Germany and other national academic research systems; that means with regard to the evaluation of academic R&D, Germany cannot claim to be a leading country but resembles something like a »late follower« – at least this is the impression of several experts. In that context, Peter Weingart was willing to raise openly in »public« (at least the »academic public«) the perhaps most pronounced and rigorous criticism by asserting that – in reference to the matureness of indicator-based evaluation systems of academic research – Germany is lagging more than twenty years behind the

---

[92] To simplify terminology we will abbreviate in this Chapter the institutions of the higher education sector simply as »university« or »universities«. Concerning the institutional diversity of Germany's higher education sector, see again our survey at the beginning of Chapter 3.1.

[93] »Ihrer Informationspflicht gegenüber der Öffentlichkeit sind die Hochschulen bisher jedoch kaum nachgekommen. Bereits 1988 hatten Experten aus Wissenschaftsorganisationen und der Wissenschaftsadministration festgestellt, daß die Erprobung von Konzepten zur Leistungsberichterstattung im Hochschulbereich im internationalen Vergleich längst überfällig sei und daß sich die Massenmedien des Themas annehmen würden, wenn nich die Universitäten selbst mit wissenschaftlich fundierten Methoden zur Leistungstransparenz beitragen würden. Nach Auffassung des Wissenschaftsrates verlangen die offenkundigen Funktionsmängel und die öffentliche Kritik an den Hochschulen nach internen und externen Verfahren der Evaluation« (Daniel, 1995, 205).

United States.[94] However, within German academic research there again seems to be a certain difference between universities and university-related research. So while the universities are still lacking a system of systematic evaluations, the university-related sector (*außeruniversitäre Forschung*) can rely in that respect, at least partially, on a certain evaluation-fostering tradition.[95] If we are willing to agree with such a statement, then this would indicate the existence of a domestic university/non-university evaluation cleavage within German academic R&D (in Chapter 4 we will discuss the evaluation policy of Germany's university-related research more focused).

In addition to those individual expert opinions, including research-assessing institutions, such as the German Science Council, we arrive at the critical conclusion that currently a comprehensive evaluation of the overall performance of German universities is still missing.[96] Furthermore, in the Science Council's assessment the evaluation of teaching is even in a worse condition than the research evaluation;[97] a statement, to which also »private« experts would agree.[98] So also teaching-oriented evaluations claim a high priority on the policy recommendation agenda of the Science Council (Wissenschaftsrat, 1996a, 42–44).[99] Thus we can state currently a broad range of initiatives that primarily aim at teaching and the possibility of assessing its efficiency and quality. The University of Mannheim, for instance, launched such a project on its own – called »Evaluation der Lehre«, that means *evaluation of teaching* (see Daniel, Thoma, and Bandilla, 1995; and Daniel, 1995b).

When we set up the hypothesis – in consensus with other experts – that currently German university research is lacking a »systematic« evaluation, then such a statement can be derived from the following two facts:

– First of all, a comprehensive *ex-post* evaluation of all of Germany's university research at the national level – for a given year or a several-year period – was never carried out. Such an exercise clearly would mark a watershed for German academic culture and

---

[94] »Die indikatorengestützte Evaluierung von Forschungsleistungen ist im deutschen Wissenschaftssystem auch etwa zwei Jahrzehnte nach deren Einführung in den USA und nach der inzwischen in allen westlichen Industrieländern zunehmenden Verbreitung noch immer ein Skandalon« (Weingart, 1995, 73).

[95] See, for instance, an article of Wilhelm Krull, published in 1995, in which he describes in great detail the self-evaluating procedures which are applied by the Max Planck Society in Germany (Krull, 1995).

[96] »In ihrer traditionellen Verfasstheit kann die Hochschule die Aufgaben der Leistungsevaluation, der leistungsgesteuerten Ressourcenverteilung und der eigenverantwortlichen Anpassung an die Anforderungen der gesellschaftlichen Umwelt nicht wahrnehmen« (Wissenschaftsrat, 1993a, 18).

[97] »An deutschen Hochschulen fehlt es an systematischen Verfahren zur Evaluation der Lehrprogramme und der Leistungen in der Lehre« (Wissenschaftsrat, 1993a, 53).

[98] Block and Krull refer to the following observation: »Whereas evaluation of research projects or institutions by peer review is widely accepted in the scientific community, the evaluation of teaching still meets with reservations in universities« (Block and Krull, 1990, 435).

[99] In that context we regard it as highly interesting that the German Science Council refers explicitly to the Dutch university system as a positive example for how university teaching can be evaluated meaningful. By this the Science Council admits the possibility of useful policy recommendations for a specific country which are international-comparison based (Wissenschaftsrat, 1996a, 46–47).

would resemble »pioneer land« for the development of a policy expertise for all the involved actors.

– Secondly, the experience of systematically evaluating single universities has also not been developed that consequently in contemporary Germany – when there is agreement that a so-called systematic evaluation would have to include substantial quantitative indicators. Those »institutional« evaluations, of a specific university, obviously could not focus only on research, but would also have to cut across other issues such as teaching and administration.

At the same time, of course, it must be admitted that at the »disaggregated« level, that means non-national or sub-national level – when such a metaphoric phrase is permitted or applicable – many evaluation initiatives are carried out on a permanent basis. Those single evaluations cover a broad spectrum: so on the one end of the spectrum we will find still relatively simple tasks such as evaluating proposals for research projects or research programs, which are earmarked funded (*drittmittelfinanziert*) and therefore already by definition must be evaluated *ex-ante* by potential funding agencies (for example, DFG, BMBF, and foundations). Some of the evaluation initiatives that are processed by the German Science Council are already of a much more complex nature: those include assessments of plans to establish a new university – in practical policy terms this means to recommend or not to recommend public funding (*e.g.,* Wissenschaftsrat, 1995a) – and attempts to assess the research performance of certain disciplines (»Forschungsfelder«) across all of Germany; the first two such initiatives focused on environmental research (*Umweltforschung*), completed in 1994 (Wissenschaftsrat, 1994a, 1994b), and on materials' research (*Materialwissenschaft*). There are discussions at the Science Council to launch similar discipline-assessing initiatives in the future, which could address health research, biotechnology, or information technology.[100]

Therefore, as a bottom-line conclusion, we believe that Germany's »evaluation problem« does not result because of a lack of single and individual evaluation initiatives at the micro level. *The main problem is the missing comprehensive picture for the national (or macro-»systemic«) level,* that means for the national German R&D system and, furthermore, for the national research output of Germany's university system. Currently the evaluation policy expertise of university research is very much fragmented and locally bound to individual cases, while the framework or masterplan for a national and large-scale systemic approach is still missing or, to be more cautious, has not been implemented. Germany's decision makers and Germany's academic communities have not yet agreed on how such a

[100] Those general assessments of disciplines obviously cover both, the university and university-related sectors. In Chapter 4 we will discuss in more detail particularly those initiatives of the German Science Council that aim primarily at university-related research (*außeruniversitäre Forschung*).

masterplan should be designed, although there appears to be an evolving consensus, at least partially, on some of the fundamentals for such a major program.

So in that respect a country, for instance, like the United Kingdom seems to have developed further, since the UK can refer to a long tradition of systematically evaluating university research at the national or macro-systemic level. In 1986 and 1989, in a first phase, already two exercises were carried out, while in 1992 and 1996 two major Research Assessment Exercises (RAEs) were launched that radically altered the public funding mode for UK university research. In a European context the UK university-research evaluation model currently is the most radical or – using a different connotation – most consequent approach that, from a purely philosophical viewpoint, certainly can claim a certain intellectual fascination. Summarized briefly, those Research Assessment Exercises feature the following characteristics: (1) All UK universities are covered, that means that the RAEs address the national or macro-systemic level. (2) The primary aim of the RAEs is to assess *ex-post* and by peer review the quality, and not the quantity, of the university research output. (3) However, this qualitative assessment is then translated into a quantitative-oriented national ranking of all university departments and, as a final consequence, of the universities themselves.[101] Those department rankings are also published. (4) The public basic funding (so-called GUF) of university research is, based on a transparent funding-formula, primarily determined by the outcome of those RAEs. This means that there is a direct link – or a systemic feedback – between the results of an evaluation and the amount of public funding, in reference to university research (UFC, 1992c; HEFCs, 1993a, 1994a, 1995a, 1995b; HEFCE, 1996a; Felderer and Campbell, 1994a, 122–126).

To which extent those UK Research Assessment Exercises are applicable to Germany or other Continental European countries, would have to be investigated in a separate analysis. In principle, of course, applied evaluation procedures always must exhibit some compatibility with the prevailing academic culture, otherwise they run the risk of being rejected by the scientific communities. On the other hand, we are personally convinced that some of the basic principles that underpin the British Research Assessment Exercises – for example, evaluating university research comprehensively at the national level and the creation of linkages between evaluation outcome and funding intensity – are clearly future-oriented and therefore should be taken seriously in the context of decision-making processes that affect the universities. Although, particularly in the German case, evaluations in general are facing two practical problems or »cultural« constraints. In parts of the academic communities and other policy-making institutions there is a substantial resistance against an extensive use of quantitative indicators and a wide-spread – although not unanimous – dislike of rankings, no matter if those rankings address universities or

[101] So one interpretation of such an assessment procedure could be to say that *quality is transformed into quantity.*

departments (»institutional« rankings), disciplines (»functional« rankings), or individual researchers. In other national innovation systems, where a different academic research culture and tradition evolved that is more oriented towards competition – as seems to be the case in the UK –, evaluation approaches are more commonly accepted that refer to quantitative indicators and that employ rankings.[102]

Now, again wrapping up our comments on the status – or state-of-the-art – of university-research evaluation in Germany, the current situation may be characterized by the following two main problems:

- Despite a whole series of individual evaluation initiatives of university research at the micro or »bottom« level, we must state that the *generic picture* or the *comprehensive overview* is still missing. So to speak – and willing to employ a provocative phrase – Germany's university evaluation expertise and knowledge appears »fragmented«, and up until now no attempt was undertaken to evaluate university research empirically at the national level for the whole university system (*e.g.,* in an *ex-post* fashion).

- The other major challenge seems to be, what should be the function of future evaluation initiatives for Germany's university research – and the universities in general. Obviously there should be some agreement on the consequences of evaluations, which must be settled in advance. The more systemic question to ask would be: *How should evaluations be implemented into the university context and what should be their feedback link to other functional key elements of the university system?* One of the most sensitive political questions, to which an adequate answer must be found in the next years, is if there should be a systematic and formal link between the outcome of an evaluation and the degree of public basic funding that is allocated to such an evaluated university unit.[103]

Before analyzing in more detail those dynamic changes that, interpreted as a trend and as a possible future scenario, will upgrade the importance of systematic evaluations and, at the same time, will also put the German university system – and the academic scientific community in general – under pressure to give up its resistance (partial resistance) against evaluations, we want to summarize those factors that until now constrained a comprehensive application of evaluations on academic research or, to be more precise, on university research. One intention of this summary is also that – seen in a comparative context – it simultaneously highlights some of the principle problems several Continental

---

[102] We already mentioned the ranking of British university departments according to the 1992 Research Assessment Exercise (see again UFC, 1992c). Beyond that also British university guides, that are designed to help individuals in their decision on selecting the »appropriate« program at the »right« university, employ the concept of a comparative ranking of universities (see, for example, O'Leary and Cannon, 1995, 13–21).

[103] Other countries, for instance the UK, have already decided to establish a tight feedback linkage between the two factors *evaluation and funding.*

European countries would face when at the decision-making level there is an agreement between the involved actors (or participants), to use systematic evaluations as a tool and as a strategy for improving the quality and efficiency of university research. Particularly for European countries with a similar or comparable academic tradition to that of Germany, an in-depth analysis of the German case promises the potential of some learning effects. Turning this argument around, this implies that the constraints against evaluating university research in Germany are not just typically German, but seem to represent a more widespread phenomenon – partially a »cultural« phenomenon – across Continental Europe. Now pinning down our summary, we want to cluster those constraining factors into two groups; *the structural and the cultural constraints.*

3.2.2 The Structural Constraints against Evaluations of University Research

(1) Partially as a reaction against and a historical lesson from the tragic experience of the totalitarian National Socialist dictatorship during the Third Reich period (1933–1945), German universities are granted with the privilege of a far-going autonomous status – concerning, for instance, the content of their teaching and research – which is also legally protected by the German constitution (*Grundgesetz*). This means that the constitution wants to safeguard the freedom of university-internal scientific academic life and furthermore aims at preventing university research from being biased as a result of government interference. In Article 5 (»Art. 5 III GG«) of the *Grundgesetz,* that belongs to the block of the basic rights (*Grundrechte*) which are defined at the beginning of the German constitution, this freedom of science (»Wissenschaftsfreiheit«) is explicitly implemented.[104] In practical terms – *i.e.,* the tradition of legal interpretation – this implies mainly the freedom of the professors (»Lehrstuhlinhaber«) (Mohler, 1995, 7–8). To give an example, Professor Friedhelm Hufen prepared a legal expertise for the German University Association (*Deutscher Hochschulverband*) in which he assessed the legal basis of evaluations of teaching at German universities. Hufen arrives at the conclusion that – with the exception of evaluations carried out »privately« by students[105] – systematic evaluations of teaching of university professors, done by the universities or the government (»Staat«), are unconstitutional (*verfassungswidrig*). Obviously, also sanctions or the re-allocation of resources depending on the outcome of such evaluations, are unconstitutional (Hufen, 1995, 44–48). As a consequence of such (and other) statements, the following situation is created: any systematic evaluation initiatives of university research (or university teaching) will automatically face the *legal constraint* that the opponents of evaluations will consider such assessments to be »unconstitutional« or even »anti-

[104] The exact wording of the crucial paragraph in German is: »Kunst und Wissenschaft, Forschung und Lehre sind frei. Die Freiheit der Lehre entbindet nicht von der Treue zur Verfassung.«

[105] However, the use of the data and results of such student-organized evaluations by other institutions, for example the government, is again – according to Hufen – unconstitutional (Hufen, 1995, 46–47).

constitutional«, by asserting that evaluations violate the constitutionally guaranteed basic right of freedom of research (and the freedom of teaching). Again speaking about university teaching, currently any systematic evaluation of teaching at a German university faces potentially legal risks (see Daniel *et al.,* 1995, 83–84).

(2) Among German academic scholars – and also among German and international experts – there is the wide-spread impression that for a long period of time the funding system of university research was generous, particularly when compared with other Western European countries. This generosity manifested itself twofold: first of all in the intensity and diversity of the German funding system. Germany developed a broad spectrum of different and very distinct institutions that financially foster university research – this institutional spectrum of R&D-funding agencies covers *Länder* and federal ministries, public intermediary agencies (most importantly DFG), companies of the business enterprise sector, and foundations; particularly those foundations occupy a strategic niche for the German research system by having developed a funding policy profile that complementarily supports the public institutions.[106] Secondly, the German research funding system was – and still is, comparatively speaking – characterized by relatively high approval rates for project (or program) proposals. To illustrate that statement, we can take the DFG as an example, the most important intermediary public agency for earmarked research funding (»Drittmittelfinanzierung«) at universities. DFG's most important funding program is the so-called Individual Grants Program (*Normalverfahren*) that is open for bottom-up research proposals, covering the whole spectrum of disciplines. In 1994, 41% of DFG's total funding – 1819.2 million DM in absolute figures – was allocated through this Individual Grants Program initiative. And the approval rate, calculated as a percentage value of the total money sum aggregated from all project applications, was no less than 46.9% (DFG, 1995, 20, 179).[107] In other words, this means that every second research proposal that is forwarded to DFG will also be approved. This represents within the European context – national as well as supranational at EU level – a high value. Analyzed in reference to the outcome, this system of a generous university-research funding produced different effects. On the one hand, it reinforced a »domestic bias« of German academic scholars also in their fund-accessing behavior with the consequence that EC (or EU) research money was not targeted that systematically (see also Felderer and Campbell, 1994a, 52–53).[108] On the other hand,

---

[106] For further information on the key role of foundations in the German context, see, for instance, Felderer and Campbell (1994a, 65–69).

[107] For a quick survey of the *Sonderforschungsbereiche* (Collaborative Research Centers), after the *Normalverfahren* the second major DFG funding initiative, see Streiter (1992).

[108] In our empirical journal-based bibliometric analysis of Germany's academic publication behavior, in Chapter 2.2, we also criticized that the German academic discourse is strongly domestic oriented and does not express enough receptiveness for the international discussions in the sciences: »One of Germany's main problems seems to be that Germany's academic research (and perhaps also national R&D) is biased towards the ›domestic pole‹ and is not enough internationally or outwardly oriented« (Chapter 2.2).

and this is perhaps more crucial for our current argument (but also somewhat of a paradox), the generous »old« funding system also acted as a *financial constraint* against evaluations: since enough research money was available – for example, basic transfer funding, called GUF –, there was not such a need to make a rigorous *ex-post* evaluation of how the research money was used by the scientific university communities. So the message could be that affluency deters evaluations.

(3) Under normal conditions the allocation of a professorship, in Germany, is un-limited or tenured. Therefore, if an individual succeeds in being appointed as a professor at a German university, he or she receives a tenured employment status (similar to that of a »unbefristetes Beamtendienstverhältnis«): this means that the employment contract is permanent and un-limited, and that he or she cannot be dismissed, except for a criminal offense; furthermore he or she receives all the benefits of public employment, however, with the main advantage and privilege of an extremely far-reaching and legally safeguarded autonomy – also protected by the German constitution – against attempts of interference from outside (the government, for example) as well as from the university hierarchy. In combination with the prior and »old« system of a generous funding of university research – including the intensity of basic funding (GUF) – this so-called *unbefristete Beamtendienstverhältnis* acted as an *institutional employment constraint* against performing *ex-post* evaluations: first of all, the public decision makers had almost no direct means at hand to force the German professors to be evaluated; and, secondly, the incentives for professors to be evaluated voluntarily were also weak – the allocation of benefits or resources was not sufficiently based on evaluations and their outcome.[109]

(4) One argument often heard and commonly referred to is that of the *constraint of size* in the case of Germany's higher education sector R&D. The simple message would be: *Germany is simply too large* for applying evaluations comprehensively to university research. And the quantitative numbers seem to support such a hypothesis. In 1994, the German higher education sector consisted all together of 325 institutions, of which 88 qualified as a »classical« university. In total, those higher education institutions counted 1.7 million students (see again our survey at the beginning of Chapter 3.1). Focusing more specifically on the higher education sector R&D potential, Germany had in 1993 no less than 67140 researchers – academic researchers with a university degree – calculated in full-time equivalents; the corresponding figure for France, in the same year, was 49862, for the UK 32000, and for Italy 33204 (OECD, 1996a, 39). Therefore, some experts argue that Germany is

[109] For instance one expert expressed the opinion that, as an estimation, only one third of the German professors forward project proposals to DFG. If this is really the case, then this would imply – as a simplified message – that only one third of the university research is systematically (*ex-ante*) pre-evaluated. Of course we must admit that we are currently not in the possession of an exact and professional survey on that topic – this obviously would be necessary for a further scientific-based discussion.

too large in size for a comprehensive application of evaluations, based on quantitative indicators, that would cover the whole national territory. In that respect, smaller European countries, such as Switzerland, the Netherlands, Scandinavia, and Finland, are in a more comfortable position since they share the same attribute of being »small-sized«; and because of that it is easier for decision makers to establish an evaluation-based overview on the quality and efficiency of university research.[110] Furthermore, a similar argument in addition emphasizes that in Germany most (or at least many) universities developed a comparable level of research quality. This would imply that in Germany there is not much deviation from an average median standard of university research – using a different statistical term for a metaphoric description, the variance of German university research appears to be low. So while – as those experts would stress – it seems to be quite clear in many European and non-European countries (for instance, the UK and the United States) which are the leading research universities, the German situation is opposite and can be characterized in the following way: in Germany we are confronted with a major pool of universities that perform at a similar level of research quality and research efficiency. This complicates university research evaluations since, to give an example, a ranking of universities might be somehow »subjective« or at least not reliable. Speaking in more scientific terms, critics could assert that depending on the evaluation methods – and keeping in mind the similar national performance quality of German universities – very different evaluation outcomes would be produced as a consequence of those circumstances. So evaluations, already in principal, are a critical tool for Germany and not appropriate for decision-making and future-oriented planning.

(5) Some experts emphasize that even more important than the sheer size is the fact that the political-geographical structure of Germany is based upon federal principles: this implies speaking of the *constraint of federal and decentralized decision-making* for evaluations. Any analysis of Germany's federal structural design quite clearly reveals its complexity as a consequence of the interactions of the many political actors at the federal, *Länder* and sub-*Länder* level. Using a simplified terminology, this diversity of actors guarantees, on the one hand, a pluralism of problem awareness and problem-solving strategies; on the other hand, however, sometimes it also leads to a certain paralysis of action – for instance concerning policy that should be applied to the universities.[111] The consequences of this actor's involvement complexity for

---

[110] The following figures demonstrate the number of academic researchers in full-time equivalents in the higher education sector of some small-sized European countries: Denmark 4627 (1993); Finland 6097 (1993); Netherlands 10630 (1993); Norway 4737 (1993); Sweden 11738 (1993); and Switzerland 7800 (1994) (see again OECD, 1996a, 39).

[111] In that context German experts like to use the term »Politikverflechtungsfalle«, which could be circumscribed in English with *mutual paralysis of multiple actors*. The meaning of »Politikverflechtungsfalle« is that because of those interactions of multiple actors the implementation of a policy will be constrained; simply speaking this implies that the actors are *trapped* (in German »Falle«). It is said that the term »Politikverflechtungsfalle« was invented by the German scholar Fritz W. Scharpf (for an application of that

university evaluation policy are manifold: firstly, it would not be easy for a central unit at the federal level to apply an evaluation policy that is not approved by the *Länder*.[112] Secondly, the German *Länder* are mutually dependent. This implies that the individual scope for a *Länder* policy is limited and must be sensitive for what is happening in the other *Länder*. So, as some experts would say, one *Land* cannot apply a university evaluation policy that is independent or ignorant of developments in that policy field in other *Länder*. Thirdly, a funding of universities based on the outcome of evaluations might create the fear among political actors at *Länder* level that such a rigorous application of elements of competition would favor those *Länder* with locations of the »good quality« universities and would harm the other *Länder*. Since, however, all *Länder* are mutually dependent and linked to each other across several issue dimensions, this finally could lead to the collapse of a general policy consensus among the *Länder:* those *Länder,* with the »poor quality« universities, could take revenge in other policy fields.

3.2.3 The Cultural Constraints against Evaluations of University Research[113]

(1) In Germany the classical paradigm or leitmotif that shaped the evolution and development of structures of the university system is the so-called *Humboldtian* principle of the *Unity of Teaching and Research* (»Einheit von Lehre und Forschung«).[114] This implies that teaching and research at a university are two academic activities that functionally overlap in their content and thus cannot be separated. In other words: they are mutually linked to each other. And as a normative statement this, thought a step further, leads to the consequence that each professor and the other academic staff at a university *should* in fact do both, that means teaching and research. This may also partially explain the emergence of the public funding base of European universities with the peculiarity of those general public transfer funds (also called public basic funding or general GUF) which are not *ex-ante* earmarked or pre-defined for a certain purpose. This basic funding is thought to support teaching as well as research (and also, of course, administration). Only *ex-post* and after applying an empirical survey is it possible to say for which purposes

---

concept in the context of an empirical analysis, see, for example, Hohn and Schimank, 1990, and Schimank, 1995).

[112] At this point we must keep in mind that in Germany the major portion of GUF funding of universities is not carried out by the national government, but by the *Länder* themselves (see BMBF, 1996b, 23, 88–90; Irvine *et al.,* 1991, 50; Felderer and Campbell, 1994a, 46–50).

[113] We personally believe that those *cultural* constraints are just as important as the *structural* constraints, that we discussed above, since the structures of a society are significantly shaped by – if not even the product of – underlying cultural fundamentals.

[114] In that respect Josef Lange speaks of the »auf *Humboldt* zurückgeführten Prinzip der Einheit von Lehre und Forschung für die deutsche Universität« (Lange, 1995a, 71) and the »Humboldt'schen Prinzip der Einheit von Forschung und Lehre in der Gemeinschaft von Lehrenden und Lernenden« (Lange, 1995b, 9).

which portions of those basic funds were used, that means, for instance, how much money or how many personnel working-hours were invested into university research and, more specifically, into which types of research activity. However, this financial formula of a common *unity funding of teaching and research*[115], derived from the Humboldtian principle of a unity of teaching and research, is already *per se* in a potential conflict with the concept of evaluations. The crucial problem of those public basic transfer funds seems to be that their actual use by the university communities is not optimally transparent.[116] A successful application of evaluations, on the other hand, depends highly on the degree of transparency that is offered by a system. So, from the perspective of successful evaluations, it would make sense to redesign, at least partially, the public funding base of universities; this means to decrease the transfer funds and to boost up earmarked funding for research and perhaps also for teaching (the financial key term in German is *Drittmittelfinanzierung*).[117] The critics assert that this would endanger the functionally necessary unity of teaching and research.[118] Therefore, the correct funding formula for universities is a controversially debated hot issue in the German context and no general consensus is in sight currently.

(2) Some German experts and scholars are inclined to characterize Germany as a consensus society, permanently attempting to find a compromise between divergent and potentially conflicting positions. And, at least according to that hypothesis, this cultural pattern is also reproduced within the university context. This implied two consequences for universities and for university research in more particular: first of all, competition – either between individual researchers (or professors), university departments, or universities in general – was never rigorously applied; in other words, competition was not seen as a principle for advancing and encouraging research (and teaching) quality, but as something that potentially could disturb those consensus-oriented mechanisms that were permanently seeking a compromise within the university system. To give an example, the German Science Council issued in 1985 a report that investigated the possibility of more competition in the German university system (Wissenschaftsrat, 1985). Although the Science Council, in principle,

---

[115] In German one could paraphrase this as: *Die Einheit der Finanzierung von Lehre und Forschung*.

[116] In Chapter 3.1 we already discussed in great detail the funding formula of German universities, particularly the two funding categories of basic funds (*Grundmittelfinanzierung*) and earmarked funds (*Drittmittelfinanzierung*).

[117]) For a further discussion on the advantages and disadvantages of *basic versus earmarked funding* of university research in two European countries, the UK and Austria, see Felderer and Campbell (1994a, 113–114, 210–211).

[118] Uwe Schimank, for instance, claims that – staying in his terminology – the »application-remote basic research« (*anwendungsferne Grundlagenforschung*) can only survive at the universities when it is interpreted as a by-product of teaching; that means teaching offers something like a »piggy-back legitimation« (»Huckepack-Legitimation«) for certain areas of basic research (Schimank, 1995, 334–335). The critical question to ask, of course, would be whether something such as an *anwendungsferne Grundlagenforschung* really exists, based on scientific premises?

recommended (or recommends) competition (*Wettbewerb*) – particularly when this competition leads to an increase of diversification between the universities[119] –, the Science Council, at the same time, also underscored potentially negative effects of an intensified competition. For an »outside observer« this appears to be a concession to the prevailing academic culture in Germany – which obviously is necessary when recommendations by institutions should be taken seriously by the academic communities.[120] The second consequence of this orientation towards consensus was – in combination with the Humboldtian principle of unity between teaching and research – the »cultural demand« to create a *medium standard quality* in research (and teaching) across the national German university system. So an emphasis was put on supporting the »weak« universities and not so much on encouraging the »stronger« universities, since a broad spectrum of quality diversification would have deteriorated the *medium quality* claim. Speaking metaphorically in statistical terms, there was an interest in keeping the variance or variability of quality – in reference to that medium standard quality – low. Therefore, competition was qualified as a principle, probably correctly, that would have questioned and delegitimated this medium-standard approach twofold: firstly normatively and secondly, understood as a process, competition almost automatically will lead over time to an increase of diversification and by this also to an extended variance of quality. So when some German scholars claim that evaluations of university research are difficult to realize because there is not enough difference between the quality performance of German universities[121], then this – obviously only postulated – quality similarity of university

---

[119] »Wenn der Wettbewerb im Hochschulbereich sich nicht darauf beschränkt, daß alle das gleiche – nur jeweils besser als der andere – zu tun versuchen, sondern dazu führt, daß alle Beteiligten ihre besonderen Stärken herauszufinden und zu entwickeln sich bemühen, wenn also Wettbewerb ein hohes Maß an Differenzierung hervorbringt, können viele Gewinn davon haben. In diesem Sinne befürwortet der Wissenschaftsrat Wettbewerb im Hochschulbereich« (Wissenschaftsrat, 1985, 9).

[120] To illustrate the Science Council's caution, we want to quote four passages out of this report (Wissenschaftsrat, 1985): »In der hochschulpolitischen Diskussion der letzten Jahre ist ›Wettbewerb‹ ein Schlüsselwort geworden. Von verschiedenen Seiten wird gefordert, Wettbewerbselemente im Hochschulsystem der Bundesrepublik Deutschland erheblich zu verstärken. Das nordamerikanische Hochschulsystem gilt dabei häufig als ein Muster, dem nachzueifern sei. Dagegen haben sich allerdings auch kritische Stimmen erhoben. Sie bezweifeln, daß dieses Modell für die deutschen Verhältnisse geeignet sei und widersprechen darüber hinaus auch seiner Idealisierung« (page 5). »Wettbewerb ist kein Selbstzweck. Auch im Hochschulsystem ist es nicht schon ein Wert an sich, daß seine Mitglieder – Institutionen und Personen – miteinander konkurrieren. Wettbewerb ist prinzipiell sinnvoll, wenn und insoweit er die Fähigkeit eines Systems verstärkt, wünschenswerte Leistungen zu vollbringen« (page 7). »Damit ist bereits angedeutet, daß das Maximum an Wettbewerbsintensität nicht das Optimum ist, gemessen an den Zielen, denen Wettbewerb dienen soll. Es gibt eine Zuspitzung der Konkurrenz, die eher lähmt und entmutigt als stimuliert, weil vernünftige, kalkulierbare Erfolgschancen nicht mehr gegeben sind. Wettbewerb kann zu einer Konzentration von Qualität führen, die mit einer geistigen Provinzialisierung außerhalb der Zentren bezahlt wird« (page 8). »Diejenigen, die bei einem solchen Verteilungssystem zurückfallen, dürfen jedoch keineswegs ihrer Arbeitsmöglichkeiten und Wettbewerbschancen beraubt werden. Dies ist schon deshalb auszuschließen, weil im Sinne deutscher Hochschultraditionen und im Einklang mit der Verfassung der Bundesrepublik Deutschland die Länder gehalten sind (und gehalten bleiben sollen), Mindeststandards für gute Forschung und Lehre einheitlich im ganzen Hochschulsystem zu gewährleisten« (page 9).

[121] We already mentioned that argument when discussing the *structural constraints* against evaluations of university research in Germany.

performance could be interpreted as the final »end-product« of a normative cultural demand. Now when this consensus orientation (and avoidance of competition) *per se* is questioned, then there is one commonly referred to argument: the economic and societal success story of West Germany after 1945 – a viable economy and a stable development of democracy – seems to verify, in a positive sense, the underlying cultural premises of German society.

(3) Detlef Müller-Böling presents one of the most fascinating analyses of German academic culture, by distinguishing between the two terms of an *ex-ante* and an *ex-post* quality control (in German »ex-ante-Steuerung« and »ex-post-Steuerung«) (Müller-Böling, 1995, 31–34). Müller-Böling argues that the Continental European and particularly the German university was guided by the idealized conceptual belief that an *ex-ante quality control* for universities and university research is possible[122]: speaking in simplified terminology, this concept implies that by implementing a rigorous quality control at the »beginning« of a process (for example, either appointing a professor or establishing a new university), this almost automatically guarantees that a high-level quality will be continuously performed – consequently, within the logic of such a conceptual framework, there is no need for systematically evaluating the quality of a process, either in parallel or *ex-post*. Therefore, policy makers – who accepted this conceptual approach – preferred to invest their activity in developing a system or regulatory framework of quality checks and quality thresholds that already in advance, this means *ex-ante,* should have the capability to promise a high-quality output of university performance. In practical policy terms this leads, on the one hand, to a close monitoring of the government over the operation of the *de facto* public universities[123] – here we can mention the missing tradition of »private« universities in Germany[124] – and, secondly, to a complex framework of public regulations (ranging from curricula issues to questions of organization), which had to be met by the universities. At the level of financing, however, this also explains the strong dominance of GUF funds: GUF implies that the universities receive substantial public transfer funds that are used by the university communities themselves and autonomously either for teaching or research (or administration); and because of the implementation of a strict public *ex-ante* quality control regulation – this means that universities and university personnel had to pass early-stage quality checks – the government and other public decision makers are convinced in advance (now speaking in »official«

---

[122] »Kennzeichnend für die Universität in der europäischen Tradition ist ein System der ex-ante-Steuerung seitens des Staates« (Müller-Böling, 1995, 31).

[123] »In der Hochschule als staatlicher Einrichtung bzw. unter staatlicher Anerkennung wacht der Staat über seine eigenen Einrichtungen oder vergibt die Anerkennung als Hochschule. Damit ist die Einrichtung einer privaten Institution, die sich Hochlschule nennen darf, verwehrt« (Müller-Böling, 1995, 31).

[124] The Universität Witten/Herdecke, Germany's first private university, was founded at the beginning of the 1980s, about 1980 (Universität Witten/Herdecke, 1994, 7, 13). In Chapter 1 we already mentioned the University of Witten/Herdecke for the first time.

terminology) that those public basic transfer funds are used properly.[125] Coming down to the individual micro-level this general belief in the possibility of an *ex-ante* quality control probably helps us to understand the institution of a German professorship, with the crucial characteristics that professorships are regularly granted with tenure (temporally unlimited) including the additional key privilege that a professorship is non-dismissable – in that respect professors occupy a status comparable to that of public employees (»unbefristetes Beamtendienstverhältnis«). The underlying »cultural« rationale for that seems to be that once an individual has passed certain quality checks and quality thresholds – most importantly the *Habilitation* procedure – then this should be acknowledged as a guarantee for a life-long and high-level quality performance that does not demand the efforts of systematic *ex-post* evaluations anymore. So he or she can be granted the privilege of a tenured professorship without *ex-post* evaluations, either institutional or procedural.

(4) There are some German academic scholars who are willing to raise the principal question whether *the quality or efficiency of university research can really be measured?* Going a step further, this implies challenging the possibility of a meaningful representation of research quality and research efficiency that relies on quantitative indicators. Already in Chapter 2.1, while investigating the theoretical basis of bibliometric analyses, we summarized those standard statements (or hypotheses) that often are brought into discussion as arguments against bibliometric comparisons that refer to journals which are covered by the SCI and SSCI data bases. And one of those arguments is an almost fundamental skepticism that surfaces when the appropriateness of a quantitative-oriented measuring of the quality and efficiency of research is at stake.[126] So while most scholars would not doubt the »measurability« of application-oriented research (for example, the number of patents), the picture turns dramatically in the case of basic research: some scholars express the opinion that it is impossible to »measure« basic research, simply because basic research represents something like a frontier where the establishment of a general scientific consensus among researchers is still not a consolidated and final body of knowledge but a process in flux; and without such a consensus also an »objective« measurement procedure of research results cannot be carried out. This is also the context which makes the argument of Uwe Schimank clearer, who asserts that the *anwendungsferne Grundlagenforschung* (»application-remote basic research«) can only survive at universities when it is understood as a by-product of teaching – the so-called »piggy-back legitimation« of teaching for basic research (Schimank, 1995, 334–336). Therefore, summarized with other words, we can say that within Germany's

[125] For example, in 1991 no less than 71.4% of Germany's higher education sector R&D was financed by such public basic-transfer funds, called GUF (see again Figures 17 and 18).

[126] In Chapter 2.1 we paraphrased such a fundamental rejection of the possibility of measuring research output meaningfully as »anti-bibliometrics«.

academic university community there seems to be a partial dislike against systematic quantification attempts of research output: the asserted claim would be that *quantification of quality* is inaccurate, meaningless, or »shallow«. So it should not be surprising that at the policy level the idea of ranking universities relative to a quantitative-oriented measuring of research represents a controversially debated issue and cannot claim a high popularity. Wolfgang Frühwald, president of DFG since 1992, stresses that in the case of the United States a ranking of universities makes sense, since in the US the university system is based on private competition; so a high-ranking position of a university guarantees and justifies a stable influx of tuition-paying students. In Germany, however, a simple application of such a university-ranking index – without recognizing and incorporating some of the traditions of German academic culture – does not appear appropriate.[127] Similarly Hans-Uwe Erichsen argues by emphasizing that there is a need for more diversification and competition between and within universities; at the same time, however, Erichsen rejects the applicability of the concept of »elite universities« (*Elite-Universitäten*) for the German higher education sector.[128]

This analytical survey of the **structural and cultural constraints** impressively demonstrates why evaluations and evaluation policy were not that systematically applied in Germany on university research (and university teaching) as was the case in other countries, for instance the United Kingdom. In that context one must remember that a national system of innovation or a national university system is always embedded in a complex domestic (and international) environment, so the structures and the cultural patterns of a society define a framework that either encourages or constrains the application of evaluations. Perhaps Germany as a society had generated such favorable conditions over a long period of time – for instance, a prosperous and rapidly expanding economy with a surplus creation of wealth – that it could afford the »luxury« of a university system that relied on some *ex-ante* control principles but without rigorous *ex-post* quality checks. This could imply the interpretation of Germany as an affluent society, and then the investigation of what the policy consequences in certain areas are (or were). On the other hand, we are also convinced that when analyzing comparatively the underlying patterns that shaped the traditions of German academic culture, then this would reveal that Germany's academic system can be characterized with some attributes that are also shared by other »neighboring« European countries. This would encourage us, in our function as »outside«

---

[127] »Das System des ›Ranking‹, also ein wirtschaftliches Wettbewerbsinstrument des freien Universitätsmarktes, in das staatlich (mehr oder weniger) garantierte, für die Studenten kostenlose und auf Verwischung aller Unterschiede angelegte Universitätssystem Deutschlands zu übertragen, ohne zugleich wesentliche Elemente des Herkunftssystems zu übernehmen, ist nichts als Augenauswischerei und führt zu den verwirrenden und letztlich folgenlosen Ergebnissen der deutschen Universitätsumfragen« (Frühwald, 1995, 211).

[128] »Wir brauchen keine Elite-Universitäten... Was wir brauchen, ist Elite und Eliteförderung in den Universitäten, ferner mehr profilbildender und leistungssteigender Wettbewerb unter und in den Hochschulen« (Erichsen, 1993, 142).

observers, to understand or to interpret developments in the German university sector as manifestations of a larger trend that takes place across Continental Europe. Since also Germany is a part of that Continental European legacy of a *particular type of academic culture and a distinct tradition of intellectual and conceptual reasoning,* problems and conflicts arising in Germany are also of relevance – at least their analysis – for other Continental European countries and their academic research systems.

Now despite the existence of those structural and cultural constraints that, up until now, blocked and successfully prevented a rigorous or pervasive application of evaluations on university research and university teaching[129], the situation or *status quo* in Germany seems to be changing. Among many leading experts – also from university communities – and also among decision makers there is a growing consensus that *evaluations essentially will gain in importance for the German university system* in the next years. This consensus may be split down into two different messages: first of all, evaluations are such a valid tool for improving the performance of a system, that a neglecting of evaluations no longer can be »afforded« or justified. Secondly, the universities themselves will come under significant internal and external pressures to incorporate principles of evaluations more determinedly. For the purpose of supporting such a conclusion, we again can cite Hans-Uwe Erichsen, who has been president of the HRK (German Rectors' Conference) since 1990 and, as a consequence of his professional function, is clearly a key actor who »plays« in favor of the German universities. Despite his skepticism against the previously mentioned concept of elite universities (Erichsen, 1993) and a ranking of German universities[130], Erichsen clearly stresses two messages: firstly, evaluations of the performance of German universities are necessary and essential, and secondly, this demands a pervasive and deep-going transparency of university activities.[131] In the following we want to investigate and analyze thoroughly those developments that will demand that also in Germany – at least this is our hypothesis – the concept of evaluating universities (and university research) will gain a crucial and strategic importance in the future. Our analysis will focus primarily on two key issues: firstly, we will present those arguments that emphasize that *research output and, consequently, research quality and research efficiency are measurable*. Secondly, we want to demonstrate why there is a *systemic need for feedback:* in practical policy terms this

---

[129] This statement appears particularly valid, when an international frame of reference is taken as a methodological basis: this means comparing Germany with other countries.

[130] »Die Praxis von Presseorganen, auf der Basis der Gewichtung und Verknüpfung einzelner Indikatoren Hochschulranglisten aufzustellen, entbehrt daher einer sauberen methodischen Grundlage und dient in erster Linie der Steigerung der Verkaufszahlen« (Erichsen, 1995b, 217).

[131] »Die Absage an Ranglisten ist daher keine Absage an Leistungsdarstellung, -messung und Transparenz im Hochschulbereich. Vielmehr dürfte die Grundsatzdebatte über das Für und Wider von Leistungsevaluation im Hochschulbereich mit einem prinzipiell positiven Ergebnis abgeschlossen sein«; and: »Notwendig ist in der Tat zunächst die Herstellung von Transparenz im Hochschulbereich. Ein weiteres Erfordernis ist es, ein quantitativ ausgerichtetes Raster zu entwickeln, mit dessen Hilfe Leistungsdaten erfaßt werden können« (Erichsen, 1995b, 218).

implies implementing transparency-producing evaluation mechanisms into the German university system.

3.2.4 The General Reasons why Evaluations of Universities and of University Research will become more Important in the Future

(1) **The measurability of research output, research quality, and research efficiency:** One fundamental and evaluation-skeptical argument sometimes raised is the assertion that research or at least basic research cannot be measured. Phrasing it somewhat differently, the objection would be that a meaningful quantification of research quality must be regarded as impossible. This argument is put forward to falsify the legitimation of evaluations of academic research in general and *per se* (already discussed in Chapter 3.2), and consequently it is also used as an argument against the methodology of specific evaluation approaches – for instance a bibliometrics-based comparison of article output in international journals (see again our analysis in Chapter 2.1 on the theoretical premises of a bibliometrics-based methodological approach). Since this notion, that research output cannot be measured, is still strongly supported by some members of Germany's academic university community, we are inclined to set up the hypothesis that the issue of *non-measurability versus measurability* of academic research seems to represent something like a fundamental cleavage that splits German academia into two groups. To give an example for that currently highly controversial debate in Germany, Hans-Dieter Daniel, who arrives at the conclusion that scientific performance can be measured properly, decided to title the corresponding article, where he published that conclusion, with the question *Ist wissenschaftliche Leistung in Forschung und Lehre meßbar?*[132] (see Daniel, 1995, 205). So to propose in the context of German discourse that research, indeed, can be measured and that research quality and research efficiency can also – under certain circumstances – be quantified, is not a commonly accepted notion in German academia, but represents a hypothesis which still must be justified and defended against the critics. In that conflict cleavage in German academia we personally would agree and align with those experts and decision makers who emphasize the measurability of research and, more particular, also of university research. Furthermore, it also appears that juxtaposed as a trend that they also will be the »winning side« over those who, almost fundamentally minded, deny any meaningful quantification possibility of research.[133] Going analytically a step further, there are, of course, many *powerful arguments* that support the hypothesis that research output, research quality, and research efficiency can be

---

[132] A possible translation of that German title into English would be: *can scientific performance in research and teaching be measured?*

[133] We already cited Hans-Uwe Erichsen, president of the HRK and a »pro-university« key actor in Germany, who expresses the conclusion that the possibility of evaluating performance in a university context has already been accepted, at least in principle (see again Erichsen, 1995b, 218).

measured adequately, also in a university environment – at least in principle. In the following we want to summarize some of the arguments which we regard as crucial – also for advancing a more precise argument as a discourse:

– There are many different definitions for the common underlying purpose or aim of the broad spectrum of the whole sciences (including the disciplinary branches of the social sciences and humanities). As the devil's advocate, of course, one could even raise the provoking question whether such a common conceptual basis really exists for all of the sciences? However, for the practical purpose of continuing our analysis we want to set up the assumption or premise that all of the sciences can be characterized by some similar attributes. So for us a key definition of the function or goal of sciences would be: *the interest to create knowledge about the empirical world or the empirical environment.*[134] In that definition a crucial term obviously is the word »empirical«, since anything that is empirical can also, at least in principle, be measured. Such a conceptual framework then implies that the sciences are significantly concerned with creating knowledge about empirical structures and processes, that means of *measurable units*. Therefore, in our opinion it is absolutely clear that when the sciences deal with the empirical world as a part of their research procedure and research mission, *i.e.,* with measurable units, then the research output »itself« consequently can also be measured – at least this would reflect the general demand. Turning this argument around, this would imply that when a research output (for instance, its quality and efficiency) cannot be measured, then such a particular research output is not the product of an empirical or scientific-based approach. Speaking in simplified terminology, only non-scientific or even anti-scientific research output would be »non-measurable«. So to claim that a research output cannot be measured, implies, at least for us, to leave the realm of sciences, and to confuse sciences – which always have a functional and interest-driven empirical research element – with »non-sciences«, perhaps »metaphysics«.[135] So our radical bottom-line conclusion would be that when research output cannot be measured, then such an output is not the result of a scientific approach; with the consequence that we can simply ignore such a research output when we speak of academic, *i.e.* scientific research, and the possibilities (and problems) of evaluating scientific research. This does not mean, of course, that we want to deny the complexity of modern sciences

---

[134] This obviously includes »ourselves«, since also humankind is a part of the empirical world.

[135] As an example for early twentieth-century based criticism of modern philosophy *versus* science-skeptical metaphysics, see the work of the members of the Austrian Vienna Circle (*Wiener Kreis*); for instance, the programmatic manifest *Wissenschaftliche Weltauffassung – Der Wiener Kreis* [Scientific World View – The Vienna Circle], authored by Rudolf Carnap, Hans Hahn, and Otto Neurath. The Vienna Circle was founded in 1907, and flourished in the 1920s under the leadership of Moritz Schlick. Other Circle members were Herbert Feigl, Kurt Gödel, Friedrich Waismann, and in its »periphery« also Karl Popper. The Circle's philosophy was crucially influenced by the early work (*Tractatus Logico-Philosophicus*) of Ludwig Wittgenstein. As a consequence of the rise of Nazism in the 1930s, the Vienna Circle finally was dissolved in 1938 (see Schleichert, 1975; and Honderich, 1995, 702–703, 899, 912–916).

and also to be ignorant of perhaps diverging developmental levels of a – to invent a new term – so-called »empiricality«[136] across or within different disciplines in the sciences.

– Among academics there is this deeply rooted notion of a distinction and split between a more applied research, on the one hand, and basic research as the opposite pole on the »other side«. While most academic scholars are willing to accept that performance on the »application end« of the scientific spectrum can be measured meaningfully, there is still some skepticism whether this is also possible for basic research. We already mentioned Uwe Schimank's concept of an *anwendungsferne Grundlagenforschung* (»application-remote basic research«) and his thesis that such a research activity can primarily only be justified by teaching (the so-called »piggy-back legitimation«).[137] However, we are more inclined to agree with those experts who emphasize that it is misleading to understand basic and applied research as two contraries that have almost nothing in common. So the challenge would be to replace such a duality-thinking with the notion of a *spectrum or continuum:* this implies acknowledging that there is much interference and overlapping between basic and applied research, so that most academic research activity incorporates elements of both.[138] Such a view point is also supported at the »theoretical« level strongly by scientists who emphasize that for research and for the further development of sciences the *context of application* is crucially important. Presented radically simple and expressed in our own words: *without application, no maturing of a theory seems possible*. A group of leading experts paraphrased this new conceptual understanding of the sciences and of scientific research as **Mode 2** and published their conclusions in a book titled *The New Production of Knowledge,* in which the »familiar way« of knowledge production is called simply Mode 1 (Gibbons *et al.,* 1994).[139] This concept

---

[136] With the term »empiricality« one could circumscribe the degree of empirical maturity (and the empirical-oriented methodological sophistication) either of a theory or of a discipline in the sciences. For a translation into German one could use the phrase of a *empirische Gehalt* of a scientific discipline, that can differ from discipline to discipline when crossing the whole disciplinary spectrum.

[137] See again Schimank, 1995, 301–336.

[138] To use a metaphoric phrase, one could say that basic and application-oriented research have more in common than what seperates them.

[139] In the following we want to cite three key sections of that book (Gibbons et al., 1994), in which this notion of the *importance of the context-of-application* is underscored: »The new mode operates within a context of application in that problems are not set within a disciplinary framework« (page vii); »... in Mode 1 problems are set and solved in a context governed by the, largely academic, interests of a specific community. By contrast, Mode 2 knowledge is carried out in a context of application« (page 3); »Mode 1 is discipline-based and carries a distinction between what is fundamental and what is applied; this implies an operational distinction between a theoretical core and other areas of knowledge such as the engineering sciences, where the theoretical insights are translated into applications. By contrast, Mode 2 knowledge production is transdisciplinary. It is characterized by a constant flow back and forth between the fundamental and the applied, between the theoretical and the practical. Typically, discovery occurs in contexts where knowledge is developed for and put to use, while results – which would have been traditionally characterised as applied – fuel further theoretical advances. Discovery in the context of application in the case of hypersonic aircraft

of Mode 2 is also reviewed in German academic discourse (Krull, 1995b, 34–36, 46–47).[140] Also at the policy level an organization such as the OECD arrives at the conclusion that the old distinction between basic and applied research does not reflect anymore properly the empirical reality of processes of innovation. So, according to the OECD, the future challenge is exactly to intensify the linkage between the basic and the applied end of research. Therefore, the OECD stresses to use a terminology of a »strategic« or »application-oriented« research that should underscore the necessity of a permanent interaction between more basic-oriented and more-application oriented elements within the overall context of a research procedure.[141] Making a final and last comment on this notion of basic *versus* applied research, also basic research deals with the empirical world – so it must feature at least some elements of »empiricality« (see our above definition); and beyond that also basic research, even when it is highly theoretical, takes place within the context of a scientific community which has some incorporated rules that decide on the quality of scientific work. Therefore, at least the publication of results of theoretically-oriented basic research – for instance, in articles in internationally renowned journals – are »empirically« countable. Referring back to our previously raised question whether a type of scientific research such as the *anwendungsferne Grundlagenforschung,* as postulated by Uwe Schimank, really exists, we believe that the following play on words (*Wortspiel*) is accurate and helpful in pointing at the crucial conclusion: something like an *anwendungs-»ferne« Grundlagenforschung* (application-»remote« basic research) probably exists, but in the case of an *anwendungs-»freie« Grundlagenforschung* (application-»free« basic research) there clearly are massive doubts whether such a *Grundlagenforschung* (basic research) still can claim to be »scientific« – when we are willing to accept some referentiality to the empirical world (phrased by us as »empiricality«) as a key attribute for describing the function of modern and contemporary sciences.

– When some German academic scholars assert that research output cannot be measured, then – in addition to the arguments we already have stressed – we want to emphasize that such a statement is not internally consistent. In other words: *scholars pushing the non-measurability hypothesis reveal a cultural »blind spot«* (in German: *blinder Fleck*) or, at least, apply a double standard of morals by treating the students, on the one hand, differently than the university research staff, most

---

is illustrated in Box 1.1. Mode 2 is characterized by a shift away from the search for fundamental principles towards modes of enquiry oriented towards contextualised results« (page 19).

[140] For possible impacts of this concept of Mode 2 on the future performance of the social sciences, see also Campbell (1995, 402).

[141] »However, the old terms of basic research, applied research, and development have become increasingly inadequate to describe the innovation process. Terms such as ›strategic‹ or ›application-oriented‹ are being added to stress the idea that fundamental investigation is needed in areas with potential for applications based on new principles or discoveries« (OECD, 1992b, 32).

prominently the professors. Concerning the students, everybody accepts it as a given and »natural« fact that the students' performance should be evaluated on a permanent basis. In addition, when students work for their first-degree thesis or – at the post-first degree level – on their dissertation, then, at least according to the theory, the »university system« demands or expects that those students generate research results of their own. And those completed theses at the Master or Doctoral level then are evaluated and graded – that means the quality of their research output is »measured« – by their supervisors, mainly professors. In the case that an academic researcher wants to go through a *Habilitation* procedure, the above said obviously is also true: firstly, a *Habilitation* work must (or should) reflect an academic in-depth research, and secondly the final *Habilitation* is judged and evaluated by a university commission. *Therefore, when in the case of students and applicants for a Habilitation it is commonly accepted that their academic research work must be consequentially evaluated and permanently graded, that means their research output is measured, so why should this not also apply to the academic university research staff, most importantly the professors?* From the institutional perspective the answer appears simple: when, based on scientific premises, it is legitimate to evaluate the research output of first-degree and second-degree level students, then it cannot be justified not to evaluate professors with the same strictness. So when professors are not evaluated, then this seems to be more the product of a hierarchical distribution of privileges than the self-logic of a scientific rationale.

– When, in principle, *research and research output are measurable* – at least according to our viewpoint that we developed now in detail and that is shared and supported by many leading experts –, this implies that *research quality and research efficiency can also be measured*. Going a step further and referring to issues of policy and policy-making, this clearly demonstrates that a rejection of systematic evaluations of university research cannot be justified by scientific arguments. Speaking positively, comprehensive evaluations of university research are easily to legitimate by those fundamentals that structure and drive the evolution of modern sciences *per se*. Coming down to the question of which methods are appropriate for research evaluations, then we are inclined to emphasize the notion of a *plurality or pluralism of methodological approaches;* as long as those different methods meet the criteria of an accurate and sound scientific analysis. Furthermore, it is plausible to assume that concerning the »general picture« those different methodological approaches should converge, at least by tendency, in their results and conclusions (see again, for example, van Raan, 1995, 93).[142] In Chapters 2.1 and 2.2 we presented one such methodological possibility for how academic research (its patterns, quality, and efficiency) can be measured and analyzed meaningfully. We referred to a

[142] Van Raan emphasizes that, in the case of the Dutch universities, peer-reviews and indicator-based bibliometric analyses generally arrive at the same conclusions.

bibliometrics-based comparison of article output in those international journals that are covered by SCI and SSCI, and developed for Germany a three-stage analysis: in Chapter 2.1 we discussed in detail the theoretical basis for such an approach, and in Chapter 2.2 – after having commented thoroughly on the article-output patterns – we clearly demonstrated how those bibliometric results were impressively compatible with other indicators for Germany (for example, the degree of international economic competitiveness and the technology balance of payments). Of course we know that the methodology of any evaluation approach is vulnerable to some form of criticism. But this is also true for all of the sciences and to all scientific-based research, since no empirical analysis – that means speaking of the empirical world – can claim to be absolutely perfect. The only guarantee for avoiding empirical mistakes would be not to conduct empirical surveys which, obviously, cannot be a goal for the sciences. So the ultimate challenge is (and can only be) to improve evaluation methodology and to offer alternatives that can also be practically applied. Criticism without such an offer of an alternative is, therefore, only of a limited value.

(2) **The systemic need for a feedback:** Taking some of the basic principles of systems theory seriously, this then implies that the quality and efficiency performance of a system depend crucially on the extent to which this system has developed internal structures that communicate a feedback. Phrased in a more simple terminology: *feedback is essential for the viability of a system*. So without internal feedback mechanisms a system runs the danger of slowly and gradually loosing its capability to learn and to adapt to changes in the environment (the context of society).[143] There are many practical examples that demonstrate the correctness of such a statement. Just to illustrate this, the superiority of market economy (»capitalism«) over a planned economy (»communism«) and of democracy over authoritarian government can be partially explained, because market economy as well as democracy have developed feedback mechanisms that enable sensitive and accurate reactions of the overall system. In the case of a market economy clearly the market forces, such as the interplay of supply and demand, and in the case of democracy electoral processes of implementing or rejecting a party (or party coalition) as a government party – and the political and electoral competition of government and opposition parties – guarantee and demonstrate certain feedback effects. Using a general metaphor of systems theory, then the »macro«-system society can be disaggregated into several sub-systems (the German term would be *Teilsystem* or *Subsystem*), such as the economic system (»market economy«) and the political system (»democracy«). So clearly also the universities – in general and also more specifically in the case of Germany – can be interpreted as a system. And such a conceptualization then leads to the following key question: *What are, concerning the universities, the feedback*

[143] For an instructive overview on cybernetical principles in systems theory, see the article *The Science of Cybernetics and the Cybernetics of Science,* authored by Stuart A. Umpleby (1990).

mechanisms that enable the »university system« to learn and find adequate responses to new problems and challenges that arise? Because one fact must be clear: when it is generally accepted (and even demanded) that a system must be able to demonstrate and perform a feedback capability – thus also those high approval rates and the impressive degree of public legitimation for market economy and democracy in the advanced industrial countries –, then this is a general demand which is also applicable to the university system. In other words: it does not appear legitimate or, at least, it will not be accepted by the public – at least in the longer run – if the universities would claim that they represent a system that can survive and perform with quality and efficiency without having incorporated feedback mechanisms into their structures; because such a claim would imply that universities would occupy a privileged position, would demand a different treatment in comparison to other institutions or sub-systems. However, this would probably create the public critique of a double standard of morals. When we consider feedback mechanisms, that means feedback with consequences, as absolutely normal and even desirable for the economic system and the political system (for example, government parties can be elected out of office), then the university system cannot claim a treatment by the public that differs substantially from the treatment of other policy areas. Searching for a possibility how feedback could be processed in practical terms within the university system, we want to stress the crucial importance of evaluations: *perhaps evaluations are the most important strategic key tool for implementing and processing feedback effects and feedback mechanisms for the university system*. Already in Chapter 3.2.1 we proposed the observation that, first of all, for the contemporary German higher education sector something like a national and comprehensive evaluation system for university research and teaching was not implemented, at least not up until now; and, secondly, the public funding formula for German universities does not take evaluation results into account – or, to phrase it somewhat differently, because a comprehensive *ex-post* evaluation system of universities (and university research) currently does not exist, the public funding formula must operate without such evaluation-based data. Derived from this observation, this would imply setting up the provoking hypothesis that within the context of Germany's university system those systemically necessary feedback mechanisms have not matured to a favorable degree. In the following we want to highlight some of those structural patterns and trends that demonstrate the urgent need to implement more determined feedback mechanisms – for example, based on evaluations – into the German university system:

- Concerning the funding resources of German universities and German university research, there are two crucial facts that we should recall: firstly, university R&D depends financially primarily on public basic transfer funds, called GUF. In 1991, 71.4% of German university R&D was funded by GUF (see Figures 17 and 18). Secondly, the German universities claim that they are currently *underfunded* or

underfinanced (*unterfinanziert*). One estimation, that was calculated by the HRK, speaks of an annual structural financial deficit for the whole German higher education sector of between 6 and 9 billion DM (see again our analysis in Chapter 3.1). From that two crucial consequences can be drawn: (1) The dominance of the funding category of GUF implies that *ex-post* evaluations represent more or less the only possibility for evaluating university research systematically, that means having a chance to project transparency onto the system. In addition, GUF funding even demands – theoretically speaking – *ex-post* evaluations since, already by definition, GUF funding denies the possibility of accurate *ex-ante* evaluations.[144] (2) The monetary resources for the German universities are seriously constrained. This determines two consequences for Germany's university sector: first of all, the universities must learn how to use their resources more efficiently and go through a painful restructuring process of adaptation. In addition, and secondly, probably the only chances of the universities to achieve an increase of their public-funding base is to convince the public of the importance of universities. And this will, in its final logic, demand the introduction of comprehensive evaluation systems.

– *Transparency is necessary and essential*. Without an adequate degree of transparency a successful rational decision-making and policy-making is very difficult to realize, since the development of intelligent and sophisticated strategies would demand, for instance, the supply of sufficient data. When such data are missing, then decisions must be based on assumptions which, of course, can be true; this means that assumptions correspond with the empirical world or with the results of an empirical inquiry that is carried out at a later moment in time than the formulation of premises. On the other hand, however, assumptions are often primarily the product of an ideological belief or prejudice which either is empirically wrong or – in a perhaps less dramatic situation – is used as an *ex-ante* argument against a systematic and scientific analysis.[145] *A key strategy to produce transparency within the context of an individual university or, more generally speaking, for the whole university sector is to carry out comprehensive and systematic evaluations on a regular basis*. Such evaluations again guarantee that systemic feedback mechanisms are incorporated and can operate within the university framework. There are several arguments that underscore why such *transparency-producing evaluations* are essential for the universities: (1) First of all, *the universities themselves must learn more about the*

[144] See again Chapter 3.1 where we discuss those different funding categories of universities and university research. Concerning the argument that GUF funding already by its structural design demands comprehensive *ex-post* evaluations, we already formulated that thesis in reference to the Austrian university system (Felderer and Campbell, 1994a, 211, 214–215).

[145] As an example for such an ideological bias we again will cite the hypothesis of a *medium-standard quality* performance in research, at German universities, in the next paragraph (see also Chapter 3.2.2); the proponents of that hypothesis emphasize that the *medium-standard quality* is an empirical fact – at the same time they argue that this empirical pattern makes an empirical investigation impossible, that means a comprehensive national evaluation of universities and of university research.

universities – this implies that there is a need for an increase of the »self-reflexive« competence of universities. A university permanently should rethink its structures and its mission, so that they are optimally prepared to increase the quality and efficiency of their performance in research and teaching. Furthermore, universities should be in a position to respond adequately to new problems and challenges that arise in the societal environment.[146] This, finally, demands that a university is sensitive to feedback and that it admits transparency – in that context the phrase of a *gläserne Hochschule*[147] is used in the German discourse (see Lange, 1995a, 72). (2) *The public of an advanced democratic and industrial society expects transparency from all institutions, particularly when those institutions claim to be of value for the whole nation*. Democracy supports the development of a »civic culture«, where society demands that main actors must be able to present themselves adequately to the public. Now drawing the line to the issue of university research and its evaluation, our hypothesis is the following: *German society expects more transparency from the German universities.* First of all, German universities are primarily public funded, which means that their primary financial resource is derived from the tax money that the Germans (and German residents) pay. Secondly, the German universities claim that their basic research activity is for the good of all of German society. In 1989, on December 11, the German weekly magazine *Spiegel* published the first ranking of German universities based on a question survey of 6000 students. This ignited a series of articles in the following years, where university rankings were presented, however, partially using a different methodology[148] – and finally, in 1993, Spiegel presented a follow-up survey that was based on a sample of 11828 questioned students and that focused on the quality of teaching (published by *Spiegel* as *Spiegel Spezial* 3/1993).[149] Of course those university rankings, in particular their methodology, was severely criticized by many analysts. So some analysts would argue that the fact that the ranking position of individual universities changes from survey to survey and – what is even more important – also changes depending on the applied methodology, just reveals how vague and poorly developed the methodological basis of such rankings is; therefore, those university rankings published by commercial journals should be read and treated with great caution (see, for example, Erichsen, 1995b, 217, and Frühwald, 1995, 211–212). Now irrespective of the validity of that criticism, other experts are inclined to offer a different opinion for those published university rankings: first of all, they express the desire of society and of the public for transparency in general and *for more transparency in the university sector.*

---

[146] Just to give an example, we discussed in the paragraph above those financial constraints under which the German universities currently suffer.

[147] A translation into English of *gläserne Hochschule* would be: a glass university, *i.e.,* a transparent university.

[148] Journals, that presented university rankings, were: *Focus* 39/1993, *Forbes* 6/93, and *Stern* 16/93.

[149] For a more professional description of those *Spiegel* surveys ranking universities, see Hornbostel and Daniel (1995), and Tarnai *et al.* (1995).

Secondly, those published university rankings also indicate a certain failure or at least a neglect on the part of the universities: since the German universities were somehow reluctant to have a thorough evaluation system implemented, the commercial media took over the role to build up a certain pressure. So even when those university rankings, which were finally presented, can be criticized because of their methodology, among other reasons, at least they were successful in creating a public sentiment that is more in favor of evaluations of universities and of university research (see again Daniel, 1995, 205; see also Weingart, 1995[150]).[151] *Another crucial argument, in that context, seems to be that probably only when the public demand for more transparency and accountability in the German university sector is met, then there might be a public consensus for significantly increasing the public funding base for universities*.[152] (3) Interestingly enough, also German unification had an effect favoring evaluations and supported this transparency demand – of the public but also at the level of experts. One consequence of unification was that the institutes of the former East German Academy of Sciences were comprehensively evaluated by the »West German« Science Council in the years 1990 and 1991 (Krull, 1994; Mayntz, 1994; Wissenschaftsrat, 1992a and 1992b). This produced two crucial consequences: first of all, this implicitly acknowledged that research institutes can be evaluated systematically and comparatively. Secondly, the evaluation of a whole sector of the old East German research system created a demand for a *policy symmetry* or a *policy balance:* When East German R&D institutions are evaluated, what is the justification not to evaluate certain sectors of the West German national research system?[153],[154] (4) As many experts indicate, the German universities must be aware of the following scenario: should the universities fail to meet the transparency expectations of the public or, to be more precise, of those public agencies that finance the universities (most prominently the ministries at the *Länder* level), then the universities run the risk of being *informally evaluated*. Decision makers at the ministries obviously have their personal impression of the quality or efficiency of the

---

[150] »Stattdessen spielen die Medien eine Vorreiterrolle. Die beiden ›Spiegel‹-Umfragen zur Attraktivität des Studiums an verschiedenen Universitäten haben die mögliche Funktion derartiger Informationen offengelegt: Vermeintliche Provinzuniversitäten wie Siegen und Bielefeld haben tausende zusätzlicher Studenten gewinnen können, die auch unter den eingeschränkten Konkurrenzbedingungen der staatlichen Kapazitätsarithmetik eine Erfolgsprämie bedeuteten« (Weingart, 1995, 74).

[151] There are other analysts, however, who would disaggree with such an interpretation. Hartmut Schiedermair, for instance, asserts that those published university rankings also emphasize a wrong perception or even prejiduce that makes the universities responsible for failures that actually occured in other policy areas: »Den Ranking-Listen der Publikumszeitschriften mag man zugute halten, daß sie die Diskussion über die schwierige Lage an den Universitäten öffentlichkeitswirksam umgesetzt haben. Sie sind allerdings auch jener merkwürdigen Theorie verfallen, die die Universitäten nun schon seit Jahren vom Opfer zum Täter einer verfehlten Bildungspolitik umdefiniert« (Schiedermair, 1995, 216).

[152] At the end of Chapter 3.1 we already mentioned that argument for the first time.

[153] »Das Tabu der institutionenbezogenen Bewertung ist durch die Evaluierung der Akademie der ehemaligen DDR durch den Wissenschaftsrat im Zuge der Wiedervereinigung gebrochen« (Weingart, 1995, 74).

[154] In Chapter 4 we will discuss this evaluation initiative of the East German Academy research institutes in more detail.

performance of individual universities. And those personal impressions – which are the sum of a series of »subjective« indicators – might be aggregated to an informal assessment which, however, can have consequences in real politics: coming down to the level of practical policy decisions, for instance the question of funding appropriation for a university or a university department, those informal evaluations will be used *de facto* as a basis for crucial public decision-making. In such a scenario the universities clearly are worse off than in the case of an »official« evaluation. *Informal evaluations* imply the following disadvantages for the universities: firstly, their methodology, implementation, and procedure is pre-mature and in no way comparable to the quality standard of an official evaluation; secondly, the methodology and the results cannot be openly discussed in the context of a professional discourse, since »officially« such an evaluation never was carried out – this implies that the universities cannot defend themselves properly; and finally, possible learning processes for the universities (and other institutions) are curtailed because there is no open and wide-spread diffusion of the results of *informal evaluations*.

– To illustrate further the above said, *i.e.* the necessity for an increased transparency, we again want to refer to the previously mentioned assertion that all German universities perform at a very similar level of standard quality – which could be designated as a *medium-standard quality* with little or no variance – in research (and in teaching). This claim is based on several assumptions: culturally it is seen to be a consequence, firstly, of the consensus-oriented and competition-critical attitudes of German society and, secondly, also a result of the Humboldtian principle of the »unity of teaching and research« (see Chapter 3.2.3). Therefore, when analyzing the structural constraints against a systematic application of evaluations – as we did in Chapter 3.2.2 –, then this argument of a *similar-or-medium-standard-quality* of university performance in research (and teaching) is used as an »anticipatory« statement that already in advance should prove that evaluations are meaningless in the case of Germany; the proponents of such a hypothesis stress that because of the similarity in performance, any evaluation-based conclusion and any university ranking would automatically be exposed to a massive methodological criticism, in the sense that the reliability of the results is uncertain and that a small variation in evaluation methodology could easily alter the outcome of a specific evaluation procedure (see again Chapter 3.2.2). We personally, however, are skeptical about such a line of argument. The crucial point seems to be that there must be a clear-cut distinction whether a statement is, in its character and essence, ideologically *ex-ante* or empirically *ex-post*. So to assert already in advance that the German universities express a similar performance standard clearly is an ideological and not an empirical argument. What, however, does not appear legitimate is when such a statement is used as an argument or even as a »proof« for the non-applicability of evaluations; because this would imply that ideology would be valued higher than empirical

research, such as scientific inquiry. From a scientific perspective such a claim, obviously, is unacceptable. Therefore, when summarizing the debate whether or not the German universities perform at a similar and low-variance level of standard quality, we want to stress the following two arguments: firstly, there is no other possibility than to test and analyze empirically if such a hypothesis is correct and an ideological argument may never be confused with an *empirical ex-post conclusion;* secondly, only after – and not already before – a comprehensive evaluation was carried out, will we really know more about the performance and quality distribution of German universities.[155]

– Germany's massive economic build up in the post-war period, and its economic strength in contemporary Western Europe, was seen by many analysts, over a long period of time, as a proof of the superiority of the German model of society and economy. Now irrespective of Germany's economic success story (*Wirtschafts-wunder*) and its economic advantages across many fields, it cannot be ignored that also Germany's economy expresses some problems of competitiveness. At the international, *i.e.* non-European markets in general and more particular at the international technology and R&D markets, Germany's economic competitiveness seems to be under pressure. This impression again is compatible with our observation that Germany's academic research system does not, when compared with other OECD countries, lead in efficiency when the publication output of articles in international journals is taken as frame of reference.[156] From that it follows that a policy that aims at improving the overall competitiveness of the German economy must also be interested in a careful monitoring of the university research. Since university research occupies a strategic niche within the context of a national system of innovation – for example, fulfilling the function of a knowledge transfer from basic to applied research –, there is the danger that a comprehensive and competitiveness-encouraging policy will only be partially successful when the university sector is not taken into account. Therefore, systematic evaluations of the quality and efficiency of university research must also be regarded, in their final consequence, as a crucial contribution to a general masterplan that wants to increase the competitiveness of the economy – particularly for those international markets.

---

[155] While discussing in Chapter 3.2.2 the structural constraints against evaluations of university research, we also mentioned (in paragraph number five) the so-called *constraint of federal and decentralized decision-making:* this means that because there is a fear that those *Länder*, with the »good-quality« universities, might be favored, there was some resistance at the political decision-making level of the *Länder* to introduce comprehensive and competition-emphasizing evaluations. This, however, clearly demonstrates that even many key decision makers have massive doubts whether this hypothesis of a similar quality performance of all German universities is really true.

[156] See again our in-depth analysis in Chapter 2.2 where we arrive at the conclusion that, in the case of Germany, the application of bibliometric indicators produces results that appear highly compatible with other indicators, such as international competitiveness or the technology balance of payments. For further information, see again: IMD, 1996; OECD, 1992b, 1994a, and 1996a.

– Summarizing often-raised statements that emphasize the necessity of feedback mechanisms within the context of a university system, we want to stress the following arguments: (1) Feedback – processed by evaluations – should create *transparency*. As a result the universities, often perceived by the public as a »black box«, should become a »white box«. Also in that context the phrase of a »glass university« (*gläserne Hochschule*) was created (see Lange, 1995a, 72). This would also help raising the legitimacy of universities in the view of the public. (2) Feedback should support the »self-reflexive« capabilities or the systemic »self-awareness« of universities: universities should learn more about themselves, which again would be a basis for improving the development of proper strategies. In practical terms this would imply correcting errors or structural deficiencies and improving the quality and efficiency of the university performance. (3) Feedback would help *defining the frame of reference,* against which the universities could position themselves; firstly, they would know *where* their location is; secondly, they could learn *what* their goals and *what* the new problems and challenges are, to which an adequate response must be found; thirdly, universities could decide more easily and more professional *into which* direction they should move in the future.

3.2.5 Evaluation Policy of University Research

In the following we will present some of the policies that are either discussed or actually even applied (or that are at the beginning of an application), and that aim at evaluating university research in Germany. Speaking more generally, such evaluation policies only rarely are *ad hoc* initiatives; rather they should be understood as consequences, that are drawn, and that mark the concluding phase of a long learning process. On the other hand, of course, we also should not forget that evaluation policies are exposed to an evolutionary drive. This means, an evaluation policy must be sensitive for new trends or new problems that permanently arise or, in other words, also *evaluations must be evaluated regularly*.

(1) **The necessity of a combination of an *ex-ante* and *ex-post* quality control system:** Already in Chapter 3.2.3 we presented the fascinating theory of Detlef Müller-Böling, who describes the German universities and German academic culture as a system, believing – at least »officially« over a long period of time – in the possibility of an *ex-ante quality control:* this implies postulating that rigorous one-time quality checks can guarantee a limitless quality performance of a process (Müller-Böling, 1995, 31–34). To give an example: this explains why in Germany the assignment of a professorship normally is tenured, that means temporally not limited.[157] In the previous chapter (Chapter 3.2.4), however, we summarized those arguments that are put forward by many experts and that emphasize that in the future even in Germany

[157] In Chapter 3.2.2 we qualified this practice as a *structural constraint* against the application of evaluations in Germany.

evaluations of universities in general and, in particular, evaluations of university research will be of a growing or even crucial importance. One of the underlying systemic core ideas is that each system must develop some feedback mechanisms that enable the system, first of all, to learn and, secondly, to be in a position to enable adequate responses in the context of a society which is permanently changing. In principle a very similar argument is that of Müller-Böling, who expresses the opinion that an *ex-ante* quality control, as the only steering instrument, does not meet and fulfill the demands that arise in modern and contemporary society. Müller-Böling names three reasons for his diagnosis: firstly, the conditions that are linked to an appointment of a professor appear to be artificial, because nobody can predict developments over the next twenty-five years in a rapidly changing world; secondly, attempts to coordinate processes that should sustain a certain homogeneity across the German university sector are too slow to be efficient;[158] thirdly, the degree and intensity of investment of public monetary resources into the university sector are no longer sufficient (Müller-Böling, 1995, 32). Therefore, Detlef Müller-Böling emphasizes that a *combination of an ex-ante and ex-post quality control or quality steering* seems absolutely necessary.[159] In other words and interpreted freely: without the incorporation of some basic elements of an *ex-post* evaluation-policy system the German universities will face, in the long run, serious constraints and perhaps also serious problems in defending their role as a key supplier of high-quality research for the German society and economy. In practical policy terms, Müller-Böling indicates two issues that should be given a salient priority (Müller-Böling, 1995, 33–34):

– The autonomous university must engage actively in a process of goal formulation; at the same time, the university is accountable to the public and to society in general.[160]

– The goals and the performance of a university must be made transparent.

(2) **The importance of indicators and of quantitative indicators for evaluations:** Potentially any evaluation policy is caught in the dilemma of emphasizing *either quality or quantity.* To phrase this dilemma somewhat differently, one can set up the

---

[158] In Chapter 3.2.3 we discussed in great detail the *cultural constraints* against evaluations of university research. In that context we referred to the Humboldtian principle of a unity of teaching and research and, in addition, mentioned the consensus-oriented and competition-critical attitudes of German academic culture. As a consequence of that there is (or at least was) a strong desire among German academics to establish a *homogenous medium standard of quality,* without too much variance, across Germany's university system. However, as Müller-Böling argues – according to our opinion —, the costs of sustaining such a »homogeneity« are much higher than the potential benefits.

[159] »Erfolgreich wird daher nur eine Kombination aus ex-ante und ex-post-Steuerung sein, die einerseits an den formulierten Zielen und andererseits am Zielerreichungsgrad (den Ergebnissen) ansetzt, wie sie im übrigen in fast allen westeuropäischen Ländern bereits praktiziert oder augenblicklich eingeführt wird« (Müller-Böling, 1995, 33).

[160] For further literature on some of the systemic fundamentals that underpin a process of goal formulation, see, for example, Umpleby and Sadovsky (1991).

following question: Should a model, that underlies the evaluation process of university research (or other areas of university performance), try to cover the »total spectrum of complexity«, that means to recognize all the details of a university institution, or, on the contrary, should an evaluation model preferably focus on a limited number of indicators which are considered to be of a crucial importance? This means, what is the role or function of *quantitative indicators* in the context of an evaluation procedure? Obviously, there are divergent opinions of experts on that issue. In 1988, for example, Rudolf Fisch and Karl Alewell published separately the blueprint for a complex model of how universities and university research could be evaluated comprehensively (Alewell, 1988; Fisch, 1988; see also Alewell, 1995; and Sinz, 1995). Hans-Uwe Erichsen, for instance, asserts that the overall performance of a university cannot be adequately expressed within only one quantitative dimension; Erichsen admits that quantitative indicators are a useful device for assessing the performance of a university;[161] but, at the same time, Erichsen also emphasizes that each indicator can only represent a partial spectrum of a university performance.[162] Other experts, on the contrary, are much more inclined to emphasize and to underscore directly the *crucial importance of indicators* – also of quantitative indicators. Their conclusion is that, regardless which model is used, only those evaluation policies of university research can be efficient and produce meaningful results that *also* employ the use of indicators.[163] Obviously, *the concept of indicators or of quantitative indicators* can have two separate and distinct meanings: (1) The first and more simple implication is to interpret the function of indicators as counting directly units which are measurable. Examples would be the size of research staff, the number of patents, the number of publications, or the number of diplomas, and so on. In reference to the concept of first-order and second-order cybernetics, those indicators, consequently, could be classified as *first-order indicators*.[164] (2) The second group of indicators – that, within the same line of argument, could be paraphrased as *second-order indicators* – are already of a higher complexity, since their functional purpose is not just to count, but to interpret, or, in other words: *to translate quality into quantity*. That means, they aim at representing structures and processes – which we might label as »qualitative« – within a quantitative setting or, speaking simply, as numbers. *So, finally, qualitative patterns of the empirical world can also be reflected and expressed by quantitative indicators*. To illustrate this, we want to give two examples. **First example:** The

---

[161] »Um gleichwohl der Leistung von Hochschulen auf die Spur zu kommen, werden Indikatoren verwendet« (Erichsen, 1995b, 216).

[162] »Mit Indikatoren lassen sich jedoch nur einzelne Aspekte der Leistung beschreiben, es gibt bisher kein überzeugendes Modell, die Gesamtleistung einer Hochschule zu erfassen und in einer Maßzahl zum Ausdruck zu bringen« (Erichsen, 1995b, 216).

[163] »Die Beurteilung der Leistungen in Forschung, Lehre und Dienstleistungen sollte jeweils anhand eines Indikatorenbündels erfolgen« (Daniel, 1995, 206).

[164] For a further discussion of the interesting concept of first-order and second-order cybernetics, see Umpleby (1990). That term was originally invented by Heinz von Foerster and published, for the first time, in 1979 (von Foerster, 1979).

Research Assessment Exercises, that are carried out in the United Kingdom for the purpose of comprehensively assessing university research, focus in their core procedure on evaluating primarily the quality and not the quantity of publications of the academic research staff. In a second step, however, this qualitative assessment is then translated into a »quantitative-like« ranking of universities and their departments.[165] **Second example:** In the Chapters 2.1 and 2.2 we compared the academic publication output and publication efficiency of Germany with that of other OECD nations, but limited our analysis to articles that are (or were) published in international journals that again are covered by the two data bases SCI and SSCI. We justified this methodological approach by arguing that such journal-oriented bibliometric indicators can be interpreted as a manifestation of »quantified quality«: since those SCI and SSCI based journals normally have a peer-review system, this automatically implies that all approved articles must pass a quality check. Therefore, counting articles in peer-reviewed journals does not mean just counting publications, but it represents a counting of publications that exhibit a standard or above-standard quality (see again our in-depth discussion on that topic in Chapter 2.1). *Now again referring to those experts who strongly emphasize that a sophisticated and professional evaluation policy of university research demands the development and extensive use of indicators,* we want to summarize those key arguments that support such a position:[166],[167]

– Qualitative interpretations are proper for representing the whole »width« or spectrum by covering all the details. One of their main deficiencies, however, seems to be that often the overview or survey is either missing or unclear. And this is exactly one of the main strengths of an indicator-based analysis or evaluation: producing a clear-cut overview or a well-defined image or model, that quickly communicates to the observer the essential information. Using indicators during an evaluation process of university research – or of universities in general – implies that the actual or potential strengths and weaknesses of the system are rather easily to detect. In addition, as Weingart would claim, an indicator-based evaluation can be carried out faster and cheaper than a complex and primarily qualitative-oriented investigation (Weingart, 1995, 77). In other words: indicators offer a quick, easy, and »cheap« overview.

---

[165] Refer to Chapter 3.2.1, where we summarize the basic features of those British Research Assessment Exercises.

[166] So our article-based bibliometric analysis and comparison of Germany with other OECD countries, in Chapters 2.1 and 2.2, demonstrates clearly how indicators (or an indicator-based model) can be used for evaluating the performance and efficiency of academic research.

[167] An excellent summary of those indicator-favoring arguments is given by Peter Weingart, who emphasizes: »Auf der anderen Seite finden sich Evaluierungsversuche, die in ihrer Anlage zu komplex, in ihrer Aussagekraft zu vieldeutig, in ihrer Erhebung zu teuer und damit letztlich geeignet sind, den längerfristigen, entscheidungsrelevanten Einsatz zu verhindern« (Weingart, 1995, 76–77).

– An argument that should prove why in the case of Germany the application of comprehensive university evaluations at the national level are impossible, is that of size. Some analysts assert that because of the »quantitative dimensions« of Germany's university system or the higher education sector in general, a systematic evaluation attempt will be confronted with serious constraints and might collapse in the face of the quantitative complexity.[168] Obviously, we personally disagree with such a statement and emphasize that the size of Germany is not an argument against carrying out evaluations, but it is an argument for how evaluations are processed. This means that the scheme of a comprehensive evaluation, already in its planning stage, must somehow take account and be sensitive of the impressive size of Germany's higher education sector. Therefore, the development of »intelligent« indicators should be judged as crucially important for Germany; because without such indicators the establishment of an overview across the diversity of the German university system, and its performance, does not appear conceivable. In the face of the complexity and size of Germany's tertiary education, an observer easily could get lost.[169]

– Closely linked to this function of creating an overview is that of offering a basis for decision-making. The crucial point for decision-making seems to be that decisions demand that a hierarchy of options should be defined, and this resembles a process similar to that of setting up indicators. What both have in common is that the empirical complexity must be simplified and transformed into a complexity-reducing structure; because if the total information complexity would be taken into consideration, perhaps in a one-to-one relationship, this could imply that the action or policy-implementation by a decision maker might be continuously paralyzed. Therefore, *the development of complexity-»interpreting« and complexity-reducing indicators substantially supports processes of policy-oriented decision-making*. Of course, each decision-making unit faces the following classical twofold dilemma: firstly, the empirical complexity must or should be analyzed in all its details; secondly, however, in a next step this complexity again must be reduced to obtain a basis for priority-opting and policy-oriented acting.

---

[168] In Chapter 3.2.2, in paragraph number four, we already mentioned this argument of size for the first time, which is interpreted by some experts as a structural constraint against evaluations of university research.

[169] In that context we want to recall that the first European country, which introduced comprehensive evaluations of its university system and established a linkage between evaluation results and funding, was the United Kingdom. And the UK certainly cannot be qualified as a small-sized European Country. In fact, when the UK is compared with former West Germany, the argument of size loses most of its credibility – at least in our opinion. Should the number of researchers (or university graduates) in the higher education sector, in full-time equivalents, be taken as reference, then we obtain for the year 1989 the following number of research-person years: Germany 38835, and in the UK 27000 (see OECD, 1996a, 39). This means, when the former West German research-person years at universities are indexed at 100, the UK still can claim a value of 69.5. And the difference between both values does not justify the argument that the one country is too large for evaluations, while in the other countries the evaluations are already a political and empirical fact.

– There is a wide-spread consensus among experts that transparency is – or would be, when thought in »how-it-should-be« categories – a vital feature for universities. Only transparency guarantees, at least to a certain extent, that sensitive feedback mechanisms can operate within a university context which again is crucial for sustaining the performance of a university system.[170] Without transparency also a performance-dependent allocation or redistribution of resources, as demanded by some analysts, does not appear possible. *So a key function of indicators is exactly to create and diffuse transparency.* In practice such indicators can be judged as an easy, cheap, and quick possibility for establishing a transparency survey. And the comprehensive set-up of such indicators, already by definition, could be the goal of a, perhaps first or preliminary, evaluation procedure. Thus transparency and evaluations are mutually linked to each other.

(3) **The practical procedure of developing indicators:** Already at the beginning of the 1990s, in the year 1991, the HRK – German Rectors' Conference – initiated a pilot project, called *Profilbildung*.[171] In the first phase, completed in 1993, the attention focused on the disciplines (*Fächer*) physics, German language (*Germanistik*), and economic sciences. In a second phase, finished in the year 1994, the disciplines covered were electrical engineering, mechanical engineering, and computer sciences (*Informatik*). The primary purpose of this project was to invent and to develop *ad hoc* models and procedures, still at the bottom-up level of individual universities, which, in a later phase, could have the potential to be regularly applied to all universities nationwide. Consequently, these assessment exercises did not cover the whole university system, but only a small sample of those universities that participated voluntarily. During the second phase those were, all together, not more than eleven universities. Functionally the assessments were institution-oriented, that means they preliminarily attempted to evaluate or, to be more precise, they described comprehensively the performance in research and teaching in those pre-defined disciplines (as listed above) at the level of university departments as well as for the whole university. Methodologically, a twofold approach was applied: first of all, descriptive statistical data were collected – a so-called *Erhebungsraster* – with the main purpose being to develop indicators that present a broad overview of the general performance. Secondly, the data overview was accompanied by a verbal or narrative description that intended to offer to the universities the opportunity to comment on their situation and their individual performance in research and teaching. Obviously, what this pilot project did not aim at was to develop and to set up for discussion a quality-based *ranking* of different universities or university departments. So, in our terminology, the design of »second-order« indicators was only a partial goal or

[170] In Chapter 3.2.4 we discussed in great detail, including several examples, why there is a need and demand to have feedback mechanisms that are structurally built into the system.
[171] *Profilbildung* could be translated into English with the phrase »development of a profile«.

intention (for further literature, see HRK 1994b, and HRK 1994c). According to Hans-Uwe Erichsen, the collected data – which should have produced some transparency within the »black box«, called university – contained the following information (Erichsen, 1995b, 219):

- The number of new student entrants and the total number of students;
- distribution of students according to the number of academic semesters or the number of subject-related semesters (*Fachsemester*);[172]
- number of »officially« planned and regularly resourced »working environments« for students (*Studienplätze*), and the number of applicants;
- number of examinations for the first-level academic degree, functionally differentiated according to divergent types of tertiary education (*Diplom, Magister,* or *Lehramt*);
- comparison of the number of new student entrants with the number of graduates (with a time lag of five to six years, in the case of the *Fachhochschulen* of only of four years);
- number of graduates with a doctoral degree (and compared with the number of those who completed their dissertation four years before);
- number of those who completed a *Habilitation;*[173]
- number of staff, differentiated according to professorships, academic or scientific staff (again distinguished between temporary and permanent positions), and non-academic or non-scientific staff;[174]
- number of scientists who are not financed by earmarked funds (Drittmittel);
- ratios that indicate the support for teaching: number of students per one academic or scientific staff member; number of students in their fourth subject-related semester (Fachsemester) per one academic or scientific staff member; number of new student entrants per one academic or scientific staff member;

---

[172] The German term *Fachsemester* is translated into English with »subject-related semester« (compare BMBF, 1995a, 276, with BMBF, 1995b, 116).

[173] As already mentioned in Chapter 2.1, the *Habilitation* is sometimes translated or circumscribed in English with »higher doctorate« (see again Irvine et al., 1991, 52).

[174] With »academic or scientific staff« we circumscribe the German term *wissenschaftliches Personal;* and, consequently, »non-academic or non-scientific staff« addresses the *nicht-wissenschaftliches Personal.*

- number of scholarship holders and prize-winners, differentiated according to the source Humboldt, Leibniz, and Hess;

- earmarked funds (Drittmittel), differentiated according to the source (DFG, federal and Länder ministries, foundations and private resources, resources of the EU).

  Now based on those collected descriptive statistical data, the following ratios were calculated, in a second phase, to obtain a more pronounced picture and also an analytically more challenging overview of the performance of those universities that participated in the Profilbildung pilot project survey (see again Hans-Uwe Erichsen, 1995b, 219):

- number of professors *per* number of academic and scientific staff members;

- number of non-academic or non-scientific staff members *per* number of academic or scientific staff members;

- new student entrants *per* number of academic or scientific staff members;

- students *per* number of academic or scientific staff members;

- students progressing in the »normal time« (*Regelstudienzeit*) *per* total number of students;

- number of examinations *per* number of professors;

- number of examinations *per* number of academic or scientific staff members;

- number of completed dissertations (*Promotionen*) *per* number of academic or scientific staff members;

- number of completed dissertations (*Promotionen*) *per* number of professors;

- earmarked funds (*Drittmittel*) *per* professorships;

- and earmarked funds (*Drittmittel*) *per* number of academic or scientific staff members.

(4) **Principles for proper use of indicators:** Hans-Dieter Daniel elaborated a set of principles which should be taken seriously when the evaluation of academic performance is at stake. In the following we will summarize some of those principles

that are qualified as crucial for a proper use, application, and employment of indicators (see Hans-Dieter Daniel, 1995, 206–207):

- A comparison of performance across distinct disciplines, particularly across science (»natural sciences«), the social sciences, and the humanities, could cause a series of problems. Different disciplines are confronted with different environmental conditions and a distinct supply of resources; for instance, the number of students and graduates, acquired earmarked funds (*Drittmittel*), and behavioral patterns of publishing and being cited in other publications. Therefore, as Hans-Dieter Daniel argues, a focus should be given to distinct and individual disciplines and to individual university curricula and study programs.[175]

- The assessment of performance in research, teaching, and academic service demands the development of a broad spectrum of indicators. These indicators should reflect quantitative and qualitative aspects of performance and can use or rely on *subjective data* – for instance, assessments by students or the expertise or knowledge of experts – or *objective data* (or, in the words of Hans-Dieter Daniel, »quasi« objective data), such as: number of academic semesters, number of graduates, amount of earmarked funds (*Drittmittel*), number of publications, citations, and patents, and so on.[176]

- The data should be presented within the framework of a »performance profile« (*Leistungsprofil*) that presents a comparison of information oriented to the strengths and weaknesses of different disciplines or of different *Fachbereiche,* as often preferably phrased in German.

- Derived from practical experience, the performance or quality variance *within* a specific discipline is often larger than *between* different disciplines – when measured or indicated on the basis of a »virtual« or »constructed« performance or quality mean for each discipline. Therefore, as Hans-Dieter Daniel stresses, it is important to determine and to identify who really the persons are who perform most of the quality and output. Daniel also cites two authors who independently claim that most of the publishing is produced by a minority within the scientific communities.

---

[175] However, as we would like to add, in conscious contrast to Hans-Dieter Daniel, also a comparison of diverging disciplines can be conceptualized meaningfully when those differing starting points of individual disciplines are recognized and somehow incorporated into the analysis.

[176] Obviously citations are just as important as publications, since scientific output that is not recognized by the scientific community – that means, to give an example, publications which are not cited – can only claim to be of limited value. Norbert Untersteiner, for instance, asserts that in the humanities (*Geisteswissenschaften*) almost ninety percent of the published articles are never cited – although, unfortunately, Untersteiner fails to name a source for his statement (Untersteiner, 1995, 145).

- An adequate assessment of performance must always take the input of resources into account. This is necessary to prevent that automatically, and by this perhaps falsely, those disciplines (*Fachbereiche*) are evaluated and interpreted more positively that operate on the basis of a generous supply of resources; efficiency cannot be derived from the degree or intensity of output, but is the consequence of a favorable input/output relationship.

- An evaluation should cover a period of three to five years. If the temporal periods are longer, then this might imply the danger that the indicators are outdated; and the problem of shorter periods is that this might lead to the delivery of unstable data and thus questionable results. As Hans-Dieter Daniel emphasizes, such a claim of a temporal optimum of a 3–5 year period for evaluations, is based on practical experience.

- Indicators, data, and results should not be published without sufficient comments or documentation. This is to prevent wrong or unreliable conclusions from being drawn.

(5) **Earmarked funds (*Drittmittel*) as a key indicator for the quality and/or relevance of academic research:** Among leading experts, and also among decision makers, many express the opinion that earmarked funds – in German called *Drittmittel* –, or the »degree« of earmarked funding, *should be valued and interpreted as a crucially important variable for the process of evaluating university research*. In other words: earmarked-funded academic research is a key indicator for *the quality as well as the relevance* of university R&D. Such a claim can be based on the following arguments: (a) Since earmarked-funded research – no matter if it is processed either in the context of a research project or a medium-scale or large-scale research program – is always pre-assessed during a peer-review procedure (that follows the application phase), this implies that such a research already automatically will be valued by an *ex-ante* evaluation. Beyond that, because the outcome or »final product« again is assessed by those institutions or organizations that provided the earmarked funds, this leads to the consequence that earmarked-funded academic research must pass a *double quality check;* an *ex-ante* evaluation at the beginning and an *ex-post* evaluation after completion. In contrast to that, university research that is financed by basic funds, is not exposed to an in-depth and rigorous *ex-ante* quality control procedure and, in addition, is only very selectively assessed *ex-post,* since a systematic and comprehensive *ex-post* evaluation system for university research still has not been implemented in Germany (see again Chapter 3.2.1, where we arrived at that observation for the current *status-quo* in Germany).[177] Therefore, because of this »double quality control«, earmarked-funded research can be interpreted as an

[177] In Chapter 3.1 we described in detail the different funding categories of universities in general and of university research.

indicator for quality. (b) Now depending on the specific source, such earmarked-funded academic research can also be *an indicator for the relevance* of that research – and its output – for society in general, or for the economy, or other decision makers in more particular. When an earmarked fund is provided by a company or firm of the business enterprise sector, or other private or public institutions (for instance, a ministry), which – all together – commonly can be circumscribed to have a strongly focused application-oriented interest into the practical results and conclusions of academic research, then it appears legitimate to interpret, in such cases, the degree of earmarked funding as an indicator of relevance. In other cases, where an institution providing earmarked funds either does not have a »political« interest at all or at least not a direct interest in practically applying the research results – for example, the DFG and, to a certain degree, also the foundations –, the function of a quality indication of those earmarked funds still is continued. (c) Earmarked funds have the great advantage that they are relatively easy to survey. In other words: the development and build up of a data collection, that gives information on the degree of earmarked funding of university research, can be done at a reasonable cost or price. So when there is, firstly, an interest in analyzing the quality and relevance of university research, and, secondly, it is accepted that earmarked funds are a proper indicator for that, then an investigation of the degree of earmarked funding is the »cheapest« and fastest way to obtain information on such an issue. Therefore, in summarizing, the simple equation would be: *the higher the degree or intensity of earmarked funds in relation to the extent of basic funding, the more relevant and/or the higher the quality of research* that is conducted by a university department or a university entity as a whole. On the other hand, of course, such earmarked funds may also not be overly interpreted as an indicator. In that context it is important to recognize possible differences between disciplines; that means that because of their intrinsic structures or goals some disciplines might be, from the beginning, in a more favorable position to acquire earmarked funds. As some expert are inclined to indicate, perhaps disciplines in science (*e.g.,* engineering) or in the social sciences (for example, business administration or economics) have an easier access to earmarked funds than, for instance, disciplines of the humanities. Therefore, the value of an indicator such as *the degree of earmarked funding* seems to be the highest when that indicator is used for assessing the *quality and relevance of research* of university departments located within the context of the same discipline; or, as a minimum condition, during the course of applying the earmarked-funding indicator one should keep in mind possible »structural deviations« between disciplines in science, the social sciences, and the humanities. This enables a strict agreement with those principles which were outlined by Hans-Dieter Daniel and are designed as guidelines for a proper use of indicators: Daniel emphasizes that an indicator-based performance assessment promises, under such a condition, the most reliable results when the comparison takes place within the same discipline or within a group or family of

similar disciplines (refer to the previous paragraph number four and, in addition, Hans-Dieter Daniel, 1995, 206).[178] Finally, arriving at a conclusion that reveals policy relevance, we want to add that in two German *Länder* a funding formula for universities is in the process of discussion (or already at the beginning of a regular implementation*) that orients itself primarily towards the degree of earmarked funding*. In other words: those universities should be »rewarded« that prove to be successful in acquiring such earmarked funds; and, consequently, universities which are not that successful in that respect might be »punished«. In the following we want to summarize shortly those two *models* that are designed to underlie the public funding formulas of two *Länder* ministries and thus have the potential of substantially altering the German university system in the future.

Model One:

- **Basic supply or funding** (*Grundausstattung*): 20% of the resources are used for that funding. The distribution key is calculated in reference to the number of professors and in reference to half of the number of academic or scientific staff members (*wissenschaftliches Personal*). The *Fachhochschulen* are weighted with the factor 0.8; the natural sciences, medicine, and engineering are weighted with the factor 1.5.

- **Additional supply or funding for teaching** (*Zusatzausstattung Lehre*): 45% of the resources are devoted to that funding type. The funding is calculated by a formula that refers, using equal weighting, to the number of students progressing in »normal time« (*Regelstudienzeit*) and to the total number of graduates for each university. The weighing factor for the *Fachhochschulen* is 0.8.

- **Additional supply or funding for research** (*Zusatzausstattung Forschung*): 30% of the resources are used for that funding. The allocation of resources depends totally on the degree and intensity of earmarked funds (*Drittmittel*) that were acquired by each university.

- **Additional supply or funding for the promotion of young scientists** (*Zusatzausstattung Ausbildung wissenschaftlicher Nachwuchs*): 5% of the resources are allocated for that purpose. The distribution is derived from the number of those who completed a dissertation and those, weighted with the factor 10, who successfully passed a *Habilitation* procedure.

---

[178] Peter Weingart expresses a very similar opinion by arguing that not all disciplines are equipped with the equal opportunity of accessing earmarked funds: »Der Drittmittelindikator ist zwar gegen diese direkte Form der ›Umkehrung‹ immun. Aufgrund der sehr unterschiedlichen Forschungskulturen und auch der politisch vorentschiedenen Prioritätensetzung der Forschungsförderung ist ein disziplinenübergreifender Vergleich jedoch fragwürdig und müßte zur Ablehnung führen. Für die insgesamt geringer alimentierten Forschungsbereiche ergeben sich keine Lernchancen, die sie zu Anpassungsstrategien motivieren könnten. Wenn dieser Indikator eingesetzt wird, sind zumindest gebietsbezogene Normalisierungen erforderlich« (Weingart, 1995, 79).

Model Two:

- **Basic supply or funding for teaching and research** (*Grundausstattung Lehre und Forschung*): 30% of the resources are allocated for that funding mode, and the distribution is calculated according to the total number of academic or scientific staff members (*wissenschaftliches Personal*), including the professors. The following weighting factors are applied: professors in science or the »natural sciences«, engineering, and veterinary medicine: 2; professors in all the other disciplines: 1; other academic or scientific staff: 0.5.

- **Additional supply or funding for teaching** (*Zusatzausstattung Lehre*): For that purpose 40% of the resources are used, which are distributed according to the total number of students.

- **Additional supply or funding for research** (*Zusatzausstattung Forschung*): 25% of the resources are directed towards that funding type. The allocation is based on the amount of acquired earmarked funds (*Drittmittel*).

- **Additional supply or funding for the promotion of young scientists** (*Zusatzausstattung wissenschaftlicher Nachwuchs*): 5% of the resources; calculated on the basis of dissertations and completed *Habilitationen*.

(6) **The university-internal redistribution of resources depending on performance:** At several German universities models are being discussed – or already implemented – that should enable *a university-internal redistribution or re-allocation of resources depending on the performance of single university departments* (or of other units within the context of a university). Such a university-internal redistribution process should fulfill several functions, such as: (a) first of all, any redistribution must be based on »objective«, »quasi-objective«, or »objectively created« data and data-based indicators. This is necessary so that redistribution decisions can claim a sufficient degree of legitimation within a university community. Therefore, as some experts would argue, such discussions or – even more – decisions on redistribution massively foster a thinking and acting *that favors the design and development of indicators* that have the capability of adequately reflecting the performance of individual universities. (b) As the next conceptual step*, those university-internal redistribution processes should help spreading transparency within the environment of a single university*. Since, as we just have elaborated, redistribution decisions demand the development of a data base, the regular and continuous reporting of data becomes a standardized routine; and, as an essential by-product, this leads to an overall increase of transparency within that university entity. (c) *The performance-based redistribution of resources should increase the »internal« rationale of the system*. In simple terms, redistribution implies the following formula: university departments or academic

university staff (*e.g.,* professors) with a good performance should be »rewarded« (by an increase of resources) and departments with a poor performance should be »punished« (by a decrease of resources). The rationale is the following: Firstly, a performance-based allocation appears to many observers or analysts to be more just or fair than an allocation which is based mainly on »historical« claims. Secondly, either in case of a general increase or decrease of the funding base for universities, such performance-based re-allocation processes offer a mechanism or key for passing on those financial shifts and changes to the level of individual departments. Thirdly, a long-duration impact on the academic culture seems possible in the sense that a general attitude is encouraged that favors performance, output, and some forms of efficiency. Fourthly, university-*internal* competition is encouraged; this appears to be highly compatible with the general demand of an overall increase of competition *between* different universities. Therefore, similar processes at the »micro« and »macro« systemic level can be linked to each other, which again possibly creates certain synergy effects. (d) *Performance-based redistribution processes within universities can help to improve substantially the general public image of universities*. Redistribution mechanisms might be interpreted by the public as well as by public decision makers as a serious attempt by the universities to increase transparency and to introduce accountability and performance-favoring attitudes within the context of a university institution. Particularly during the current phase, in which the universities claim that they are underfinanced by the public[179], such a university policy – that of implementing a university-internal redistribution – could be used in the academic discourse and in public discussion as a key argument, with a crucial symbolic meaning, for the interests of universities. (e) The pivotal question, which remains, obviously is whether such a re-allocation process should be regarded primarily as *a setting of incentives* that can stimulate university research activities into a desired direction or, contrarily, whether *a fundamental and far-going redistribution of resources* should be the ultimate goal? Currently, as it appears to us, there is no consensus among experts on that important issue. While some experts qualify it as sufficient to use and interpret the university-internal re-allocation as a means for defining incentives, others are more inclined to favor the concept of a fundamental redistribution: thus, in the opinion of those experts, also the basic university funds should be addressed by the university-internal re-allocation decisions. Assessed in empirical terms, Uwe Schimank claims that up until now those university-internal redistribution processes are still far away from a substantial re-allocation of resources. Therefore, as a correct interpretation of the current *status quo,* this university-internal redistribution resembles more an attempt of implementing symbolic incentives. Uwe Schimank also gives a convincing empirical example: during the period of 1975 until 1990 only 5% of the academic positions were subject

[179] See again Chapter 3.1, in which we discuss this issue of an asserted »underfinancing« or »underfunding« of the German universities.

to redistribution decisions – calculated in average statistical terms this means one *per* 300 academic positions *per* year. Schimank explains this phenomenon of a *de facto* resistance against a radical redistribution with the fears of the academic staff that such a process might create tremendous internal turmoil and tensions (Schimank, 1995b, 65–66).

The Free University of Berlin (FU Berlin)[180] already has implemented a scheme that processes a university-internal redistribution of resources depending on performance. Since this scheme is qualified by many experts as interesting and innovative, we want to present and discuss, in the following, the key features of this »FU Berlin« model in more detail[181] – for an excellent summary and overview see also an article by Peter Wex, published in 1995 (Wex, 1995):

– The »FU Berlin« model – that means a university-internal redistribution of resources as an outcome or consequence of performance – was implemented back in 1992 and has been continuously operative. In 1992 the re-allocated money sum amounted to 1.6 million DM; in 1993 this sum increased to 2.4 million DM, and in 1994 to 2.5 million DM.

– To assess the financial »weight« of those 2.5 million DM, which were redistributed in 1994, two different references – and consequently two alternative interpretations – seem possible. First of all, those 2.5 million DM are a part of a volume of 23 million DM (in 1994), which are used for additional facility and non-labor expenditure (in German called *Sachmittel*). So of that money approximately 10% – in 1994, exactly 10.9% – are subject to redistribution. On the other hand, however, personnel or labor expenditure (*Personalmittel*) or the basic supply or basic funding (*Grundausstattung*) for the FU Berlin are not addressed by the re-allocation scheme. Therefore, secondly, when the total costs of the FU Berlin are taken as a frame of reference, we obtain a reverse picture: in 1994 the total costs amounted to approximately 1300 million DM; consequently, the redistributed 2.5 million DM represented only a share of 0.19% of the overall monetary supply for FU Berlin. So arriving at a, perhaps preliminary, bottom-line conclusion, it is probably fair to interpret and describe the current redistribution model of the FU Berlin primarily as a system that wants to set some incentives, and not so much as a system that aims at a »real« and fundamental re-allocation of resources.[182] Of course, there is an intense debate at FU Berlin on the future objectives of the redistribution scheme; while some experts prefer the *status*

---

[180] In German: *Freie Universität* Berlin.

[181] As one expert indicated, the »FU Berlin« model could be interpreted as a first and perhaps preliminary or somewhat pre-mature design and step towards a more complex and in-depth model for a *formula-based funding* of a whole university.

[182] This assessment is similar to the above cited claim of Uwe Schimank that currently the university-internal redistribution attempts in Germany are facing serious structural constraints (see again Schimank, 1995b).

qou, i.e. the development and set-up of incentives that should stimulate self-organizing processes, other analysts would opt for an expansion of the scheme's scope – that means that in the future also the labor costs and segments of the basic supply should be integrated, at least partially, into a re-allocation of resources that is closely linked to performance.

– The seventeen disciplines or faculties, or, as phrased preferably in German in reference to the university-internal institutional organization, the seventeen *Fachbereiche,* that are addressed by the redistribution scheme, are grouped into two main clusters. The one cluster includes the social sciences and humanities, the other cluster contains five science or natural sciences disciplines: those are physics (*Physik*), chemistry (*Chemie*), pharmacy (*Pharmazie*), biology (*Biologie*), and earth and related environmental sciences (*Geowissenschaften*). The underlying rationale for that is that the closer related disciplines (or *Fachbereiche*) are, the easier – and also methodologically »safer« – is an indicator-based comparison of performance of exactly those disciplines. Therefore, in practical policy terms the recommended implication would be to compare disciplines of the social sciences and humanities primarily with the social sciences and humanities, and science or natural science disciplines with disciplines in science or the natural sciences.[183] Comparisons across the fundamental disciplinary border of those two clusters, at a so-called »meta level«, are always vulnerable to severe criticism – however, they might be conceptually or intellectually challenging.

– For the performance assessment of each discipline or faculty (*Fachbereich*) five indicators are used. Those are, first of all, (a) **the number of scientific publications** during the last three years, and put in relation to the number of academic or scientific staff members *per* faculty. The only »authoritative« source for counting publications is the so-called *Universitätsbibliographie,* a book volume that is published each year – since the first half of the 1980s – by the university library of FU Berlin. In it all publications are documented that are reported by the academic or scientific university staff to the library, normally with a time lag of two years: so the 1995 edition covers the year 1993 (see, for example, FU Berlin, 1995). Concerning the structure of the book, for each *Fachbereich* (faculty) – down to the level of individual departments – and for each academic or scientific staff member, in alphabetical order, each reported publication is listed. This institution-oriented individual listing is then additionally supported by a person or author index (*Autorenregister*) at the end of the book. Experts indicate that since this *Universitätsbibliographie* is being used as a source

[183] See again paragraph number four, where we summarized, according to Hans-Dieter Daniel, the key principles for a proper use of indicators: there the hypothesis was put up for discussion that comparisons across disciplines are in such cases the most difficult where disciplines are very different in their structure and content.

for publication performance assessment, the frequency of publication reporting has increased considerably, so that currently it is fair to assume that the *Universitätsbibliographie* can claim a high degree of representiveness of the publication activities of the university staff. In the future it is planned to add the following weighting scheme to those individual publications, which should improve the current procedure of just simply counting the quantity of publications; single or multi-authored books (monographs): 10 points; journal articles, edited publications or contributions to an edited book: 4 points; and reviews: 1 point. (b) **Earmarked funds (*Drittmittel*):** The total expenditure of earmarked funds during the last three years *per* academic or scientific staff of each faculty (*Fachbereich*). (c) **Dissertations:** The total number of dissertations, during the preceding three-year period, again *per* the number of professors of a faculty (*Fachbereich*). (d) **Graduations:** The number of graduations during the last three years *per* the number of professors and multiplied with the *Curricular-Normwert* which indicates the supervision input or the supervision supply for each student during his or her study program. (e) **Students:** The number of students from the first to the eighth »subject-related semester« (*Fachsemester*), again multiplied with the *Curricular-Normwert,* and put in relation to the number of professors and half of the number of the other academic and scientific staff *per* faculty (*Fachbereich*).

- The actual performance assessment procedure then is that each discipline or faculty (*Fachbereich*) is ranked in relation to the other disciplines or faculties – within the context of one of the two major clusters that separate the social sciences and humanities (*Sozial- und Geisteswissenschaften*), on the one hand, from science or the natural sciences (*Naturwissenschaften*) on the other hand –, using a scale with a value range from 1 to 6. The best performing discipline gets a 1, and the discipline, performing the weakest, receives consequently a 6. In a first step each discipline or faculty (*Fachbereich*) is being ranked for each of the five indicators, as defined and described above; afterwards, in a second step, an *average ranking position* is calculated for each discipline or faculty (*Fachbereich*) on the basis of those five individual rankings. Those average ranking scores then represent the empirical data input into the formula that decides on the university-internal redistribution of those monetary resources that are at stake.

- The above described procedures of the »FU Berlin« model of a performance-depending and university-internal redistribution of resources focuses on the individual faculties (*Fachbereiche*) as the smallest unit of assessment. In other words: the redistribution is primarily a redistribution *between,* and not *within* the faculties. However, there is a growing general expectation and consequently pressure is being built up that possible monetary gains or losses at the level of the faculties be also passed on and distributed within the context of a faculty (*Fachbereich*) to the

individual departments or, even further, to the individual scientists in relation to their bottom-level performance. This exactly also indicates one of the currently most controversially debated discussions at FU Berlin, in the sense of *which institutional level represents the best frame of reference for redistribution decisions:* the faculties (as currently is the case), or the departments, or the academic individuals? As empirical investigations demonstrate, most of the financial surpluses were not redistributed within a faculty depending on individual performance, but, instead, were used for general or »conflict-avoiding« goals of a faculty; in practical terms, about 80% of the surplus money was transferred in a lump sum to the faculty library (*Fachbereichsbibliothek*). In reaction to that the *Fachbereich Wirtschaftswissenschaft* (faculty of the »economic sciences«) has developed a detailed and formalized scheme that – based on the distribution of performance points[184] – decides on how *within* the context of that faculty financial gains or losses again are redistributed; so this clearly resembles a model for a *combined inter-faculty and intra-faculty redistribution of resources oriented towards performance*.

(7) **Institutional reforms of the German university system to foster evaluations:** There is an intensive debate going on in Germany that focuses on the question of which *institutional reforms* are necessary or at least desirable for improving the quality and efficiency of the overall performance of the German universities. In that context, one of the key conclusions is that specific institutional structures can either favor or constrain the comprehensive use of evaluations.[185] There are powerful arguments that convincingly demonstrate the »systemic« reasons for why evaluations in general, but also why evaluations of university research in more particular, will gain a crucial importance in the future – see again our analysis in Chapter 3.2.4 –, so the challenge can exactly be summarized as: *to design and to implement institutional structures into a university environment that, firstly, promise a high compatibility with the demands of evaluations and, secondly, beyond that even foster the systematic application of ex-ante and ex-post evaluations*.[186] In that respect there are two sensitive issues, which we want to discuss in the following in more detail: Should professorships be granted on a permanently tenured or only a temporally limited basis? And: Are there still rational and scientifically-based arguments that justify the demand of a *Habilitation* for a professional academic career?

– In Germany a professorship is normally granted as a tenure, this means without a temporal limit. Some experts are willing to qualify this practice as a structural

---

[184] In German: *Zuweisung von Mitteln nach Leistungs- und Belastungskriterien*.

[185] In the Chapters 3.2.2 and 3.2.3 we presented an overview of the structural and cultural constraints against a comprehensive evaluation of university research in Germany.

[186] For a further and interesting literature reference, see the published summary of a symposium which was titled *Hochschulreform durch Leistungswettbewerb und Privatisierung?* and which was organized in December 1994 (Stifterverband für die Deutsche Wissenschaft, 1995).

constraint against evaluations, since such a system is not in a position to offer »strong« incentives to the individual professor to expose himself or herself to an evaluation procedure that is carried out by an external unit – particularly, when such an evaluation could lead potentially to negative consequences; for instance, a decrease in the supply of resources.[187] Therefore, a group of analysts is willing to express the opinion that in the future the employment status of German professors should be radically altered. In their view the employment contracts ought to be »privatized«, this means that professors should no longer be treated, in their contractual status, like »public employees« that, firstly, cannot be dismissed and, secondly, are permitted the privilege of an extensive autonomy that protects them against interference from the outside (the German term would be *unbefristetes Beamtendienstverhältnis*). Obviously, the current employment contracts of the German professors cannot be changed or reversed anymore. However, based on predictions it is legitimate to assume that the German university system is standing at the beginning of a replacement process of almost a whole generation of professors (*Generationswechsel in der Professorenschaft*): within the next fifteen years almost three fourths or 75% of the current professors will go into retirement (HRK, 1996a, 20–21). So this defines a »window«, through which the employment status of a German professorship can be changed substantially and comprehensively. Among those experts, who are in favor of such a change, the following two different schools of thought seem to be prevailing: (a) The one group emphasizes that tenured professorships ought to be the exception, particularly when he or she is appointed as a professor for the first time. Therefore, under normal standard conditions the first appointment as a professor should be temporally limited to a period of five to eight years. Towards the end of such a professional period, the performance of that professor would have to be systematically *ex-post* evaluated, and also his or her future plans should, in addition, be exposed to an *ex-ante* assessment. The prolongation of the employment contract, for a next functional period, then would depend primarily on the outcome and the results of the *ex-post* evaluation. In case of a non-prolongation, he or she should receive some financial compensation. (b) The other group of experts argues that the principle of a tenured professorship should be continued. They justify this basically by referring to the empirical fact that in Germany the average age of an academic, after completing a *Habilitation* – the formal prerequisite for applying for a professorship –, is about 39 years; and for persons at that age a temporally limited academic position would be too much of a risk. However, the additional supply of resources for a professor – which can be used for employing assistants, for example –, that normally accompany the assignment of a professorship (in German called *Berufungszusagen*), should be temporally limited. Therefore, in a worst case scenario, where an *ex-post* evaluation of the performance of

[187] See again paragraph number three in Chapter 3.2.2, where we raised, for the first time, this issue of tenured professorships in Germany and discussed their consequences on the application of evaluations.

a professor arrives at a negative conclusion, the professor still would receive his or her basic salary, but all the additional support and supply with university resources would be terminated.

– As already mentioned and briefly discussed in Chapter 2.1 the completion of a *Habilitation* – in English sometimes paraphrased as a »higher doctorate« (see again Irvine *et al.,* 1991, 52) – represents a necessary condition for applying for a professorship in Germany. The only commonly accepted possibility for an academic to by-pass this *Habilitation* demand is to enter an academic career outside of Germany and after succeeding in becoming a professor, to apply from »abroad«. In practical terms such a *Habilitation* procedure implies that the individual academic is expected to write a »thick book« – this is particularly the case in the social sciences and humanities – which afterwards must be defended *vis-à-vis* a commission that is recruited from the local university. Some leading experts are willing to criticize this *Habilitation* demand. The following complaints are brought forward: firstly, the compilation of a »thick book«, often equipped with a theoretical bias, can distract from the core process of conducting substantial academic research; in that context also the opinion is expressed that often the published *Habilitationen* do *not* belong to the most interesting or to the most frequently cited books. Secondly, by putting a prime emphasis on the *Habilitation* demand, this implies consequently that the catalogue of crucial criteria for an application is almost »artificially« narrowed down; therefore, some experts recommend redefining the *Habilitation* from a necessary to only one of several optional condition. This would give potential candidates for a professorship, who did not complete a *Habilitation,* the chance that they could compensate this deficit with achievements in other areas, such as: practical experience, international professional experience, a sufficient and impressive publication record, and so on. Thirdly, the *Habilitation* is seen by many as a »traditional« instrument for integrating and for conventionally or »conservatively« socializing young scientists at the beginning of their career into the academic community; therefore, according to Peter J. Brenner, the primary purpose of a *Habilitation* is not to promote creativity or to induce conceptual innovations, but to reproduce the mainstream body of knowledge that is being taught in a university context – as a result, *Habilitationen* are biased towards the »conservative« or structure-maintaining end of the spectrum.[188] Fourthly, *Habilitationen* are a crude

[188] Peter J. Brenner offers the following summary interpretation for the function of the *Habilitation* for the German academic science system: »Hier liegt aber auch schon die Grenze des wissenschaftlichen Ertrags der Habilitation: Sie führt fort, was im Fach schon angelegt ist; in aller Regel gehen innovative Impulse von ihr nicht aus. Die Habilitation als Verfahren ist so konzipiert, daß sie zur personellen wie sachlichen Konsolidierung etablierter Forschung beiträgt; die Förderung von Originalität und Innovation gehört nicht zu ihren Aufgaben. Sie ist vielmehr Instrument der Integration von Wissenschaftlern durch Internalisierung von Verhaltensformen. Zugleich dienen sie der Stabilisierung einer Disziplin durch Fortschreibung von methodischen Konventionen anhand von in der Regel neuen Materialien. Der Verzicht auf Originalität freilich ist kein Versagen der Institution Habilitation, sondern ihr Ziel« (Brenner, 1993, 344).

selection mechanism that should reduce the number of eligible candidates for announced or vacant professorship positions. As some experts emphasize, during the 1960s and the first half of the 1970s – a period of a massive expansion of the former West German university system – the *Habilitation* already was selectively *de facto* abolished as a formal prerequisite in some disciplines or university departments; however, later, when the tertiary educational output of qualified academics increased and, even more importantly, increased faster than the installation of new university positions, the comprehensive *Habilitation* demand was again »re-introduced«. This development exactly underscores the current dilemma and explains why the total or partial abolishment of the *Habilitation* resembles such a taboo in Germany: Why should those professors, who successfully passed the *Habilitation* filter or threshold, be interested in downgrading the value of the *Habilitation* and therefore making it easier for new applicants to penetrate the labor market of professorships? Thus even those experts, who would qualify it as rational and perfectly compatible with the intrinsic demands of the scientific and academic research system to »soften« the formal importance of the *Habilitation,* admit or have the impression that currently still a majority of the German professors wants to sustain the *Habilitation* demand. In other words: only a minority fraction among the German professors would favor or opt for an abrogation or »functional suppression« of the *Habilitation* prerequisite. Interestingly enough, there is no consensus among those *Habilitation*-critical experts whether the younger professors are more in favor, or not, of abolishing the *Habilitation*. Also the German Science Council put forward some question-raising comments in reference to the *Habilitation,* by emphasizing that from the turn of the century up until the end of the 1980s the average age of an academic, who completes a *Habilitation,* increased from 30 to 39 (Wissenschafsrat, 1988, 182). In addition, the Science Council criticizes several practical aspects of the current *Habilitation* procedure; however, when finally arriving at the concluding bottom-line, the Science Council does not, at least up until now, recommend abolishing the *Habilitation* requirement (Wissenschaftsrat, 1988, 189–190).[189] The *Habilitation* represents such a key institution in German academic university life, that its replacement by an alternative set of criteria would radically alter the German university system. But perhaps, as some experts indicate, this would also stimulate positive impulses.

[189] So a core message of the German Science Council is: »Der Wissenschaftsrat empfiehlt, an der Habilitation als Qualifikationsnachweis für die Professoren an Universitäten festzuhalten, soweit dies in den Fächern üblich ist. Die Universitäten sind jedoch gefordert, die Funktionsfähigkeit der Habilitation dauerhaft zu sichern. Dies kann nur gelingen, wenn der langjährige Trend zum höheren Habilitationsalter gebrochen wird und die qualifizierten Nachwuchswissenschaftler sich wieder in einem früheren Alter, vor Mitte dreißig, habilitieren« (Wissenschaftsrat, 1988, 189).

4. The Evaluation of Germany's University-Related Research: Discourse and Policy

At least some German experts are willing to set up the hypothesis that in the German context the university-related research[190] already has been more systematically evaluated or performs at a higher level of sophistication than university research (see, for instance, Krull, 1994, 206).[191] If this is true, then this would indicate a certain *unbalance or policy asymmetry* between the university and the university-related sectors or it would imply, to phrase it somewhat differently, that the diagnosed lack of evaluations seems to be primarily a problem in the case of the universities. Obviously, the reasons for such a situation would be manifold and complex, and we are certainly not in a position to offer now a thorough and in-depth analysis. However, we briefly would like to refer to the following two issues. First of all, university-related research institutes are oriented more clear-cut towards research and the career promotion of young scientists, whereas in the case of the universities the functional overlapping of teaching and research – concerning the performance output and the input of funding – might create certain ambiguities or some unclearness[192]; so, perhaps, it is easier for a university-related research institute to develop a »corporate identity« that, at the same time, defines a window through which a systematic application of evaluations is fostered. Secondly, and this identifies another major difference in contrast to the universities, the funding base of university-related research is more vulnerable; mostly they rely on a dual public funding mode where the federal government (the *Bund*) and the *Länder* ministries divide the costs between each other according to a specific key of burden distribution (BMBF, 1996b, 23). So this potentially »endangered« position leads to a situation in which the university-related research institutes must permanently convince – as a general message – their funders and the public of their importance for society and of the quality and relevance of the research they carry out. Of course, the university-related research institutes do not represent a homogenous sector, but a sector, which again in itself can be characterized by a complex diversity. However, this does not imply that certain features cannot be defined which function as common characteristics when comparing the

---

[190] As already suggested in Chapter 1, we will circumscribe throughout our analysis the German term and concept of an *außeruniversitäre Forschung* in English with »university-related research«. See again the discussion in Chapter 1, in which we defend our suggested terminology.

[191] »Mit Blick auf die Hochschulen gestaltet sich die Tätigkeit des Wissenschaftsrates ungleich schwieriger als im außeruniversitären Bereich; denn während letzterer in der Bundesrepublik – sowohl bei den deutschen Akteuren selbst als auch bei vielen ausländischen Betrachtern – als in vieler Hinsicht vorbildlich strukturiert gilt (die arbeitsteilige Vielfalt, die effiziente Wahrnehmung der Teilfunktionen, die Fähigkeit, komplementäre Effekte zu erzielen etc.), wird die westdeutsche Massenuniversität seit geraumer Zeit als ein schwerkranker Patient oder gar als ›Achillesferse der Bundesrepublik Deutschland‹ bezeichnet« (Krull, 1994, 206). Also Stefan Kuhlmann, a German expert of the Fraunhofer Institute *Systemtechnik und Innovationsforschung* (ISI) in Karlsruhe, arrives at the estimation that currently the research, which is performed by the university-related research sector, is more systematically evaluated than the university research (see Kuhlmann, 1997).

[192] Although many experts would qualify this functional duality within a university environment as essential. This means that a university could not survive or sustain its legitimation for society without simultaneously providing research and teaching.

university-related research institutes with the university sector that also reveals an intra-sectoral diversity (for a broad overview of the German university-related research sector, see; Massow, 1986; Hohn and Schimank, 1990; and BMBF, 1996b).

To illustrate and support this hypothesis, that in the German context the university-related research cluster has advanced and matured to a further degree than the university system in regard to the application of evaluations, we briefly want to describe those internal evaluation procedures that are processed as standard routines by the research institutes of the Max Planck Society (MPG). The Max Planck Society was newly founded, in 1948, as the successor organization to the Kaiser-Wilhelm-Gesellschaft which already was established in the year 1911. In 1995 the Max Planck Society consisted of approximately one-hundred research institutes and a total staff of almost 12000 members and consequently occupies a key role within Germany's university-related research sector (BMBF, 1996b, 409–410). According to an analysis by Wilhelm Krull, the MPG employs the following five internal standard procedures for self-evaluating the performance of MPG member institutes (see Krull, 1995, 442):

- controlling procedures by internal auditing units;
- *ex-ante* evaluation procedures of persons as well as concepts by panels of experts;
- assessments of institutes on a regular basis by scientific advisory boards;
- the preparation and implementation of restructuring procedures of MPG institutes by presidential committees;
- and assessment procedures by state and federal auditors.

Concerning those different self-evaluating procedures at MPG, the interesting trend can be observed – according to the analysis of Wilhelm Krull – that quantitative indicators start to play an increasingly important role. Such quantitative indicators reflect data such as publication output, citation frequency, prize awards, the intensity of acquired earmarked funds (*Drittmittel*), budgetary plans, and so on. Of course, these indicators are not used as a single or isolated source of information. However, they are of a crucial complementary value for the traditional quality assessment procedures that are carried out by the expert panels. Or in other words: those experts panels are increasingly inclined to refer and to incorporate, in addition to the »traditional« methods, the information value of those quantitative indicators for the purpose of arriving at a general conclusion on the quality and efficiency of individual MPG member institutes (see Krull, 1995, 449).[193]

[193] For us such a diagnosis and observation is interesting, since it supports our conclusions in Chapter 3.2.5 (in paragraph number 2), where we discussed the importance of general indicators, as well as quantitative-oriented indicators, for the process of evaluating university research.

Of a tremendous and crucial importance for the whole evaluation process of academic research in Germany was the unification of former West (BRD) and East Germany (DDR) to a new state, and ultimately to a new nation, in the year 1990 (Mayntz, 1994, 18). One major challenge arising out of this unification was: What should happen to the East German science and research system? Reflecting the power distribution of real politics (*Realpolitik*), finally, what resulted was the basic decision to incorporate the East German research institutions into the West German research system, which then – by definition of real world politics – represented the structures for the research system of unified Germany. Andreas Stucke interpreted this process as a transfer of institutions (*Institutionentransfer*) from the West to the East.[194] Concerning the former university-related research cluster in East Germany, which consisted primarily of Academy institutes, two fundamental decisions were processed. Firstly, those Academy institutes should not be sustained, but »dissolved« and re-integrated into the structural framework and the institutional design of the West German academic research system (Meske, 1993, 17–18). Secondly, this re-integration process should be preceeded by an in-depth evaluation of those former Academy institutes, which, as a final outcome, would decide which institutes or which staff members are of a scientific value for the unified German academic research system and thus should continue to carry out research. The »West German« Science Council was officially asked by the West German and East German governments, in July 1990[195], to carry out this tremendous task of evaluating and of putting forward recommendations for the university-related research in the DDR: in practical policy terms this implied that all research institutes of the Academy of Sciences (*Akademie der Wissenschaften*) – often abbreviated as AdW-*Insitute* or AdW institutes –, the Academy of Constructing (*Bauakademie*), and the Academy of Agriculture (*Akademie der Landwirtschaften*) were exposed to a thorough investigation (Maurer, 1996, 9). For that purpose the Science Council set up nine working groups (*Arbeitsgruppen*), in which, all together, more than 300 experts (*Sachverständige*) participated. The major part of that evaluation procedure, and the main conclusions, were finally completed by the Science Council in July 1991 (Wissenschaftsrat, 1992a, 7).

Concerning those evaluations of the East German Academy institutes, experts like to emphasize the following issues:

(1) It was not just a coincidence that the Cologne-based Science Council was assigned with that major exercise and task. Since its establishment in the year 1957, the Science Council already had individually evaluated almost 60 research institutes by the end of the 1980s – primarily institutes of the so-called »Blue List« (in German

---

[194] »Wie läßt sich die Genese der institutionellen Grundlagen für die deutsche Vereinigung im Wissenschaftsbereich erklären, d. h., wie kam es – letztlich – zum Artikel 38 Einigungsvertrag, der die formale Basis für die Umsetzung der Einheit im Wissenschaftsbereich darstellt und im Kern einen Institutionentransfer von West nach Ost vorsieht?« (Stucke, 1992, 3).

[195] Unification between the BRD and DDR was finally legally and politically processed on October 3, 1990.

Blaue-Liste-Institute or *Blaue Liste-Einrichtungen*).[196] So the Science Council had acquired a professional expertise in assessing the performance of research institutes, and many analysts and public decision makers labeled the Science Council as the only public or semi-public agency that could carry out such an initiative. At the same time, however, this evaluation exercise also present the Science Council with a completely new challenge: whereas those earlier assessments focused on individual institutes, the evaluations of individual Academy institutes had to be put and integrated into a comprehensive performance evaluation of a whole research sector.[197] Therefore, as an essential by-product of that exercise, the Science Council could significantly improve its evaluation expertise (Maurer, 1996, 1, 3–4, 9–10).[198]

(2) Some experts regret that during this German-German unification process only the university-related research sector of former East Germany was evaluated, *i.e.,* the Academy research institutes, whereas the East German universities were not exposed to such an in-depth assessment of quality and efficiency.[199] In the opinion of those experts this represents a missed opportunity. The reasons for this divergent development, obviously, are manifold: on the one hand, already before 1990, the attention of the Science Council traditionally focused on the university-related research sector which is binary funded by the *Bund* (federal government) and *Länder* ministries; since universities – also those of the newly established *Länder* in former East Germany – are placed within the primary competence realm of *Länder* ministries, this resembled an efficient structural constraint against possible evaluation attempts of the Science Council. In addition, a few experts also express the opinion that the existing resources and means of the Science Council would have to be

---

[196] The Blue List institutes (BLIs) represent a heterogeneous group within Germany's university-related research sector that covers almost the whole disciplinary spectrum within the sciences. At the beginning of 1996, no less than 83 BLIs existed, with a total staff size of approximately 10000 persons. Historically the BLIs originated from the *Königsteiner Institute* and were formally established in the 1970s. BLIs are co-funded on a fifty-fifty percent basis, *i.e.,* by equal shares, from the *Bund* (federal government) and the *Länder* (provincial governments). Membership of research institutes in the Blue List is not necessarily unlimited. Theoretically, and also practically, research institutes can be excluded or included into the Blue List and its generous public funding scheme, whereas the total number of Blue List member institutes – at one time – is thought to be more or less pre-defined or non-variable. This also explains why in the past most of the institute-oriented evaluations, which were carried out by the Science Council, focused on the BLIs: there is a permanent structural demand for assessing the legitimation for a *Bund-Länder* co-funding of individual BLIs, which implies that their importance must be cross-regional (*überregional*) or even national and that their research performance is of a general interest for the society and economy (see Hohn and Schimank, 1990, 135–170; BMBF, 1996b, 23, 457–487).

[197] In that respect Michael Maurer argues: »Zum anderen ist bei den Bewertungen in Ostdeutschland deutlich geworden, daß über die Bewertung einzelner Institute hinaus eine zweite Ebene erschlossen werden kann, die sich aus der Integration von Einzelbewertungen zu einem Gesamtbild ergibt« (Maurer, 1996, 10).

[198] So some experts also express the opinion that in case that a comprehensive output and quality evaluation of Germany's university research performance should be carried out at the national level, then in such a scenario the Science Council should play a crucial role in operating and guiding such a procedure.

[199] Wilhelm Krull, for example, stresses such a hypothesis; see again Krull, 1994, 205–206.

expanded so that the Science Council could undertake such an exercise of evaluating comprehensively the whole university sector with comfort.

(3) Now analyzing the direct effects those evaluations of the Academy research institutes had for the post-East German science and academic research system, then to many experts the following two consequences appear crucial: firstly, and most dramatically, a massive reduction of university-related research potential occurred. According to an estimation by Werner Meske, when the year 1989 serves as temporal reference point, then, in 1992, only one third of the university-related research personnel still was active in R&D (Meske, 1993, 27). Since the former East German university-related sector more or less coincided with the cluster of the Academy research institutes (Meske, 1993, 17), this could lead to the conclusion that the evaluation-accompanied integration of the East German university-related research sector into the »all«-German research system had a »high price« for the former East German research community: on the average the integration addressed only one out of three researchers (see also Meske, 1993, 29–33).[200] Secondly, many of those former East German Academy research institutes, as a whole or in segments, were transformed into newly established institutes of the so-called Blue List. If the Blue List in former West Germany at the end of the 1980s contained approximately 50 member institutes, their number almost doubled by increasing to 83 BLIs in 1996 (BMBF, 1996b, 457; Maurer, 1996, 9). This increase only occurred because BLIs were founded in former East Germany. Therefore, speaking in the context of the all-German academic institutional context, one could set up the hypothesis that the Blue List was the major »institutional winner« in Germany's university-related research sector. This significant expansion of the Blue List also redefined some of the well-established ratios and distributions – and perhaps also some of the allocation keys of R&D-funding resources – between different institutions or research organizations of the »old« West German university-related research system. According to the analysis of Werner Meske, now the Max Planck Society perceives the Blue List as a serious competitor (Meske, 1993, 22).[201]

In addition, those Science Council-based evaluations of the East German Academy research institutes imposed the following consequences or provoked the following serious questions for the national academic research system of now unified Germany:

---

[200] For the university and the business enterprise sectors we also obtain a very similar ratio (see again Meske, 1993, 27).

[201] For further general information on this evaluation exercise and its consequences, see: Gläser, 1992; Gläser *et al.,* 1995.

- Firstly, it was admitted for a whole research sector that research output, research quality, and research efficiency are in principle measurable.[202]

- Secondly, if the East German academic research system was subjected to a thorough investigation, then how can it be justified not to evaluate comprehensively the former West German academic research system?

- Thirdly, if it is acceptable to assess the university-related research sector, so why is it not possible to carry out a comprehensive research evaluation of the university system at the national level?

As already elaborated, one major consequence of the evaluation of the East German Academy research institutes was that the number of Blue List institutes increased significantly. Consequently almost over night, unified Germany had a »new« and very strong research organization, located in the university-related sector, which also is perceived by the other »traditional« and well-established institutions, such as the Max Planck Society and the Fraunhofer Society, as a serious competitor (see again Meske, 1993, 21–22). This growing and increased self-confidence of the Blue List also manifests itself in the strategy to redefine its public image: whereas in the past the individual Blue List institutes cooperated only on a voluntary basis, there are now serious initiatives being undertaken for a more focused coordination of the activities of the different Blue List institutes. As an ultimate goal, this could imply designing and developing a »corporate identity« for the Blue List which is comparable to that of Max Planck or Fraunhofer Societies.[203] A first initiative, in November 1991, in which a working group (*Arbeitsgemeinschaft*) between the Blue List institutes was launched, was again emphasized and reinforced in March 1995, by establishing the Science Community Blue List, abbreviated as WBL (*Wissenschaftsgemeinschaft Blaue Liste*). The WBL has a permanent administrative office, which is funded on the basis of membership contributions of individual Blue List institutes, and which seeks to coordinate more effectively the activities of the Blue List institutes. Furthermore, the WBL attempts to communicate the common interests of the Blue List *vis-à-vis* the public, and public and private decision makers. Most Blue List institutes are members of the WBL (BMBF, 1996b, 457).

Also in reaction to this growing influence of the Blue List institution, the Science Council published, in 1993, a series of recommendations that addressed (and still address) the future performance of the Blue List institutes. In the following we will summarize the most important recommendations (Wissenschaftsrat, 1993c, 38–40; see also BMBF, 1996b, 457):

---

[202] See again our analysis in Chapter 3.2.4 (see also Weingart, 1995, 74).

[203] For an overview of the history of the Max Planck and Fraunhofer Societies, see: Massow, 1986b, 32–34, 34–35.

(1) Each individual Blue List institute should install scientific advisory boards (*Wissenschaftliche Beiräte*): the purpose of such boards is to support the performance of the institutes by regularly assessing, in parallel to the institute's activities, the research output and the service offer.

(2) To encourage the staff flexibility of the institutes, between 30% and 50% of the scientific personnel – depending, obviously, on the specific tasks – should receive only temporary (temporally limited) contracts of employment.

(3) The collaboration of the Blue List institutes with the universities and *Fachhochschulen* should be intensified. To achieve this goal, a particular emphasis should be intentionally put on the appointment of scientists and researchers who also perform active duties at universities; for instance professors, who then are expected, in parallel to their Blue List employment, to teach at universities, supervise Master's and Doctoral theses, and who promote the academic career of talented young scientists.

(4) Concerning the pivotal question of individual institute membership in the Blue List, with the main consequence of accessing to and benefiting from the public funding scheme for the Blue List, the Science Council put forward a very clear message: with respect to the inclusion and acceptance as well as the exclusion of research institutes in the Blue List, there should be more flexibility. This, simultaneously, implies that the chances for admission of a high-quality research institute increase significantly; on the other hand, however, the risk for an institute of again being excluded also would increase. So, all together, institute-oriented membership mobility and flexibility of the Blue List institution will advance. To have a sophisticated basis for decision-making on that crucial issue, the Science Council demands that each Blue List institute should be evaluated externally, and regularly, once during a time period of four to five years. This task will (and should) be carried out by the Science Council (see also Wissenschaftsrat, 1993c, 33–35).[204]

(5) Beyond this principal mobility demand on the issue of membership to the Blue List, also within the framework of the Blue List there should be a greater flexibility concerning, for instance, the allocation of resources. This means that a redistribution of resources or funds between different Blue List institutes should be regarded as an option which should be applied. Those external evaluations could supply the data and rational basis for such a decision-making.

---

[204] As the Science Council comments on that issue: »Geht man von einer regelmäßigen Evaluation der Institute in etwa fünfjährigem Abstand aus, so müßten durchschnittlich knapp 20 Evaluationsberichte pro Jahr erstellt erden« (Wissenschaftsrat, 1993c, 33).

(6) For the purpose of improving the cooperation between the different institutes and, at the same time, also to support those necessary external evaluations, the Blue List institutes should organize and self-organize themselves more determinedly. Therefore, the establishment of an *Arbeitsgemeinschaft,* in 1991, and of the WBL in 1995 – as already described above – represent a direct consequence of that recommendation.

Concerning the methodology, how the Science Council evaluates individual institutes, normally the following standard procedure is applied – in our description we will follow an analysis which is presented by Michael Maurer (see Maurer, 1996, 5–6, 17):

(1) First of all, at the beginning, the Science Council would establish a working group (*Arbeitsgruppe*) that is primarily responsible for carrying out the specific evaluation task of an institute. Essential for the composition of those working groups is the participation of external experts who are invited to join in.

(2) In a second phase, a questionnaire is prepared and sent to the institute. The questionnaire is very detailed and attempts to cover the whole functional profile of that institute thoroughly. Normally the institute would be granted a time period of six to eight weeks for responding, answering, and returning the questionnaire. In addition, it is expected that the institute provides detailed information on issues, for example, such as: annual reports, the statutes and articles (*Satzung*), the budget, publication records, figures on the staff, and acquired earmarked funds (*Drittmittel*).

(3) After a first assessment and analysis of the information returned by the institute, including the questionnaire, the working group will carry out a direct local visit (*Ortsbesuch*) to the institute that normally will last one or two days. During such a visit, the following activities are conducted:

- talks with the director of the institute;
- talks with the department chairs (without the director);
- talks with members of the academic or scientific staff (without the director or the department chairs);
- a walk or tour through the institute and talks with institute members at their place of work;
- talks with external partners of cooperation (for example, local universities and local users).

(4) Afterwards, the working group compiles a report. At that point there is an interest in the working group to achieve unanimous consensus among the members. Then this report, and particularly its recommendations, are forwarded to the two main sections

of the Science Council, the *Wissenschaftliche Kommission* (Scientific Commission), consisting primarily of scientists, and the *Verwaltungskommission* (Administrative Commission) that is setup of politically appointed members by the federal government (*Bund*) and the *Länder* governments: in that phase, the recommendations (*Empfehlungen*) of the working group might again be modified and changed. In the *Vollversammlung* (Plenary Assembly), the unifying body of those two commissions, an agreement is finally decided on the report that now contains the »official« recommendations of the Science Council.

Concerning such institute-oriented recommendations of the Science Council, they mostly are very detailed and cover a whole spectrum of issues. In the following we present a short overview of possible recommendations (see again Maurer, 1996):

(1) **Funding:** In the case of an evaluation of a Blue List institute – which normally was the case in the past (Maurer, 1996, 3–4) – it is of crucial importance, whether the Science Council recommends that the binary co-funding (*gemeinsame Förderung*) of *Bund* (federal government) and *Länder* (provincial governments) should be continued; because, in practical policy terms, the termination of this public co-funding implies that the very existence of that institute is then at stake. Normally, the Science Council recommends that the *Bund/Länder* co-funding for that evaluated institute should be continued. But such a general recommendation is mostly linked to additional recommendations, which indicate how the performance of the institute can be improved in the future. And since there always is the possibility of a termination of the public co-funding, the institutes take those additional recommendations very seriously. Until 1995 there were only three cases in which the Science Council directly put forward the explicit message of stopping the co-funding of an institute. Those institutes were:

- *Forschungsinstitut für Rationalisierung* [Research Institute for Rationalization], Aachen (in the year 1982);

- *Gesellschaft für Information und Dokumentation* [Society for Information and Documentation], Frankfurt (in the year 1984);

- *Institut für Erdöl- und Erdgasforschung* (IfE) [Institute for Petroleum and Natural Gas Research], Clausthal (in the year 1995).

(2) **Additional recommendations:** Those additional recommendations put forward by the Science Council, which go beyond this central and pivotal question of continuing (or discontinuing) the public funding base for the evaluated institute, cover a whole spectrum of issues. Addressing the prime funders *Bund* and *Länder,* the features

commented can be: the equipment of an institute, its staff structures, or the necessity for advisory boards. Concerning the institute more directly, the Science Council's suggestions can discuss aspects such as: the research topics, the applied methodology, the ratio of tenured and temporally limited employment contracts, and the publication profile of the academic staff members. Concerning the cooperation of the institute with external partners, possible recommendations can touch on the following questions: personnel and functional overlapping with local universities; common use of equipment and common research projects with other institutes; collaboration with industry or other commercial businesses.

In April 1994 the German Science Council was asked by the BLK (Federal-*Länder* Commission)[205] to evaluate – beginning in 1995 – within five years all institutes of the Blue List. Two reasons were crucial that particularly the Science Council was asked to carry out such a mission: firstly, up until the early 1990s the Science Council had already evaluated in former West Germany about sixty individual institutes, most of them belonging to the Blue List (Maurer, 1996, 3–4); and, secondly, the comprehensive evaluation of the East German Academy institutes in the years 1990 and 1991 implied, as a consequence, that the Science Council's expertise and know-how in the application of evaluations increased dramatically and significantly. Now concerning the comprehensive and comparative evaluation exercise of all Blue List institutes the Science Council installed, in May 1995, an *Ausschuß Blaue Liste* or, as phrased in English, a Permanent Panel »Blue List« that coordinates and carries out the individual institute evaluations. Those general recommendations on the Blue List, which were published by the Science Council in 1993 – and which we already discussed in detail above[206] –, served as a conceptual framework or as a masterplan that helped in structuring and designing the currently conducted evaluations. The practically conducted evaluation procedure is similar to those earlier and pre-1990s individual institute assessments; however, there are some minor, but not unimportant differences (see, again, Wissenschaftsrat, 1993c, and Maurer, 1996):

– The Permanent Panel implements or sets up for each Blue List institute a locally oriented *ad hoc* working group (*ad hoc-Arbeitsgruppe*), in which also foreign experts should participate. Those *ad hoc* working groups then compile evaluation reports in reference to those individual Blue List institutes.

– In addition, the Science Council designed and developed a comprehensive six-page questionnaire, which is sent to all Blue List institutes and to which the institutes also must respond. The purpose of the questionnaire is to receive broad and detailed information on the structure, performance, and future plans of each institute. This

[205] In German the BLK is called *Bund-Länder-Kommission*. For further information on the BLK, see: Massow, 1986b, 46–49.
[206] See again, in the bibliography: Wissenschaftsrat, 1993c.

information should help in creating an accurate picture of the general state-of-the-art of the Blue List research institutions.

– One crucial difference to earlier individual institute evaluations is that now the evaluation reports, and their conclusions, of the *ad hoc* working groups cannot be changed or altered anymore by the two Commissions or the Plenary Assembly of the Science Council.[207] Only the formulation of the recommendations is still the prime responsibility of those three bodies of the Science Council; however, the unchanged evaluation reports of the working groups must be simultaneously published together with the Science Council's recommendations. This new practice and policy should, already in advance, prevent more effectively that the evaluation exercise might be politically biased as a result of the influence of certain interest groups or lobbyists – or, as Maurer phrases it, consequently certain policy or recommendation-paralyzing effects of the so-called *Politikverflechtungsfalle* should be avoided (see Maurer, 1996, 10).[208]

– Another major difference is that in the past, in only three cases, the Science Council explicitly recommended that the public co-funding by *Bund* and *Länder* should be terminated (see again our documentation above). This implies that the Science Council was very cautious in putting forward such a harsh recommendation. This situation, however, seems to have changed dramatically concerning the currently conducted evaluation exercise of the Blue List institutes. Now, as it appears, the Blue List institutes evaluations are much more consequence-driven since the Science Council lost its reluctance of recommending the cancellation of public funding for an individual institute. Therefore, those evaluations also must be taken more seriously by the particular research community that is subjected to an evaluation. Through such a policy the Science Council emphasizes consciously its previously formulated recommendation that there is a demand for more flexibility and dynamics concerning the crucial issue of acceptance of »new« institutes into or the exclusion of »old« institutes from the Blue List funding scheme (Wissenschaftsrat, 1993c, 39). To illustrate our statement, we want to mention and refer to the following four examples: whereas, in the context of the current Blue List evaluation, the Science Council recommended that in the case of the *Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie* (HPI), *Institut für Zeitgeschichte* (IfZ), and *Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel* (IPN) the public co-funding should be continued, it arrived in the case of the HWWA-*Institut für Wirtschaftsforschung Hamburg* and *Niedersächsisches Landesamt für*

---

[207] In German: *Wissenschaftliche Kommission, Verwaltungskommission,* and *Vollversammlung*.

[208] In Chapter 3.2.2, in paragraph number 5, we already mentioned this German term of a *Politikverflechtungsfalle* – which is said to have been invented by Fritz W. Scharpf – for the first time and briefly discussed its meaning.

Bodenforschung – Geowissenschaftliche Gemeinschaftsaufgaben Hannover (NLfB-GGA) at the totally opposite conclusion, by recommending that, concerning those two Blue List institutes, the public co-funding should be canceled (see Wissenschaftsrat: 1996b, 1996c, 1996d, 1996e, and 1996f).

– In parallel to this »external« evaluation exercise conducted by the Science Council, several Blue List institutes are also in the process of carrying out »internal« self-evaluations. The WZB in Berlin, for instance, is a prominent example for such an institutional self-evaluation.[209] Partly those initiatives can be understood as a consequence of the 1993-recommendations, put forward by the Science Council (see Wissenschaftsrat, 1993c), in which it was suggested that each Blue List institute should install scientific advisory boards (*Wissenschaftliche Beiräte*) that regularly monitor the performance of the institute. However, beyond that, those institutional self-evaluations should furthermore be interpreted as a new mode or a new paradigm that influences and shapes the strategic behavior of research institutes. The crucial argument would be that as a prerequisite for employing adequate research strategies, an institute must develop »self-reflecting« means, which again implies implementing within the organization some »domestic« self-evaluating procedures: in practice this could mean asking for external peer-reviews or designing a set of indicators that represents the institute's performance, quality, and efficiency in research. Only through such an internal mechanism – so the argument – can an institute adequately self-assess its past performance, build up learning capabilities, and seek future-oriented strategies that optimally place the research institute within the context of an increasingly competitive environment.

Arriving at a bottom-line conclusion, many experts are willing to assert that the past evaluation of the East German Academy institutes and the currently conducted comprehensive evaluation of the Blue List institutes represent an important watershed for Germany's evaluation policy of academic research. In that context the two following arguments are often put forward: those evaluation exercises were important, firstly, for developing an expertise in evaluation policy and, secondly, in altering the attitudes of the German academic community and of the German public towards evaluations *per se*. However, as many experts stress, this should only be the beginning of a much longer and much deeper-going policy process or even of a policy evolution. Those experts demand that in the future also other areas or sectors of Germany's academic research system should be subjected to a thorough evaluation. Finally, as an ultimate consequence, this would imply that also the output and quality of Germany's university research would have to be evaluated comprehensively at the national level, covering all of Germany. In the opinion of many leading experts those evaluations are qualified as crucial and as a key tool for improving the quality standards and the global competitiveness of Germany's academic

[209] WZB is the abbreviation for *Wissenschaftszentrum Berlin,* which means in English: Science Center Berlin.

research performance, which again is crucial for maintaining the general wealth of German society and the competitiveness of the German economy.

5. Summary

In this country study the patterns and policies of *the evaluation of academic research* in Germany are analyzed. Our report is structured into three main sections:

- In Chapters 2.1 and 2.2 we investigate patterns of publication output and again comment shortly on Germany's publication efficiency. We will refer to only one form of publication output, that consists of articles which are published in international journals as covered by SCI (Science Citation Index) and SSCI (Social Sciences Citation Index). Our interest will be to put Germany's publication profile into relation with other OECD countries.

- In Chapter 3, our main section, our attention focuses on university (or higher education sector) R&D (research and experimental development) – the so-called »Hochschulforschung«. We will analyze the contemporary discourse and different opinions on how university research should be (or should not be) evaluated, what the practical experiences are, and, beyond that, which strategic scenarios should be developed for the future. In Chapter 3.1 we present an overview of the contemporary general trends in Germany's higher education sector. In Chapter 3.2.1 the current situation, concerning the evaluation of university research, is reviewed. In the Chapters 3.2.2 and 3.2.3 we discuss the structural and cultural constraints that, up until now, prevented a more comprehensive application of evaluations of university research; in Chapter 3.2.4, however, we summarize those arguments that stress why also in Germany the evaluation of university research will become more important in the future. In Chapter 3.2.5, finally, we give an overview of those evaluation initiatives of university research which are currently carried out in Germany.

- In Chapter 4, we summarize those evaluation procedures that focus on Germany's university-related research cluster – called in German the »außeruniversitäre Forschung«. A particular emphasis will be placed on the current evaluation exercise of the »Blue List« institutes, which also involves the development of a more generic masterplan which, in principle, also could be applied to other university-related institutes (in other Central European countries).

Section I: Bibliometric Analysis of Germany's Academic Publication Output and Publication Efficiency

One possibility to measure and consequently to evaluate the output and efficiency of a national academic research system, is to use a bibliometrics-based analysis that refers to articles in journals which are covered by SCI and SSCI. However, particularly in the German context often criticism is raised against this methodological approach, by formulating

arguments such as: (a) research quality *per se* cannot be measured and cannot be quantified; (b) there are different types of knowledge representation, and article-based knowledge is biased towards a fragmented understanding of our world; (c) and the theory of the specific and historic growth of academic cultures, proposing that while Anglo-American scholars prefer to write articles for journals, German (and perhaps also other Continental European) scholars developed an inclination for publishing books – the German tradition of *Habilitation*[210] can be used as an additional reference to reinforce such a statement. Notwithstanding that such criticism must be taken very seriously, we agree with other leading experts who emphasize the value of bibliometric analyses; particularly, when the specific strengths of such an approach are used sensitively and adequately. The following arguments we consider as key arguments that speak in favor of a bibliometric approach which focuses on articles in international journals:

– The decision of inclusion or exclusion (or acceptance or rejection) of forwarded article manuscripts by those journals, which are listed in SCI and SSCI, is generally based on a peer-review system. This comes close to something like a »built-in quality threshold« that guarantees that all published articles reveal a minimum and comparable basic standard of quality. Therefore, it is legitimate to interpret the number of articles in such journals as an indicator for or approximation of *quantified quality,* what again demonstrates the possibility of measuring quality – and falsifies the assertion that quality *per se* is beyond the scope of measurement.

– At the level of individual authors, obviously, one can imagine that a great variance of quality exists with regard to different articles. However, at the aggregated level of those different national academic research systems probably some phenomena of a *normal distribution of quality* will come into effect. If this is the case (which we and other experts believe), then this consequently implies that the variance between the aggregated median quality of articles of different national academic research systems is smaller – particularly when countries reveal a similar economic performance and similar socioeconomic standards, which is the case for most OECD nations – than the median publication quality of authors at the level of the individual researcher. Therefore, the frequency counting of articles makes more sense or appears more valid when conclusions should be drawn about the publication quality of different national research systems; at least more valid than a frequency-counting based comparison and assessment of the performance of individual researchers.

– Certainly, there are very different types of knowledge representation. Regarding the publication markets more directly, the two perhaps most classical but also most diverse examples would be books on the one hand (no matter if they are »thick« or

[210] The German term and concept of a *Habilitation* is sometimes translated into English as »higher doctorate« (see Irvine *et al.,* 1991, 52).

»thin«), and, on the other hand, articles in journals. Each of those different publication media has, of course, its very distinct profile, which implies that it can refer to a set of functional strengths and weaknesses (or advantages and disadvantages). Especially this relative view is important, since it means that every observer must realize that no particular publication medium offers only advantages or only disadvantages – when, for instance, considering the competition between books and articles. For us the crucial point is the following: the overall performance and competitiveness of a national academic research system is defined by the extent to which it is in a position to develop very different types of knowledge representation and its capability to operate along a wide spectrum of diverse publication means.

– Books, obviously, demonstrate certain strengths. Articles, on the other hand, also have their advantages – particularly when they are published in international and reviewed journals – which make them a powerful tool for storing and presenting information. In that respect the following characteristics of articles should be valued highly: first of all, the limited page space of an article forces the author to focus or to concentrate his arguments. One could say that those constraints on page resources encourage innovativeness and information efficiency. An important side-effect of this is that the chances that a »short« article is read and cited are probably higher than in the case of »thick« books; also the time for reading has developed into an ultimately scarce resource of scientists. Secondly, since a short (»few-page«) article can be written faster than a long (»many-page«) book, articles often demonstrate a temporal lead of competitiveness when compared with books (although the tedious review process of article manuscripts can often create substantial time lags). Thirdly, the data bases and retrieval systems which administrate articles have developed contemporarily a higher degree of sophistication than the computer-based documentation of books – for instance, article-oriented data bases contain information on the vocational address of authors and mostly also provide article abstracts. This, by itself, has encouraged a trend in which scientists are inclined to publish articles in such journals, since this guarantees a higher visibility of their research results (so this would be an excellent example for how certain »structural advantages« can lead to an increase of the quality of the »content«). Fourthly, and finally, we also must keep in mind that scientists write not just one, but normally a whole series of articles. This prevents an exaggerated fragmentation of knowledge and helps crafting a more holistic image or a more comprehensive scientific understanding of the world. In that context we also invented the term of publication-cluster strategies, which are applied by many scientists who publish. This means that many authors would place at the center of such a self-created publication cluster one or two (perhaps »thick«) books, which are reinforced in the periphery by a series of articles that demonstrate a strong referential linkage with those so-called »center« books.

– In reference to the *cultural argument* which claims an Anglo-American or English-speaking dominance in those international journals (as represented by SCI and SSCI), two arguments appear particularly crucial to us: firstly, patterns of cultural hegemony always existed in human history, also in sciences – at this point one could recall the pivotal role of Latin during the medieval and early modern period. So, consequently, when a country or a national academic research system does not belong to the *cultural core* of global sciences, then it must develop strategies of how to cope adequately with such a situation. In that respect Germany (or any other country) cannot demand an exceptional treatment. And, secondly, as a proper frame of reference – for the purpose of comparison – we always use for Germany the geographical concept of *non-English speaking Western Europe:* all together fifteen countries – Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland –, covering all of Northern and Western Continental Europe. Since those countries are all defined by the same cultural criterion that they are *not English-speaking,* the »cultural point of departure« for those academic research communities appears quite similar.

Concerning the empirical outcome of such a bibliometric analysis that is oriented towards journals, there are two crucial results: (1) When the absolute number of articles is used as reference[211], then *Germany represents one of the most important and salient article-producing nations* among the OECD countries (see Figure 7). Only the United States, UK, and Japan lie ahead of Germany, while Germany can outpace any Continental Western European country. In comparison to France, as an example, Germany managed, in the year 1993, an aggregated total article output for science and the social sciences which was higher by a factor of 28%. So within the context of global sciences Germany occupies a strong position, which again implies that at the level of international knowledge exchange and dialogue the German national research system developed into a »collective actor« which cannot be ignored by the international scientific research community. German academic research is a key element for the ongoing process of the global evolution of sciences. (2) If, however, not the absolute number of articles but different dynamical or »output/output« relations are taken into account, then the empirical assessment of Germany's article output performance leads to a more critical result. Such a conclusion we can base on several key observations:

– When article output is put in relation to population for the purpose of calculating a population/publication ratio, then this ratio has a significant positive correlation with gross domestic expenditure on R&D and the number of researchers per thousand labor force (see Figures 8, 9, and 10). Only three countries deviate from this pattern substantially: Germany, France, and Japan. On the one hand they can be

[211] That means, articles in SCI and SSCI journals.

characterized as nations that practically devote a major investment of resources into R&D. On the other hand they only achieve a below-average population/publication ratio with regard to articles. So a critical observer could raise the provoking question whether those three countries represent »worst-case scenarios«?

– When Germany's aggregated article output in science is expressed as a percentage of the total OECD[212] output, then one must state a gradual relative decline over the period 1980–1990. Only because of the geographical enlargement as a consequence of unification, could unified Germany of the early 1990s achieve a similar percentage value like »old« West Germany in the years 1980 and 1981 (see Figure 11).[213] At the same time the fifteen-country cluster of non-English speaking Western Europe – including Germany as one country case – expanded, during the late 1980s and early 1990s, within the overall OECD context in aggregated science as well as in the aggregated social sciences faster and more successful than Germany (see Figure 13). This, ultimately, implies that some of the non-English speaking Western European countries managed a growth rate of their article output that clearly outpaces Germany's growth rate.

– If on a discipline-by-discipline basis Germany's article output is calculated as a percentage value of the total output of non-English speaking Western Europe, then one general trend can be diagnosed: with the only exception of four disciplines, the average value for the five-year period 1989–1993 is higher than the value for 1993 (see Figure 12). This additionally supports our previously mentioned observation that the aggregated article output of non-English speaking Western Europe expands faster than that of Germany.

Under the premise that efficiency of academic research is modeled by referring the article output in journals (covered by SCI and SSCI) to the input in academic research – monetary and personnel resources of the higher education, government, and private non-profit sectors –, an empirical evaluation of Germany's academic research performance will arrive at a considerably critical conclusion: *Germany's academic research system exhibits some efficiency problems that should be taken seriously*. Such a conclusion can be based on the following two observations:

– In Figure 8a the interrelation between monetary input and publication output is expressed, by offering and displaying the following ratio: the number of articles *per* one million $ (in PPP) which are invested into academic R&D (mainly for the year

---

[212] Without Iceland, Mexico, and Turkey.

[213] In the social sciences Germany's national academic research system performed, during the years 1989–1993, a more successful expansion – in relative terms – than in science.

1992). By occupying only the seventeenth position out of a comparative ranking of 21 OECD countries, Germany clearly places only in the bottom third.

– In Figure 8b the focus concentrates on the relationship of personnel and publication output through investigating the ratio of: the quantity of articles *per* one »person year« measured in full-time equivalents (the reference year is 1991). In reference to that indicator, Germany's academic research system can demonstrate a somewhat better efficiency performance. However, by positioning at rank 11 – again out of a sample of 21 OECD member countries – Germany lies only within a medium and not a top efficiency range.

For any analyst the challenge now seems to be to decide how those German article output deficiencies should be correctly interpreted? The core content of our hypothesis – that clearly emphasizes the necessity of a comprehensive or overall (holistic) concept and understanding of academic research – can be stated as follows: *Germany's academic research system in particular and German R&D in general exhibit certain deficiencies and competitiveness problems when the international science and technology markets are taken as frame of reference.* In that context two observations appear to us crucial:

– The German national system of research and innovation, in its broadest sense[214], can be characterized as having developed a strong orientation towards Europe and, in economic terminology, towards the EU domestic market; at the same time, however, Germany exhibits certain problems of competitiveness at the international and global R&D and technology markets. In a current study, the OECD emphasizes that with regard to high-tech trade Germany is doing best in medium-high and medium-low technology, but concerning high technology Germany clearly reveals trade deficits and competitiveness problems. In some of the most dynamic technology markets the German industry – or national R&D system all together – suffers from a loss of market shares (OECD, 1994a, 201–202). The internationally renowned IMD, that is located in Lausanne, Switzerland, also concludes a decline of the current competitiveness of Germany's economy. Compared with OECD and other industrial countries, Germany ranked at position six in 1995; in 1996, however, Germany was pushed down to the rank position ten (IMD, 1996, 18–29).

– Another key term for evaluating the international competitiveness of R&D of a given national research system is the technology balance of payments (also abbreviated as TBP) that covers the exports and imports of technology. Those technology products and services, which are taken into account, easily can be understood and interpreted

[214] This means looking at the academic and industrial research and technology performance comprehensively.

as an outcome of R&D activities.[215] In Figure 19 we express to which extent the technology imports are covered by technology exports of several OECD nations at the beginning of the 1990s.[216] This means that while the United States, in 1993, exported more than four times as many technology products and services than they imported, for example in Austria the exports only covered a ratio of 28% of the imports. Out of a ranking of sixteen OECD countries, Germany rates only at position ten and by this lies below the average (see again Figure 19). What counts perhaps even more is the fact that with a value of 69% (for the year 1993) Germany imports more technology than it exports.[217] Therefore, speaking in the context of a general perspective, this not very favorable (although, of course, also not completely unfavorable) technology balance of payments for Germany could be used as a reference indicator for setting up the hypothesis that in its technology or technology-driven R&D activities Germany is not oriented enough towards the international markets.

When arriving, however preliminary, at a more comprehensive assessment of Germany's academic or total national research performance, then we are inclined to set up for discussion the following hypothesis: *one of Germany's main problems seems to be that Germany's academic research (and perhaps also national R&D) is biased towards the »domestic pole« and is not enough internationally or outwardly oriented. Germany's academic life can be characterized to have cultivated a very strong* **domestic discourse** *which could be criticized for not demonstrating enough openness for reflecting what is happening internationally.* In that respect Germany's academic research system seems to feature some similarities with France – and perhaps also with Japan. Those three countries can be characterized as »larger medium-sized« nations, when the global context is taken as a frame of reference, so they developed different academic research strategies than the small-sized Western European countries. Whereas countries such as Switzerland, the Netherlands, and Sweden, partially out of a scarcity of national R&D funds and a lack of national publication means, developed much earlier a willingness to publish in English and in international journals (Felderer and Campbell, 1995a, 62–63), the German academic research system expresses several weaknesses. Two we want to mention explicitly in that context:

– Currently, when a provoking thought should be thrown into discussion, Germany's academic research system – at least partially – may be paraphrased as having developed a *self-referentially closed national domestic discourse* in the humanities, social sciences, and to a certain extent also in science. Of course we know that such

---

[215] The OECD offers the following definition for TBP: »It consists of money paid or received for the use of patents, licences, trademarks, designs, know-how and closely related technical services (including technical assistance) and for industrial R&D carried out abroad, etc.« (OECD, 1996a, 64).

[216] The mathematical formula is to divide the exports (receipts) by the imports (payments).

[217] The value 100% would indicate a perfect equilibrium between exports and imports, and any value higher than 100% would imply an export surplus.

a statement certainly is an exaggeration, but sometimes also exaggerations are helpful in pinpointing problems; and a problem definition is a prerequisite for developing strategies that aim at improvements. So when the famous German scholar Niklas Luhmann, who became an influential thinker on modern systems theory, uses *self-referentiality* as a key term, then we could go a step further an ask ourselves the question, whether not the whole German university and academic research system could be described as a system biased for self-referentiality? When an interaction between thoughts and the socio-economic environment is accepted, then the theoretical or abstract models of German scholars might contain much »hidden« or indirect information about German society. Characteristics such as a reluctance against the use of English, the inclination for books and the tradition of the *Habilitation*-system would underpin such an interpretation or impression of a partially closed national discourse which an outside observer might have who should assess Germany's universities.

– German academia still puts a very strong emphasis on the German language and on publishing in German. On the one hand, such a behavior appears reasonable and can further be legitimated by the reference that all nations try to encourage in intellectual life their own native language (or languages). On the other hand, it is a given fact and condition that English has become the most important lingual means for communication and the exchange of ideas in global sciences. Therefore, each non-English speaking national academic research system must take this into account and accept – as a consensus and compromise for the advancement of sciences – the necessity to publish in English. So this should provoke a demand for Germany's academic scholars to publish more in English.

Section II: The Evaluation of Germany's University Research: Discourse and Policy

Concerning the evaluation of university research (*Hochschulforschung*) in Germany, many experts express the opinion that in reference to the current situation *something like a consequent evaluation system is still not extant and is still not applied for Germany's university research*. In other words, it is an extensively debated issue how such an evaluation system should be designed and what the adequate means for implementation are. Such an assessment can be derived from the following two facts:

– First of all, a comprehensive *ex-post* evaluation of all of Germany's university research at the national level – for a given year or a several-year period – was never carried out. Such an exercise clearly would mark a watershed for German academic culture and would resemble »pioneer land« for the development of a policy expertise for all the involved actors.

– Secondly, the experience of systematically evaluating single universities has also not been developed that consequently in contemporary Germany – when there is agreement that a so-called systematic evaluation would have to include substantial quantitative indicators. Those »institutional« evaluations, of a specific university, obviously could not focus only on research, but would also have to cut across other issues such as teaching and administration.

Therefore, the current situation may be characterized by the following two main problems:

– Despite a whole series of individual evaluation initiatives of university research at the micro or »bottom« level, we must state that the *generic picture* or the *comprehensive overview* is still missing. So to speak – and willing to employ a provocative phrase – Germany's university evaluation expertise and knowledge appears »fragmented«, and up until now no attempt was undertaken to evaluate university research empirically at the national level for the whole university system (*e.g.,* in an *ex-post* fashion).

– The other major challenge seems to be, what should be the function of future evaluation initiatives for Germany's university research – and the universities in general. Obviously there should be some agreement on the consequences of evaluations, which must be settled in advance. The more systemic question to ask would be: *How should evaluations be implemented into the university context and what should be their feedback link to other functional key elements of the university system?* One of the most sensitive political questions, to which an adequate answer must be found in the next years, is if there should be a systematic and formal link between the outcome of an evaluation and the degree of public basic funding that is allocated to such an evaluated university unit.[218]

However, there are developments that will demand that also in Germany the concept of evaluating universities will gain a crucial and strategic importance in the future. In the following we want to summarize some of those trends or reasons that emphasize the necessity of evaluations.

– There are many different definitions for the common underlying purpose or aim of the broad spectrum of the whole sciences (including the disciplinary branches of the social sciences and humanities). As the devil's advocate, of course, one could even raise the provoking question whether such a common conceptual basis really exists for all of the sciences? However, for the practical purpose of continuing our analysis we want to set up the assumption or premise that all of the sciences can be characterized by some similar attributes. So for us a key definition of the function or

[218] Other countries, for instance the UK, have already decided to establish a tight feedback linkage between the two factors *evaluation and funding.*

goal of sciences would be: *the interest to create knowledge about the empirical world or the empirical environment.*[219] In that definition a crucial term obviously is the word »empirical«, since anything that is empirical can also, at least in principle, be measured. Such a conceptual framework then implies that the sciences are significantly concerned with creating knowledge about empirical structures and processes, that means of *measurable units*. Therefore, in our opinion it is absolutely clear that when the sciences deal with the empirical world as a part of their research procedure and research mission, *i.e.,* with measurable units, then the research output »itself« consequently can also be measured – at least this would reflect the general demand. Turning this argument around, this would imply that when a research output (for instance, its quality and efficiency) cannot be measured, then such a particular research output is not the product of an empirical or scientific-based approach. Speaking in simplified terminology, only non-scientific or even anti-scientific research output would be »non-measurable«.

– When, in principle, *research and research output are measurable,* this implies that *research quality and research efficiency can also be measured*. Going a step further and referring to issues of policy and policy-making, this clearly demonstrates that a rejection of systematic evaluations of university research cannot be justified by scientific arguments. Speaking positively, comprehensive evaluations of university research are easily to legitimate by those fundamentals that structure and drive the evolution of modern sciences *per se*. Coming down to the question of which methods are appropriate for research evaluations, then we are inclined to emphasize the notion of a *plurality or pluralism of methodological approaches;* as long as those different methods meet the criteria of an accurate and sound scientific analysis. Furthermore, it is plausible to assume that concerning the »general picture« those different methodological approaches should converge, at least by tendency, in their results and conclusions (see, for instance, van Raan, 1995, 93).

– When some German academic scholars assert that research output cannot be measured, then – in addition to the arguments we already have stressed – we want to emphasize that such a statement is not internally consistent. In other words: *scholars pushing the non-measurability hypothesis reveal a cultural »blind spot«* (in German: *blinder Fleck*) or, at least, apply a double standard of morals by treating the students, on the one hand, differently than the university research staff, most prominently the professors. Concerning the students, everybody accepts it as a given and »natural« fact that the students' performance should be evaluated on a permanent basis. In addition, when students work for their first-degree thesis or – at the post-first degree level – on their dissertation, then, at least according to the theory, the »university system« demands or expects that those students generate research

[219] This obviously includes »ourselves«, since also humankind is a part of the empirical world.

results of their own. And those completed theses at the Master or Doctoral level then are evaluated and graded – that means the quality of their research output is »measured« – by their supervisors, mainly professors. In the case that an academic researcher wants to go through a *Habilitation* procedure, the above said obviously is also true: firstly, a *Habilitation* work must (or should) reflect an academic in-depth research, and, secondly, the final *Habilitation* is judged and evaluated by a university commission. *Therefore, when in the case of students and applicants for a Habilitation it is commonly accepted that their academic research work must be consequentially evaluated and permanently graded, that means their research output is measured, so why should this not also apply to the academic university research staff, most importantly the professors?* From the institutional perspective the answer appears simple: when, based on scientific premises, it is legitimate to evaluate the research output of first-degree and second-degree level students, then it cannot be justified not to evaluate professors with the same strictness. So when professors are not evaluated, then this seems to be more the product of a hierarchical distribution of privileges than the self-logic of a scientific rationale.

– Concerning the funding resources of German universities and German university research, there are two crucial facts that we should recall: firstly, university R&D depends financially primarily on public basic transfer funds, called GUF (General University Funds). In 1991, 71.4% of German university R&D was funded by GUF (see Figures 17 and 18). Secondly, the German universities claim that they are currently *underfunded* or *underfinanced* (*unterfinanziert*). One estimation, that was calculated by the HRK, speaks of an annual structural financial deficit for the whole German higher education sector of between 6 and 9 billion DM (see our analysis in Chapter 3.1). From that two crucial consequences can be drawn: (1) The dominance of the funding category of GUF implies that *ex-post* evaluations represent more or less the only possibility for evaluating university research systematically, that means having a chance to project transparency onto the system. In addition, GUF funding even demands – theoretically speaking – *ex-post* evaluations since, already by definition, GUF funding denies the possibility of accurate *ex-ante* evaluations.[220] (2) The monetary resources for the German universities are seriously constrained. This determines two consequences for Germany's university sector: first of all, the universities must learn how to use their resources more efficiently and go through a painful restructuring process of adaptation. In addition, and secondly, probably the only chances of the universities to achieve an increase of their public funding-base is to convince the public of the importance of universities. And this will, in its final logic, demand the introduction of comprehensive evaluation systems.

---

[220] See Chapter 3.1 where we discuss those different funding categories of universities and university research.

– *Transparency is necessary and essential*. Without an adequate degree of transparency a successful rational decision-making and policy-making is very difficult to realize, since the development of intelligent and sophisticated strategies would demand, for instance, the supply of sufficient data. When such data are missing, then decisions must be based on assumptions which, of course, can be true; this means that assumptions correspond with the empirical world or with the results of an empirical inquiry that is carried out at a later moment in time than the formulation of premises. On the other hand, however, assumptions are often primarily the product of an ideological belief or prejudice which either is empirically wrong or – in a perhaps less dramatic situation – is used as an *ex-ante* argument against a systematic and scientific analysis. *Therefore, a key strategy to produce transparency within the context of an individual university or, more generally speaking, for the whole university sector is to carry out comprehensive and systematic evaluations on a regular basis*. Such evaluations again guarantee that systemic feedback mechanisms are incorporated and can operate within the university framework. This, finally, demands that a university is sensitive to feedback and that it supports transparency – in that context the phrase of a *gläserne Hochschule*[221] is used in the German discourse (see Lange, 1995a, 72).

– The public of an advanced democratic and industrial society expects transparency from all institutions, particularly when those institutions claim to be of value for the whole nation. Democracy supports the development of a »civic culture«, where society demands that main actors must be able to present themselves adequately to the public. Now drawing the line to the issue of university research and its evaluation, our hypothesis is the following: German society expects more transparency from the German universities. First of all, German universities are primarily public funded, which means that their primary financial resource is derived from the tax money that the Germans (and German residents) pay. Secondly, the German universities claim that their basic research activity is for the good of all of German society. Another crucial argument, in that context, seems to be that probably only when the public demand for more transparency and accountability in the German university sector is met, then there might be a public consensus for significantly increasing the public funding base for universities.

In the following we will present some of the policies that are either discussed or actually even applied (or that are at the beginning of an application), and that aim at evaluating university research in Germany:

[221] A translation into English of *gläserne Hochschule* would be: a glass university, *i.e.,* a transparent university.

(1) **The importance of indicators and of quantitative indicators for evaluations:** Potentially any evaluation policy is caught in the dilemma of emphasizing *either quality or quantity.* To phrase this dilemma somewhat differently, one can set up the following question: Should a model, that underlies the evaluation process of university research (or other areas of university performance), try to cover the »total spectrum of complexity«, that means to recognize all the details of a university institution, or, on the contrary, should an evaluation model preferably focus on a limited number of indicators which are considered to be of a crucial importance? This means, what is the role or function of *quantitative indicators* in the context of an evaluation procedure? Obviously, there are divergent opinions of experts on that issue. However, many experts are inclined to emphasize and to underscore directly the *crucial importance of indicators* – also of quantitative indicators. Their conclusion is that, regardless which model is used, only those evaluation policies of university research can be efficient and produce meaningful results that *also* employ the use of indicators (see, for example, Daniel, 1995, 206). Obviously, *the concept of indicators or of quantitative indicators* can have two separate and distinct meanings: (1) The first and more simple implication is to interpret the function of indicators as counting directly units which are measurable. Examples would be the size of research staff, the number of patents, the number of publications, or the number of diplomas, and so on. In reference to the concept of first-order and second-order cybernetics, those indicators, consequently, could be classified as *first-order indicators*.[222] (2) The second group of indicators – that, within the same line of argument, could be paraphrased as *second-order indicators* – are already of a higher complexity, since their functional purpose is not just to count, but to interpret, or, in other words: *to translate quality into quantity*. That means, they aim at representing structures and processes – which we might label as »qualitative« – within a quantitative setting or, speaking simply, as numbers. *So, finally, qualitative patterns of the empirical world can also be reflected and expressed by quantitative indicators*.

(2) **The practical procedure of developing indicators:** Already at the beginning of the 1990s, in the year 1991, the HRK – German Rectors' Conference – initiated a pilot project, called *Profilbildung*.[223] In the first phase, completed in 1993, the attention focused on the disciplines (*Fächer*) physics, German language (*Germanistik*), and economic sciences. In a second phase, finished in the year 1994, the disciplines covered were electrical engineering, mechanical engineering, and computer sciences (*Informatik*). The primary purpose of this project was to invent and to develop *ad hoc* models and procedures, still at the bottom-up level of individual universities, which, in

---

[222] For a further discussion of the interesting concept of first-order and second-order cybernetics, see Umpleby (1990). That term was originally invented by Heinz von Foerster and published, for the first time, in 1979 (von Foerster, 1979).

[223] *Profilbildung* could be translated into English with the phrase »development of a profile«.

a later phase, could have the potential to be regularly applied to all universities nationwide. Consequently, these assessment exercises did not cover the whole university system, but only a small sample of those universities that participated voluntarily. During the second phase those were, all together, not more than eleven universities. Functionally the assessments were institution-oriented, that means they preliminarily attempted to evaluate or, to be more precise, they described comprehensively the performance in research and teaching in those pre-defined disciplines (as listed above) at the level of university departments as well as for the whole university. Methodologically, a twofold approach was applied: first of all, descriptive statistical data were collected – a so-called *Erhebungsraster* – with the main purpose being to develop indicators that present a broad overview of the general performance. Secondly, the data overview was accompanied by a verbal or narrative description that intended to offer to the universities the opportunity to comment on their situation and their individual performance in research and teaching. Obviously, what this pilot project did not aim at was to develop and to set up for discussion a quality-based *ranking* of different universities or university departments. So, in our terminology, the design of »second-order« indicators was only a partial goal or intention (for further literature, see HRK 1994b, and HRK 1994c). In Chapter 3.2.5 we document which indicators and ratios exactly were developed.

(3) **Earmarked funds (*Drittmittel*) as a key indicator for the quality and/or relevance of academic research:** Among leading experts, and also among decision makers, many express the opinion that earmarked funds – in German called *Drittmittel* –, or the »degree« of earmarked funding, *should be valued and interpreted as a crucially important variable for the process of evaluating university research*. In other words: earmarked-funded academic research is a key indicator for *the quality as well as the relevance* of university R&D. Such a claim can be based on the following arguments: (a) Since earmarked-funded research – no matter if it is processed either in the context of a research project or a medium-scale or large-scale research program – is always pre-assessed during a peer-review procedure (that follows the application phase), this implies that such a research already automatically will be valued by an *ex-ante* evaluation. Beyond that, because the outcome or »final product« again is assessed by those institutions or organizations that provided the earmarked funds, this leads to the consequence that earmarked-funded academic research must pass a *double quality check;* an *ex-ante* evaluation at the beginning and an *ex-post* evaluation after completion. In contrast to that, university research that is financed by basic funds, is not exposed to an in-depth and rigorous *ex-ante* quality control procedure and, in addition, is only very selectively assessed *ex-post,* since a systematic and comprehensive *ex-post* evaluation system for university research still has not been implemented in Germany. Therefore, because of this »double quality control«, earmarked-funded research can be interpreted as an *indicator for quality*. (b)

Now depending on the specific source, such earmarked-funded academic research can also be *an indicator for the relevance* of that research – and its output – for society in general, or for the economy, or other decision makers in more particular. When an earmarked fund is provided by a company or firm of the business enterprise sector, or other private or public institutions (for instance, a ministry), which – all together – commonly can be circumscribed to have a strongly focused application-oriented interest into the practical results and conclusions of academic research, then it appears legitimate to interpret, in such cases, the degree of earmarked funding as an indicator of relevance. In other cases, where an institution providing earmarked funds either does not have a »political« interest at all or at least not a direct interest in practically applying the research results – for example, the DFG and, to a certain degree, also the foundations – the function of a quality indication of those earmarked funds still is continued. (c) Earmarked funds have the great advantage that they are relatively easy to survey. In other words: the development and build up of a data collection, that gives information on the degree of earmarked funding of university research, can be done at a reasonable cost or price. So when there is, firstly, an interest in analyzing the quality and relevance of university research, and, secondly, it is accepted that earmarked funds are a proper indicator for that, then an investigation of the degree of earmarked funding is the »cheapest« and fastest way to obtain information on such an issue. Therefore, in summarizing, the simple equation would be: *the higher the degree or intensity of earmarked funds in relation to the extent of basic funding, the more relevant and/or the higher the quality of research* that is conducted by a university department or a university entity as a whole. Finally, arriving at a conclusion that reveals policy relevance, we want to add that in some German *Länder* (»Federal States« or provinces) a funding formula for universities is in the process of discussion *that orients itself primarily towards the degree of earmarked funding*. In other words: those universities should be »rewarded« that prove to be successful in acquiring such earmarked funds; and, consequently, universities which are not that successful in that respect might be »punished«.

(4) **The university-internal redistribution of resources depending on performance:** At several German universities models are being discussed – or already implemented – that should enable *a university-internal redistribution or re-allocation of resources depending on the performance of single university departments* (or of other units within the context of a university). Such a university-internal redistribution process should fulfill several functions, such as: (a) first of all, any redistribution must be based on »objective«, »quasi-objective«, or »objectively created« data and data-based indicators. This is necessary so that redistribution decisions can claim a sufficient degree of legitimation within a university community. Therefore, as some experts would argue, such discussions or – even more – decisions on redistribution massively foster a thinking and acting *that favors the design and development of indicators* that

have the capability of adequately reflecting the performance of individual universities. (b) As the next conceptual step*, those university-internal redistribution processes should help spreading transparency within the environment of a single university*. Since, as we just have elaborated, redistribution decisions demand the development of a data base, the regular and continuous reporting of data becomes a standardized routine; and, as an essential by-product, this leads to an overall increase of transparency within that university entity. (c) *The performance-based redistribution of resources should increase the »internal« rationale of the system*. In simple terms, redistribution implies the following formula: university departments or academic university staff (*e.g.,* professors) with a good performance should be »rewarded« (by an increase of resources) and departments with a poor performance should be »punished« (by a decrease of resources). The rationale is the following: firstly, a performance-based allocation appears to many observers or analysts to be more just or fair than an allocation which is based mainly on »historical« claims. Secondly, either in case of a general increase or decrease of the funding base for universities, such performance-based re-allocation processes offer a mechanism or key for passing on those financial shifts and changes to the level of individual departments. Thirdly, a long-duration impact on the academic culture seems possible in the sense that a general attitude is encouraged that favors performance, output, and some forms of efficiency. Fourthly, university-*internal* competition is encouraged; this appears to be highly compatible with the general demand of an overall increase of competition *between* different universities. Therefore, similar processes at the »micro« and »macro« systemic level can be linked to each other, which again possibly creates certain synergy effects. (d) *Performance-based redistribution processes within universities can help to improve substantially the general public image of universities*. Redistribution mechanisms might be interpreted by the public as well as by public decision makers as a serious attempt by the universities to increase transparency and to introduce accountability and performance-favoring attitudes within the context of a university institution. Particularly during the current phase, in which the universities claim that they are underfinanced by the public, such a university policy – that of implementing a university-internal redistribution – could be used in the academic discourse and in public discussion as a key argument, with a crucial symbolic meaning, for the interests of universities. (e) The pivotal question, which remains, obviously is whether such a re-allocation process should be regarded primarily as *a setting of incentives* that can stimulate university research activities into a desired direction or, contrarily, whether *a fundamental and far-going redistribution of resources* should be the ultimate goal? Currently, as it appears to us, there is no consensus among experts on that important issue. While some experts qualify it as sufficient to use and interpret the university-internal re-allocation as a means for defining incentives, others are more inclined to favor the concept of a fundamental redistribution: thus, in the opinion of those experts, also the basic

university funds should be addressed by the university-internal re-allocation decisions. The Free University of Berlin (FU Berlin)[224] already has implemented a scheme that processes a university-internal redistribution of resources depending on performance. Since this scheme is qualified by many experts as interesting and innovative, we present and discuss the key features of this »FU Berlin« model in more detail in Chapter 3.2.5[225] – for an excellent summary and overview see also an article by Peter Wex, published in 1995 (Wex, 1995).

(5) **Institutional reforms of the German university system to foster evaluations:** There is an intensive debate going on in Germany that focuses on the question of which *institutional reforms* are necessary or at least desirable for improving the quality and efficiency of the overall performance of the German universities. In that context, one of the key conclusions is that specific institutional structures can either favor or constrain the comprehensive use of evaluations.[226] There are powerful arguments that convincingly demonstrate the »systemic« reasons for why evaluations in general, but also why evaluations of university research in more particular, will gain a crucial importance in the future, so the challenge can exactly be summarized as: *to design and to implement institutional structures into a university environment that, firstly, promise a high compatibility with the demands of evaluations and, secondly, beyond that even foster the systematic application of ex-ante and ex-post evaluations*. In that respect there are two sensitive issues: Should professorships be granted on a permanently tenured or only a temporally limited basis? And: Are there still rational and scientifically-based arguments that justify the demand of a *Habilitation* for a professional academic career?

Section III: The Evaluation of Germany's University-Related Research: Discourse and Policy

At least some German experts are willing to set up the hypothesis that in the German context the university-related research (*außeruniversitäre Forschung*) already has been more systematically evaluated or performs at a higher level of sophistication than university research (see, for instance, Krull, 1994, 206). If this is true, then this would indicate a certain *unbalance or policy asymmetry* between the university and the university-related sectors or it would imply, to phrase it somewhat differently, that the diagnosed lack of evaluations seems to be primarily a problem in the case of the universities.

---

[224] In German: *Freie Universität* Berlin.

[225] As one expert indicated, the »FU Berlin« model could be interpreted as a first and perhaps preliminary or somewhat pre-mature design and step towards a more complex and in-depth model for a *formula-based funding* of a whole university.

[226] In the Chapters 3.2.2 and 3.2.3 we present an overview of the structural and cultural constraints against a comprehensive evaluation of university research in Germany.

Of a tremendous and crucial importance for the whole evaluation process of academic research in Germany was the unification of former West (BRD) and East Germany (DDR) to a new state, and ultimately to a new nation, in the year 1990 (Mayntz, 1994, 18). One major challenge arising out of this unification was: What should happen to the East German science and research system? Concerning the former university-related research cluster in East Germany, which consisted primarily of Academy institutes, two fundamental decisions were processed. Firstly, those Academy institutes should not be sustained, but »dissolved« and re-integrated into the structural framework and the institutional design of the West German academic research system (Meske, 1993, 17–18). Secondly, this re-integration process should be preceeded by an in-depth evaluation of those former Academy institutes, which, as a final outcome, would decide which institutes or which staff members are of a scientific value for the unified German academic research system and thus should continue to carry out research. The »West German« Science Council (*Wissenschaftsrat*) was officially asked by the West German and East German governments, in July 1990[227], to carry out this tremendous task of evaluating and of putting forward recommendations for the university-related research in the DDR. The major part of that evaluation procedure, and the main conclusions, were finally completed by the Science Council in July 1991 (Wissenschaftsrat, 1992a, 7).

Speaking in more general terms, those Science Council-based evaluations of the East German Academy research institutes imposed the following consequences or provoked the following serious questions for the national academic research system of now unified Germany:

- Firstly, it was admitted for a whole research sector that research output, research quality, and research efficiency are in principle measurable.

- Secondly, if the East German academic research system was subjected to a thorough investigation, then how can it be justified not to evaluate comprehensively the former West German academic research system?

- Thirdly, if it is acceptable to assess the university-related research sector, so why is it not possible to carry out a comprehensive research evaluation of the university system at the national level?

In April 1994 the German Science Council was asked by the BLK (Federal-*Länder* Commission)[228] to evaluate – beginning in 1995 – within five years all institutes of the so-called Blue List (*Blaue Liste-Einrichtungen*). Two reasons were crucial that particularly the Science Council was asked to carry out such a mission: firstly, up until the early 1990s the

[227] Unification between the BRD and DDR was finally legally and politically processed on October 3, 1990.
[228] In German the BLK is called *Bund-Länder-Kommission*.

Science Council had already evaluated in former West Germany about sixty individual institutes, most of them belonging to the Blue List; and, secondly, the comprehensive evaluation of the East German Academy institutes in the years 1990 and 1991 implied, as a consequence, that the Science Council's expertise and know-how in the application of evaluations increased dramatically and significantly. Now concerning the comprehensive and comparative evaluation exercise of all Blue List institutes the Science Council installed, in May 1995, an *Ausschuß Blaue Liste* or, as phrased in English, a Permanent Panel »Blue List« that coordinates and carries out the individual institute evaluations. Those general recommendations on the Blue List, which were published by the Science Council in 1993 (see Wissenschaftsrat, 1993c), served as a conceptual framework or as a masterplan that helped in structuring and designing the currently conducted evaluations. The practically conducted evaluation procedure is similar to those earlier and pre-1990s individual institute assessments; however, there are some minor, but not unimportant differences (see, again, Wissenschaftsrat, 1993c, and Maurer, 1996):

- The Permanent Panel implements or sets up for each Blue List institute a locally oriented *ad hoc* working group (*ad hoc-Arbeitsgruppe*), in which also foreign experts should participate. Those *ad hoc* working groups then compile evaluation reports in reference to those individual Blue List institutes.

- In addition, the Science Council designed and developed a comprehensive six-page questionnaire, which is sent to all Blue List institutes and to which the institutes also must respond. The purpose of the questionnaire is to receive broad and detailed information on the structure, performance, and future plans of each institute. This information should help in creating an accurate picture of the general state-of-the-art of the Blue List research institutions.

- One crucial difference is that in the past, in only three cases, the Science Council explicitly recommended that the public co-funding by *Bund* and *Länder* (federal and provincial governments) of a particular research institute should be terminated. This implies that the Science Council was very cautious in putting forward such a harsh recommendation. This situation, however, seems to have changed dramatically concerning the currently conducted evaluation exercise of the Blue List institutes. Now, as it appears, the Blue List institutes evaluations are much more consequence-driven since the Science Council lost its reluctance of recommending the cancellation of public funding for an individual institute. Therefore, those evaluations also must be taken more seriously by the particular research community that is subjected to an evaluation. Through such a policy the Science Council emphasizes consciously its previously formulated recommendation that there is a demand for more flexibility and dynamics concerning the crucial issue of acceptance of »new« institutes into or the

> exclusion of »old« institutes from the Blue List funding scheme (Wissenschaftsrat, 1993c, 39).

Arriving at a bottom-line conclusion, many experts are willing to assert that the past evaluation of the East German Academy institutes and the currently conducted comprehensive evaluation of the Blue List institutes represent an important watershed for Germany's evaluation policy of academic research. In that context the two following arguments are often put forward: those evaluation exercises were important, firstly, for developing an expertise in evaluation policy and, secondly, in altering the attitudes of the German academic community and of the German public towards evaluations *per se*. However, as many experts stress, this should only be the beginning of a much longer and much deeper-going policy process or even of a policy evolution. Those experts demand that in the future also other areas or sectors of Germany's academic research system should be subjected to a thorough evaluation. Finally, as an ultimate consequence, this would imply that also the output and quality of Germany's university research would have to be evaluated comprehensively at the national level, covering all of Germany. In the opinion of many leading experts those evaluations are qualified as crucial and as a key tool for improving the quality standards and the global competitiveness of Germany's academic research performance, which again is crucial for maintaining the general wealth of German society and the competitiveness of the German economy.

Glossary

AIDS	Acquired Immune Deficiency Syndrome
approx.	approximately
BLIs	Blue List institutes (Blaue Liste-Einrichtungen or Blaue-Liste-Institute) (Germany)
BLK	Federal-Länder Commission for Educational Planning and Promotion of Research (Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung) (Germany)
BMBF	Federal Ministry of Education, Science, Research and Technology (Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie) (Germany)
BMFT	Federal Ministry of Research and Technology (Bundesministerium für Forschung und Technologie) (Germany)
BMWFK	Federal Ministry of Science, Research and the Arts (Bundesministerium für Wissenschaft, Forschung und Kunst) (Austria)
BMWVK	Federal Ministry of Science, Transport and the Arts (Bundesministerium für Wissenschaft, Verkehr und Kunst) (Austria)
BRD	Federal Republic of Germany [»West Germany«] (Bundesrepublik Deutschland)
DDR	German Democratic Republic [»East Germany«] (Deutsche Demokratische Republik)
DFG	German Research Society (Deutsche Forschungsgemeinschaft)
EC	European Community
EU	European Union
ed.	editor
eds.	editors
e.g.	exempli gratia (in German: »zum Beispiel«)
EU	European Union
F&E	Research and Experimental Development (Forschung und experimentelle Entwicklung)
FU Berlin	Free University of Berlin (Freie Universität Berlin) (Germany)
FWF	Austrian Science Fund (Fonds zur Förderung der wissenschaftlichen Forschung)
FOM	Foundation for Fundamental Research on Matter (Netherlands)
FTE	full-time equivalents
GDP	Gross Domestic Product
GG	German Constitution (»Grundgesetz«)
GNP	Gross National Product
GUF	General University Funds
HEFCE	Higher Education Funding Council for England (UK)

HEFCs	Higher Education Funding Councils (UK)
HRK	German Rectors' Conference / Conference of Rectors and Presidents of Universities and other Higher Education Institutions in the Federal Republic of Germany (Hochschulrektorenkonferenz)
i.e.	id est (in German: »das heißt«)
IHS	Institute for Advanced Studies (Institut für Höhere Studien) (Vienna, Austria)
IMD	International Institute for Management Development (Lausanne, Switzerland)
ISI	Institute for Scientific Information (Philadelphia, USA)
ISSRU	Information Science & Scientometrics Research Unit (Budapest, Hungary)
IWT	Institute for Science and Technology Research (Institut für Wissenschafts- und Technikforschung) (University of Bielefeld, Germany)
KMK	Standing Conference of Ministers of Education of the Länder (Ständige Konferenz der Kultusminister der Länder) (Germany)
MPG	Max Planck Society (Max-Planck-Gesellschaft) (Germany)
MPIFG	Max Planck Institute for the Study of Societies (Max-Planck-Institut für Gesellschaftsforschung) (Cologne, Germany)
NAFTA	North American Free Trade Association
NSB	National Science Board (USA)
NWO	Netherlands Organisation for Scientific Research
OECD	Organisation for Economic Co-Operation and Development
ÖSTAT	Austrian Central Statistical Office (Österreichisches Statistisches Zentralamt)
PPP(s)	Purchasing Power Parities
RAE	Research Assessment Exercise (UK)
RAEs	Research Assessment Exercises (UK)
R&D	Research and Experimental Development
SCI	Science Citation Index
SPSS	Statistical Package for the Social Sciences
SSCI	Social Sciences Citation Index
SERC	Science and Engineering Research Council (UK)
TBP	Technology Balance of Payments
TEKES	Technology Development Centre (Finland)
UFC	Universities Funding Council (UK)
UK	United Kingdom
US	United States (of America)
USA	United States of America
WBL	Science Community Blue List (Wissenschaftsgemeinschaft Blaue Liste) (Germany)
Wissen-schaftsrat	(German) Science Council
WZB	Science Center Berlin for Social Research (Wissenschaftszentrum Berlin für Sozialforschung) (Germany)

Index of Contacted Experts in Germany

Directly Contacted Experts (*Face-to-Face* Interviews)

Prof. Dr. Wolfgang van den Daele. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Forschungsschwerpunkt Technik, Arbeit, Umwelt, Abteilung Normbildung und Umwelt, Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-260 or 261. Fax: ++49-(0)30-25491-219.

Prof. Dr. Hans-Dieter Daniel. Wissenschaftliches Zentrum für Berufs- und Hochschulforschung (WZ I), Universität Gesamthochschule Kassel, Henschelstrasse 4, D-34109 Kassel, Germany. Telephone: ++49-(0)561-804-2419 or 2415. Fax: ++49-(0)561-804-7415 or 3301.

Dr. Jochen Gläser. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Forschungsgruppe Wissenschaftsstatistik, Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-593. Fax: ++49-(0)30-25491-541.

Dr. Dietmar Goll-Bickmann. Wissenschaftsrat (Geschäftsstelle), Brohler Strasse 11, D-50968 Cologne, Germany. Telephone: ++49-(0)221-3776-245. Fax: ++49-(0)221-388440.

Dr. Dieter Grühn. Freie Universität Berlin, Projekt Pro Lehre, Malvenstrasse 6, D-12203 Berlin, Germany. Telephone: ++49-(0)30-838-4659. Fax: ++49-(0)30-838-5271.

Prof. Dr. Hans-Dieter Klingemann. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Abteilung Institutionen und sozialer Wandel, Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-321 or 320. Fax: ++49-(0)30-25491-684. E-Mail: Klingem@medea.wz-berlin.de

Dipl.-Ing. Traugott Klose. Freie Universität Berlin, Abteilung Angelegenheiten von Lehre, Studium und Weiterbildung, Kaiserswerther Strasse 16-18, D-14195 Berlin, Germany. Telephone: ++49-(0)30-838 73 500 or 501. Fax: ++49-(0)30-838 73 505.

Dr. Wilhelm Krull. Volkswagen-Stiftung, Kastanienallee 35, Postfach 81 05 09, D-30505 Hannover, Germany. Telephone: ++49-(0)83-81-215. Fax: ++49-(0)83-81-235.

Dr. Josef Lange. Hochschulrektorenkonferenz (HRK), Ahrstrasse 39, D-53175 Bonn, Germany. Telephone: ++49-(0)228-887-115. Fax: ++49-(0)228-887-110.

Konstanza Prinzessin zu Löwenstein. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-302. Fax: ++49-(0)30-25491-308.

Dr. Norbert Marahrens. Bundesverband Deutscher Stiftungen (e.V.), Adenauerallee 15, D-53111 Bonn, Germany. Telephone: ++49-(0)228-218031 extension 32. Fax: ++49-(0)228-214526.

Prof. Dr. Renate Mayntz. Max-Planck-Institut für Gesellschaftsforschung (MPIFG), Lothringer Strasse 78, D-50677 Cologne, Germany. Telephone: ++49-(0)221-3360510. Fax: ++49-(0)221-3360555. E-Mail: mayntz@mpi-fg-koeln.mpg.de

Prof. Dr. Friedhelm Neidhart. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-501 or 503. Fax: ++49-(0)30-25491-514.

Dr. Reinhard Ost. Freie Universität Berlin, Abteilung VI (Forschungsförderung und Forschungsvermittlung), Kaiserswerther Strasse 16-18, D-14195 Berlin, Germany. Telephone: ++49-(0)30-838-736 10. Fax: ++49-(0)30-838-736 04.

Dr. Uwe Schimank. Max-Planck-Institut für Gesellschaftsforschung (MPIFG), Lothringer Strasse 78, D-50677 Cologne, Germany. Telephone: ++49-(0)221-33605-27. Fax: ++49-(0)221-3360555.

Dr. Christoph Schneider. Deutsche Forschungsgemeinschaft (DFG), Kennedyallee 40, D-53175 Bonn, Germany. Telephone: ++49-(0)228-885-2251. Fax: ++49-(0)228-885-2777.

Dr. Kunigunda Schrüfer. Freie Universität Berlin, Abteilung VI (Forschungsförderung und Forschungsvermittlung), Kaiserswerther Strasse 16-18, D-14195 Berlin, Germany. Telephone: ++49-(0)30-838 736 11. Fax: ++49-(0)30-838 736 04.

Dr. Dagmar Simon. Wissenschaftszentrum Berlin für Sozialforschung (WZB), Reichpietschufer 50, D-10785 Berlin, Germany. Telephone: ++49-(0)30-25491-588.

Dr. Andreas Stucke. Wissenschaftsrat (Geschäftsstelle), Referat Hochschulplanung, Brohler Strasse 11, D-50968 Cologne, Germany. Telephone: ++49-(0)221-3776-244. Fax: ++49-(0)221-388440.

Dr. Peter Wex. Freie Universität Berlin, Abteilung Haushalt (II), Harnackstrasse 5, D-14195 Berlin, Germany. Telephone: ++49-(0)30-838 30 96. Fax: ++49-(0)30-838 34 48

Matthias Winterhager. Institut für Wissenschafts- und Technikforschung (IWT), Universität Bielefeld, Universitätsstrasse 25, P.O. Box 100 131, D-33501 Bielefeld, Germany. Telephone: ++49-(0)521-106-4657. Fax: ++49-(0)521-106-6033. E-Mail: mw@iwt.uni-bielefeld.de

Additionally Contacted Experts

Dr. Peter Lange. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), Godesberger Allee 185-189, D-53175 Bonn, Germany. Telephone: ++49-(0)228-57-3641. Fax: ++49-(0)228-57-3601.

Prof. Dr. Detlef Müller-Böling. Centrum für Hochschulentwicklung (CHE), Carl-Bertelsmann-Strasse 256, D-33311 Gütersloh, Germany. Telephone: ++49-(0)5241-9761-21. Fax: ++49-(0)5241-816098.

Dr. Erika Rost. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF), Heinemannstrasse 2, D-53175 Bonn, Germany. Telephone: ++49-(0)228-57-3233. Fax: ++49-(0)228-57-3601.

Dr. Sonnenburg. Deutsche Forschungsanstalt für Luft- und Raumfahrt (e.V.), DLR Projektträgerschaften, Südstrasse 125, D-53175 Bonn, Germany. Telephone: ++49-(0)228-3821-223.

Prof. Dr. Peter Weingart. Universität Bielefeld, Zentrum für interdisziplinäre Forschung (ZiF), Wellenberg 1, D-33615 Bielefeld, Germany. Telephone: ++49-(0)521-106-2778. Fax: ++49-(0)521-106-2782.

Bibliography

Alewell, Karl (1988). Beurteilung der Leistungen von Hochschulen, 41–58, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung. Band 4. Constance: Universitätsverlag Konstanz.

Alewell, Karl (1995). Gestaltung der Hochschulorganisation, 85–94, in: Klaus Dieter *Wolff* (ed.): Qualitätskonzepte einer Universität. Differenzierung, Effektivierung und Vernetzung. Erfurter Beiträge zur Hochschulforschung und Wissenschaftspolitik. Band 1. Munich: Iudicium Verlag.

Atkinson, Harry / Philippa *Rogers* / Richard *Bond* (1990). Research in the United Kingdom, France and West Germany: A Comparison. Volume 1 and 2. Swindon: SERC.

Averch, Harvey (1993). Annoted Bibliography on Evaluation of Research 1985–1990, 279–300, in: Barry *Bozeman* / Julia *Melkers* (eds.): Evaluating R&D Impacts: Methods and Practice. Boston / Dordrecht / London: Kluwer Academic Publishers.

Block, H.-J. / W. *Krull* (1990). What are the Consequences? Reflections on the Impact of Evaluations conducted by a Science Policy Advisory Body. *Scientometrics* 19 (5–6): 427–437.

BMBF (1995a). Grund- und Strukturdaten 1995/96. Bonn.

BMBF (1995b). Basic and Structural Data 1995/96. Bonn.

BMBF (1996a). Zur technologischen Leistungsfähigkeit Deutschlands. Bonn: Druckpartner Moser Druck + Verlag.

BMBF (1996b). Bundesbericht Forschung 1996. Bonn: Bonner Universitäts-Buchdruckerei.

BMFT (1993). Bundesbericht Forschung 1993. Bonn: Bonner Universitäts-Buchdruckerei.

Bozeman, Barry / Julia *Melkers* (eds.) (1993). Evaluating R&D Impacts: Methods and Practice. Boston / Dordrecht / London: Kluwer Academic Publishers.

Brenner, Peter J. (1993). Habilitation als Sozialisation, 318–356, in: Peter J. *Brenner* (ed.): Geist, Geld und Wissenschaft. Arbeits- und Darstellungsformen von Literaturwissenschaft. Frankfurt am Main: Suhrkamp Verlag.

Campbell, David F.J. (1994). European Nation-State under Pressure: National Fragmentation or the Evolution of Suprastate Structures? *Cybernetics and Systems: An International Journal* 25 (6): 879–909.

Campbell, David F.J. (1995). Forschung und Forschungspolitik in Österreich. Ein strategisches Aktionsprogramm für die Sozialwissenschaften. *SWS-Rundschau* 35 (4): 395–404.

Daniel, Hans-Dieter (1988a). Evaluation der biotechnologischen und molekulargenetischen Forschung in der Bundesrepublik Deutschland anhand des ISI Atlas of Science. Eine Momentanaufnahme bedeutender Forschungsbeiträge, 121–150, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Daniel, Hans-Dieter (1988b). Methodische Probleme institutsvergleichender Analysen der Forschungsproduktivität. Untersucht am Beispiel des Faches Psychologie, 215–241, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Daniel, Hans-Dieter (1989). Ansätze zur Messung und Beurteilung des Leistungsstandes von Forschung und Technologie. Eine Bestandsaufnahme biblio-, sziento- und technometrischer Forschungen in der Bundesrepublik Deutschland. *Beiträge zur Hochschulforschung* 3: 223–253.

Daniel, Hans-Dieter (1993). Guardians of Science. Fairness and Reliability of Peer Review. Weinheim: VCH Verlagsgesellschaft.

Daniel, Hans-Dieter (1995). Ist wissenschaftliche Leistung in Forschung und Lehre meßbar? *Universitas* 50 (3): 205–209.

Daniel, Hans-Dieter (1995b). Das Modellprojekt »Evaluation der Lehre« an der Universität Mannheim. Teil 2: Statistische Auswertung von Befragungen in Lehrveranstaltungen, 93–104, in: Peter Ph. *Mohler* (ed.): Universität und Lehre. Ihre Evaluation als Herausforderung and die Empirische Sozialforschung. Münster: Waxmann Verlag.

Daniel, Hans-Dieter / Rudolf *Fisch* (eds.) (1988). Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Daniel, Hans-Dieter / Michaela *Thoma* / Wolfgang *Bandilla* (1995). Das Modellprojekt »Evaluation der Lehre« an der Universität Mannheim. Teil 1: Planung und Durchführung von Befragungen in Lehrveranstaltungen, 83–91, in: Peter Ph. *Mohler* (ed.): Universität und Lehre. Ihre Evaluation als Herausforderung and die Empirische Sozialforschung. Münster: Waxmann Verlag.

DFG (1995). Deutsche Forschungsgemeinschaft. Jahresbericht 1994. Band 1. Aufgaben und Ergebnisse. Bonn: Bonner Universitäts-Buchdruckerei.

European Commission (1994a). The European Report on Science and Technology Indicators 1994. Report EUR 15879 EN. Brussels.

Europäische Kommission (1995). Grünbuch zur Innovation. Luxembourg.

Erichsen, Hans-Uwe (1993). Braucht Deutschland Elite-Universitäten? Nein. *Focus* 39/1993: 142.

Erichsen, Hans-Uwe (1995a). Qualitätssicherung in Forschung, Lehre und Management, 19–26, in: Detlef *Müller-Böling* (ed.): Qualitätssicherung in Hochschulen. Forschung – Lehre – Management. Gütersloh: Verlag Bertelsmann Stiftung.

Erichsen, Hans-Uwe (1995b). Ranglisten nein – Transparenz und Evaluation ja. *Universitas* 50 (3): 216–220.

Felderer, Bernhard (1995a). The Importance of R&D for Future Industries and the Wealth of a Nation. *IHS Newsletter* 3 (4): 1–3.

Felderer, Bernhard / David F.J. *Campbell* (1994a). Forschungsfinanzierung in Europa: Trends, Modelle, Empfehlungen für Österreich. Vienna: Manz-Verlag.

Felderer, Bernhard / David F.J. *Campbell* (1994b). Die Evaluation der akademischen Forschung im internationalen Vergleich: Strukturen, Trends und Modelle. Vienna / IHS: Research Project Proposal.

Felderer, Bernhard / David F.J. *Campbell* (1995a). Die Evaluation der akademischen Forschung im internationalen Vergleich: Strukturen, Trends und Modelle. 1. Zwischenbericht: Patterns of Publication Output and Publication Efficiency of 23 Developed Industrial Countries (1980–1993). Vienna / IHS: First Report of Current Research Project (*status: unpublished*).

Fisch, Rudolf (1988). Ein Rahmenkonzept zur Evaluation universitärer Leistungen, 13–30, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung. Band 4. Constance: Universitätsverlag Konstanz.

Finkenstaedt, Thomas / Marlene *Fries* (1988). Forschungsmessung in der Anglistik. Eine Längsschnittuntersuchung, 151–187, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Frühwald, Wolfgang (1995). Mehr als ein intellektuelles Spiel? *Universitas* 50 (3): 209–212.

FU Berlin (1995). Universitätsbibliographie 13 (1993). Berlin: Universitätsbibliothek of the Free University of Berlin.

Gibbons, Michael / Camille *Limoges* / Helga *Nowotny* / Simon *Schwartzmann* / Peter *Scott* / Martin *Trow* (1994). The new production of knowledge. The dynamics of science and research in contemporary societies. London: SAGE Publication.

Gläser, Jochen (1992). Die Akademie der Wissenschaften nach der Wende: erst reformiert, dann ignoriert und schließlich aufgelöst. *Aus Politik und Zeitgeschichte – Beilage zur Wochenzeitung Das Parlament* (B 51/92): 37–46.

Gläser, Jochen / Gabriele *Groß* / Marion *Höppner* / Charles *Melis* / Werner *Meske* (1995). Die aufgeschobene Integration. Erste Befunde zur Integration neugegründeter Blaue-Liste-Institute in die deutsche Wissenschaftslandschaft. Berlin: WZB P 95–404.

Goodin, Robert E. / Hans-Dieter *Klingemann* (eds.) (1996). A New Handbook of Political Science. Oxford: Oxford University Press.

Gripp-Hagelstange, Helga (1995). Niklas Luhmann. Eine erkenntnistheoretische Einführung. Munich: Wilhelm Fink Verlag.

HEFCE (1993a). Annual Report 1992–93: Promoting Quality and Opportunity. Bristol.

HEFCE (1993b). Assessment of the Quality of Education. Bristol.

HEFCE (1993c). Research Funding Method. Bristol.

HEFCE (1993d). Accountability for Research Funds. Bristol.

HEFCE (1993e). Strategies for Developing Research. Bristol.

HEFCE (1996a). A Guide to Funding Higher Education in England. How the HEFCE Allocates its Funds. Bristol.

HEFCs (1993a).A Report for the Universities Funding Council on the conduct of the 1992 Research Assessment Exercise. Bristol.

HEFCs (1994a). 1996 Research Assessment Exercise. Bristol.

HEFCs (1995a). 1996 Research Assessment Exercise. Guidance on Submission. Bristol.

HEFCs (1995b). 1996 Research Assessment Exercise. Criteria for Assessment. Bristol.

Herbertz, Heinrich / Benno *Müller-Hill* (1993). Qualität und Effizienz. Molekularbiologische Institute im internationalen Vergleich. *Futura* 8 (3): 8–21.

Hohn, Hans-Willy / Uwe *Schimank* (1990). Konflikte und Gleichgewichte im Forschungssystem. Akteurskonstellationen und Entwicklungspfade in der staatlich finanzierten außeruniversitären Forschung. Frankfurt am Main: Campus Verlag.

Honderich, Ted (ed.) (1995). The Oxford Compandium to Philosophy. Oxford: Oxford University Press.

Hornbostel, Stefan / Hans-Dieter *Daniel* (1995). Das SPIEGEL-Ranking. Mediensensation oder ein Beitrag zur hochschulvergleichenden Lehrevaluation?, 29–44, in: Peter Ph. *Mohler* (ed.): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung. Münster: Waxmann Verlag.

HRK (1992a). Konzept zur Entwicklung der Hochschulen in Deutschland. Dokumente zur Hochschulreform 75/1992. Bonn: Self-published.

HRK (1994). Hochschulrektorenkonferenz. Funktion, Struktur, Geschichte. Bonn: Self-published.

HRK (1994b). Profilbildung in den Hochschulen II. Werstattbericht über ein Pilotprojekt der Hochschulrektorenkonferenz. Dokumente zur Hochschulreform 89/1994. Bonn: Self-published.

HRK (1994c). Profilbildung im internationalen Vergleich. Fachtagung der Hochschulrektorenkonferenz am 24. Februar 1994. Dokumente zur Hochschulreform 98/1994. Bonn: Self-published.

HRK (1996a). Zur Finanzierung der Hochschulen. Last-version manuscript (July 1996). Bonn: Self-published.

Hufen, Friedhelm (1995). Rechtsfragen der Lehrevaluation an wissenschaftlichen Hochschulen. Rechtsgutachten im Auftrag des Deutschen Hochschulverbandes. Forum Heft Nr. 62. Bonn: Deutscher Hochschulverband.

IHS / ISSRU (1995). A research cooperation on the basis of a research contract between IHS and ISSRU for the purpose of bibliometric analyses (ISSRU served as contractor and the IHS is the only copyright owner of the results). Vienna and Budapest.

IMD (1996). The World Competitiveness Yearbook 1996. Lausanne: Self-published.

Irvine, John / Ben R. *Martin* / Phoebe *Isard* (1991). Investing in the Future: An International Comparison of Government Funding of Academic and Related Research. Aldershot, England: Edward Elgar.

ISI (1992a). Science Citation Index. Philadelphia.

ISI (1992b). Social Sciences Citation Index. Philadelphia.

ISSRU – A. Schubert / W. Glänzel / T. Braun (1989). Scientometric Datafiles. A Comprehensive Set of Indicators on 2649 Journals and 96 Countries in all Major Science Fields and Subfields 1981–1985. *Scientometrics* 16 (1–6): 3–478.

ISSRU / FOM (1994). Physics in the European Union in the 1980s. A Scientometric Study. FOM: 94.1483.

Klingemann, Hans-Dieter (1988). Zitierhäufigkeit als Qualitätsindikator. Eine Rangordnung der amerikanischen politikwissenschaftlichen Fachbereiche in den 80er Jahren, 201–214, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Krull, Wilhelm (1994). Im Osten wie im Westen – nichts Neues? Zu den Empfehlungen des Wissenschaftsrates für die Neuordnung der Hochschulen auf dem Gebiet der ehemaligen DDR, 205–225, in: Renate *Mayntz* (ed.): Aufbruch und Reform von oben.

Ostdeutsche Universitäten im Transformationsprozeß. Frankfurt am Main: Campus Verlag.

Krull, Wilhelm (1995). The Max Planck Experience of Evaluation. *Scientometrics* 34 (3): 441–450.

Krull, Wilhelm (1995b). Centers of Excellence innerhalb und außerhalb der Universitäten oder: Was sind und zu welchem Ende betreibt man Höhere Studien?, 33–47, in: Klaus Dieter *Wolff* (ed.): Qualitätskonzepte einer Universität. Differenzierung, Effektivierung und Vernetzung. Erfurter Beiträge zur Hochschulforschung und Wissenschaftspolitik. Band 1. Munich: Iudicium-Verlag.

Krull, Wilhelm / Ekkehard *Winter* (1996). A Reader on the Seminar »Von der Evaluation zur Prospektion? – Neue Herausforderungen für die deutsche Wissenschafts- und Forschungspolitik« (Munich, November 10–12, 1995). Essen / Stifterverband für die Deutsche Wissenschaft: Self-published.

Kuhlmann, Stefan (1997). Evaluationspraktiken im deutschen Forschungssystem (geschätzt) (*unpublished figure*). Karlsruhe.

Lange, Josef (1994). Entwicklung und künftige Perspektiven der Hochschulen. *Hochschulforschung* (3): 335–357.

Lange, Josef (1995a). Strategische Planung. Grundlage zukünftiger Hochschulentwicklung. *Wissenschafsmanagement* (March/April 1995): 71–75).

Lange, Josef (1995b). Hochschulpolitik – Die Kunst des Möglichen? Phantasie, Strategie, Bürokratie oder das langsame Bohren harter Bretter. Festvortrag zur Feierstunde anläßlich des 60. Geburtstags des Gründungspräsidenten der Universität Bayreuth, Dr. Dr. h.c. Klaus Dieter Wolff, gehalten am 17. Juli 1995. Bayreuth: University of Bayreuth.

Lehrl, Siegfried / Walter *Kinzel* / Bernd *Fischer* (1988). Der Science Impact Index. Untersucht an den Ordinarien der bundesdeutschen Psychiatrie und Neurologie, 291–305, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Link, Albert (1993). Methods for Evaluating the Return on R&D Investments, 1–16, in: Barry *Bozeman* / Julia *Melkers* (eds.): Evaluating R&D Impacts: Methods and Practice. Boston / Dordrecht / London: Kluwer Academic Publishers.

Luhmann, Niklas (1988). Neuere Entwicklungen in der Systemtheorie. *Merkur* 42 (4): 292–300.

Lundvall, Bengt-Åke (ed.) (1992). National Systems of Innovation. Towards a Theory of Innovation and Interactive Learning. London: Pinter Publishers.

Maddison, Angus (1986). Phases of Capitalist Development. Oxford: Oxford University Press.

Massow, Valentin von (1986a). Wissenschaft und Wissenschaftsförderung in der Bundesrepublik Deutschland. Bonn: Inter Nationes.

Massow, Valentin von (1986b). Organisation and Promotion of Science in the Federal Republic of Germany. Bonn: Inter Nationes.

Maurer, Michael (1996). Evaluation von Forschungsinstituten und -feldern durch den Wissenschaftsrat, in: Wilhelm *Krull* / Ekkehard *Winter* (eds.). A Reader on the Seminar »Von der Evaluation zur Prospektion? – Neue Herausforderungen für die deutsche Wissenschafts- und Forschungspolitik« (Munich, November 10–12, 1995). Essen / Stifterverband für die Deutsche Wissenschaft: Self-published.

Mayntz, Renate (1994). Deutsche Forschung im Einigungsprozeß. Die Transformation der Akademie der Wissenschaften der DDR 1989 bis 1992. Frankfurt am Main: Campus Verlag.

Melkers, Julia (1993). Bibliometrics as a Tool for Analysis of R&D Impacts, 43–61, in: Barry *Bozeman* / Julia *Melkers* (eds.): Evaluating R&D Impacts: Methods and Practice. Bosten / Dordrecht / London: Kluwer Academic Publishers.

Meske, Werner (1993). Die Umgestaltung des ostdeutschen Forschungssystems. Eine Zwischenbilanz. Berlin: WZB P 93–401.

Mohler, Peter Ph. (1995). Vorwort des Herausgebers, 7–9, in: Peter Ph. *Mohler* (ed.): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung. Münster: Waxmann.

Müller-Böling, Detlef (1995). Qualitätssicherung in Hochschulen. Grundlage einer wissenschaftsbasierten Gesellschaft, 27–45, in: Detlef *Müller-Böling* (ed.): Qualitätssicherung in Hochschulen. Forschung – Lehre – Management. Gütersloh: Verlag Bertelsmann Stiftung.

Münzinger, Frank / Hans-Dieter *Daniel* (1992). Biochemie und Molekularbiologie in der Bundesrepublik Deutschland. Bibliometrische Indikatoren 1973–1987. *Futura* 7 (4): 4–11.

Nelson, Richard R. (ed.) (1993). National Innovation Systems. A Comparative Analysis. Oxford: Oxford University Press.

NSB (1991). Science & Engineering Indicators – 1991. Washington, D.C.: U.S. Government Printing Office.

OECD (1989a). The Measurement of Scientific and Technical Activities. R&D Statistics and Output Measurement in the Higher Education Sector. »Frascati Manual« Suplement. Paris.

OECD (1992a). Education at a Glance. OECD Indicators. Paris.

OECD (1992b). Science and Technology Policy. Review and Outlook 1991. Paris.

OECD (1992c). Main Science and Technology Indicators. Paris.

OECD (1992d). Main Economic Indicators. Paris.

OECD (1993a). Main Science and Technology Indicators. Paris.

OECD (1993b). Basic Science and Technology Statistics. Paris.

OECD (1994a). Science and Technology Policy. Review and Outlook 1994. Paris.

OECD (1994b). Main Science and Technology Indicators. Paris.

OECD (1994c). OECD in Figures. Supplement to the OECD Observer No. 188 June/July 1994. Paris.

OECD (1994d). The Measurement of Scientific and Technological Activities. Proposed Standard Practice for Surveys of Research and Experimental Development. Frascati Manual 1993. Paris.

OECD (1994e). Main Economic Indicators. Paris.

OECD (1995a). Education at a Glance. OECD Indicators. Paris.

OECD (1995b). Main Science and Technology Indicators. Paris.

OECD (1995c). Basic Science and Technology Statistics. Paris.

OECD (1996a). Main Science and Technology Indicators. Paris.

OECD (1996b). Main Economic Indicators. Paris.

O'Leary, John / Tom *Cannon* (1995). The Times Good University Guide 1995–1996. London: Times Books.

ÖSTAT (1993). Statistisches Jahrbuch für die Republik Österreich 1993. Vienna.

Rau, Einhard / Thomas *Hummel* (1988). Die Besten und die Berühmtesten. Forschungsproduktivität wirtschaftswissenschaftlicher Fachbereiche in der Bundesrepublik Deutschland und in den USA, 243–258, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Schiedermair, Hartmut (1995). Im Spiegel des Ranking. *Universitas* 50 (3): 213–216.

Schimank, Uwe (1988). The contribution of university research to the technological innovation of the German economy: Societal auto-dynamic and political guidance. *Research Policy* 17: 329–340.

Schimank, Uwe (1992a). Forschungsbedingungen der Professoren an den westdeutschen Hochschulen – Daten aus einer Befragung im Wintersemester 1990/91. MPIFG Discussion Paper 92/2. Cologne. Self-published.

Schimank, Uwe (1992b). The Worsening of Research Conditions at German Universities: Individual Coping Makes the Best of Corporate Coping's Failure. Paper prepared for the conference »Coping With Trouble: How Scientists and Research Institutes React to Political Disturbances of Their Research Conditions« at the Max-Planck-Institut *für Gesellschaftsforschung* in Cologne, November 3–5, 1992. Cologne.

Schimank, Uwe (1995). Hochschulforschung im Schatten der Lehre. Frankfurt am Main: Campus Verlag.

Schimank, Uwe (1995b). Probleme der westdeutschen Hochschulforschung seit den siebziger Jahren. *Leviathan* 23 (1): 56–77.

Schleichert, Hubert (ed.). (1975). Logischer Empirismus – Der Wiener Kreis. Ausgewählte Texte mit einer Einleitung. Munich: Wilhelm Fink Verlag.

Schmidt, Manfred G. (1982). Wohlfahrtsstaatliche Politik unter bürgerlichen und sozialdemokratischen Regierungen. Ein internationaler Vergleich. Frankfurt am Main: Campus Verlag.

Schmidt, Manfred G. (1983). The Welfare State and the Economy in Periods of Economic Crisis: A Comparative Study of Twenty-three OECD Nations. *European Journal of Political Research* 11 (1): 1–26.

Sinz, Elmar J. (1995). Das Informationssystem der Universität als Instrument zur zielgerichteten Lenkung von Universitätsprozessen, 65–83, in: Klaus Dieter *Wolff* (ed.): Qualitätskonzepte einer Universität. Differenzierung, Effektivierung und Vernetzung. Erfurter Beiträge zur Hochschulforschung und Wissenschaftspolitik. Band 1. Munich: Iudicium Verlag.

SPSS (1990). SPSS/PC+ Statistics 4.0 for the IBM PC/XT/AT and PS/2. Chicago: SPSS Inc.

SPSS (1992). SPSS for Windows. Base System User's Guide. Release 5.0. Chicago: SPSS Inc.

Stifterverband für die Deutsche Wissenschaft (1995). Hochschulreform durch Leistungswettbewerb und Privatisierung? Dokumentation eines Symposiums. Materialien zur Bildungspolitik 14. Essen: Stifterverband für die Deutsche Wissenschaft.

Streiter, Axel (ed.) (1992). 20 Jahre Sonderforschungsbereiche. DFG – Deutsche Forschungsgemeinschaft. Weinheim: VCH Verlagsgesellschaft.

Stucke, Andreas (1992). Die deutsche Wissenschaftspolitik auf dem Weg zur deutschen Einheit. *Aus Politik und Zeitgeschichte – Beilage zur Wochenzeitung Das Parlament* (B 51/92): 3–14.

Tarnai, Christian / Holger *Grimm* / Dirk *John* / Rainer *Watermann* (1995). Studienbedingungen des Fachs Soziologie aus der Sicht seiner Studierenden. Eine Auswertung der Neidhart-SPEGEL-Studie unter dem Blickwinkel individueller Urteile, 45–60, in: Peter Ph. *Mohler* (ed.): Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung. Münster: Waxmann Verlag.

UFC (1992a). The Funding of Research by the Higher Education Funding Council for England. Bristol.

UFC (1992b). Research Assessment Exercise 1992. Bristol.

UFC (1992c). Research Assessment Exercise 1992: The Outcome. Bristol.

Umpleby, Stuart A. (1990). The Science of Cybernetics and the Cybernetics of Science. *Cybernetics and Systems: An International Journal* 21 (1): 109–121.

Umpleby, Stuart A. / Vadim N. *Sadovsky* (1991). A Science of Goal Formulation. American and Soviet Discussions of Cybernetics and Systems Theory. New York: Hemisphere Publishing Corporation.

Universität Witten/Herdecke (1994). Vorlesungs- und Personalverzeichnis Wintersemester 1994/95. Witten: Self-published.

Untersteiner, Norbert (1995). Theorien, Experimente und der Computer in den Erdwissenschaften, 126–147, in: Valentin *Braitenberg* / Inga *Hosp* (eds.): Simulation. Computer zwischen Experiment und Theorie. Hamburg: Rowohlt Verlag.

van Raan, Anthony F.J. (1995). Bewertung von Forschungsleistungen: Fortgeschrittene bibliometrische Methoden als quantitativer Kern von Peer-review-basierten Evaluationen, 85–102, in: Detlef *Müller-Böling* (ed.): Qualitätssicherung in Hochschulen. Forschung – Lehre – Management. Gütersloh: Verlag Bertelsmann Stiftung.

von Foerster, Heinz (1979). Cybernetics of Cybernetics, 5–8, in: Klaus *Krippendorff* (ed.): Communication and Control in Society. New York: Gordon and Breach Science Publishers.

Weingart, Peter / Jörg *Strate* / Matthias *Winterhager* (1992). Forschungslandkarte Schweiz 1990. Schweizer Wissenschaftsrat: Forschungspolitik (FOP) 11/1992.

Weingart, Peter / Roswitha *Sehringer* / Jörg *Strate* / Matthias *Winterhager* (1989) (edited by Schweizer Wissenschaftsrat and Schweizer Nationalfonds zur Förderung der wissenschaftlichen Forschung). Der Stand der schweizerischen Grundlagenforschung im internationalen Vergleich. Wissenschaftsindikatoren auf der Grundlage bibliometrischer Daten. Wissenschaftspolitik Beiheft/Supplément No. 44.

Weingart, Peter (1995). Forschungsindikatoren: Instrumente politischer Legitimierung oder organisatorischen Lernens, 73–84, in: Detlef *Müller-Böling* (ed.): Qualitätssicherung in Hochschulen. Forschung – Lehre – Management. Gütersloh: Verlag Bertelsmann Stiftung.

Wex, Peter (1995). Die Mittelverteilung nach Leistungs- und Belastungskriterien. Ein Beitrag zum Leistungswettbewerb in der Hochschule. *Wissenschaftsmanagement* 1 (4): 168–174.

Winterhager, Matthias / Peter *Weingart* / Roswitha *Sehringer* (1988). Die Cozitationsanalyse als bibliometrisches Verfahren zur Messung der nationalen und institutionellen Forschungsperformanz, 319–358, in: Hans-Dieter *Daniel* / Rudolf *Fisch* (eds.): Evaluation von Forschung. Konstanzer Beiträge zur sozialwissenschaftlichen Forschung, Band 4. Constance: Universitätsverlag Konstanz.

Wissenschaftsrat (1985). Empfehlungen zum Wettbewerb im deutschen Hochschulsystem. Cologne: Self-published.

Wissenschaftsrat (1988). Empfehlungen des Wissenschaftsrates zu den Perspektiven der Hochschulen in den 90er Jahren. Cologne: Self-published.

Wissenschaftsrat (1992a). Stellungnahmen zu den außeruniversitären Forschungseinrichtungen in den neuen Ländern und in Berlin – Allgemeiner Teil –. Charakteristika der Forschungssituation in der ehemaligen DDR und künftige Entwicklungsmöglichkeiten einzelner Fachgebiete. Cologne: Self-published.

Wissenschaftsrat (1992b). Stellungnahmen zu den außeruniversitären Forschungseinrichtungen der ehemaligen Akademie der Wissenschaften der DDR auf dem Gebiet der Geisteswissenschaften und zu den Forschungs- und Editionsabteilungen der Akademie der Künste zu Berlin. Cologne: Self-published.

Wissenschaftsrat (1993a). 10 Thesen zur Hochschulpolitik. Cologne: Self-published.

Wissenschaftsrat (1993b). Drittmittel der Hochschulen 1970 bis 1990. Cologne: Self-published.

Wissenschaftsrat (1993c). Empfehlungen zur Neuordnung der Blauen Liste. Cologne: Self-published.

Wissenschaftsrat (1994a). Stellungnahme zur Umweltforschung in Deutschland. Band I. Cologne: Self-published.

Wissenschaftsrat (1994b). Stellungnahme zur Umweltforschung in Deutschland. Band II. Cologne: Self-published.

Wissenschaftsrat (1995a). Stellungnahme zur Aufnahme der Universität Erfurt in das Hochschulverzeichnis des Hochschulbauförderungsgesetzes. Cologne: Self-published.

Wissenschaftsrat (1996a). Empfehlungen zur Stärkung der Lehre in den Hochschulen durch Evaluation. Cologne: Self-published.

Wissenschaftsrat (1996b). Stellungnahme zum Heinrich-Pette-Institute für Experimentelle Virologie und Immunologie (HPI). Cologne: Self-published.

Wissenschaftsrat (1996c). Stellungnahme zum Institut für Zeitgeschichte (IfZ). Cologne: Self-published.

Wissenschaftsrat (1996d). Stellungnahme zum Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel (IPN). Cologne: Self-published.

Wissenschaftsrat (1996e). Stellungnahme zum HWWA-Institut für die Wirtschaftsforschung Hamburg. Cologne: Self-published.

Wissenschaftsrat (1996f). Stellungnahme zum Niedersächsischen Landesamt für Bodenforschung – Geowissenschaftliche Gemeinschaftsaufgaben, Hannover (NLfB-GGA). Cologne: Self-published.

Figures 1–19