

PROCEDURES FOR RESEARCH EVALUATION IN GERMAN HIGHER EDUCATION: CURRENT FRAGMENTATION AND FUTURE PROSPECTS

Dominic Orr

Higher Education Information System (HIS), Germany

Mathias Paetzold

Academic and Research Commission Lower Saxony, Germany

The practice of carrying out systematic research evaluations between the state and institutions of higher education can be characterised as fragmented in Germany. This has a lot to do with the multi-layered character of the system. The German scientific landscape is one of the largest and most diverse in Europe, but it also encompasses a multitude of actors, each with different interests and procedures for evaluation. Indeed, the fragmented character of Germany's higher education system is likely to become further pronounced under Federal reforms of the present government. A forerunner in the field of research evaluation in higher education is the Academic and Research Commission of the state of Lower Saxony. As this body has been active since 1999, its organisation and procedures will be presented as a case study for possible developments and the impact of such procedures in the German scientific system. The article closes with considerations on future prospects of systematic research evaluations in Lower Saxony and in Germany.

Objectives of German scientific and university policy and the task of research evaluation

Science is seen within the context of knowledge-based societies as a key to development and social and economic well-being (cf. European Commission, 2002). The general move of European economies towards knowledge-based economies, therefore, has led to a growth in prestige, but also in expectations of universities (and institutions of higher education), which are seen both as the originators of much new research and knowledge and the breeding ground for new generations of scientists and knowledge-workers. Many countries, including Germany, have been introducing instruments of change into their higher education sectors to promote effectiveness – i.e. the quality of the outcome or process – and efficiency – i.e. the cost of an activity as measured by input to output.

On matters affecting higher education and science, the 16 Federal states in Germany (*Länder*) have a broad sovereignty, which has resulted in differing practices in the *Länder*. Most of the *Länder*, for example, have introduced performance-based funding, but despite broad trends,

there are also differences in the selection and weighting of the indicators used and the proportion of university budgets, which they determine (Orr *et al.*, in press). Such procedures emphasise by their nature efficiency gains, but have only an indirect relation to quality of output and to the processes leading to this output. Therefore, it can be argued that quality evaluation is a necessary complementary instrument for governance and steering systems, which can counterbalance incentives to improve efficiency at the cost of quality (Orr, 2005a).

Quality assurance through evaluation has been given a clear role in the higher education laws of the 16 *Länder* in Germany and further definition of this role can be expected for the revision of these laws, which is occurring currently. The focus of developments has, however, been on teaching and learning. The reasons for this general neglect of research evaluation are multifaceted: most reforms in the last decade have centred on teaching and learning and the challenge of expansion of student numbers. Additionally, scientific research funding via third parties (i.e. not institutional funding) is organised in a competitive system, so that evaluation may have been viewed as less necessary in many states. This situation has been criticised recently, precisely because the organisation of research and the assurance of a high competitive performance are important to the scientific and economic standing of Germany. However, this state of affairs has a lot to do with the fragmented structure of research activities in Germany.

It is the purpose of this article to describe the structure of research in Germany, with a focus on public sector higher education, and to present the different forms in which the state or its agents evaluate public research in Germany. A forerunner in the field of research evaluation in higher education is the Academic and Research Commission of the state of Lower Saxony. As this body has been active since 1999, its organisation and procedures will be presented as a case study for possible developments and to assess the impact of such procedures in the German science system. The article closes with a discussion of the future prospects of systematic research evaluations in Lower Saxony and in Germany.

The scale of Germany's research landscape

Germany has a very large research sector, which can be characterised best in international comparisons (see Figure 1). Germany has by far the largest total number of researchers in Europe, and accounts for 17% of all researchers in European higher education. Indeed, Germany's share in the total number of graduates of advanced research degrees (doctorates; ISCED level 6)1 in Europe is 26%. This high share is directly related to the comparatively high proportion of graduates of lower degrees, who go on to study a doctorate. Similar to the much smaller countries of Austria and Sweden, graduate doctorates make up more than 7% of all higher education graduates in Germany. This huge volume of research capacity is also reflected in a comparative indicator for research productivity: number of scientific publications. On this indicator, Germany is one of the four leading producers of scientific publications, alongside the United Kingdom, France and Italy, who together account for 27.6% of the total world share. Germany alone accounts for 8.4% of the world share and thereby ranks fourth in the world topped in Europe only by the United Kingdom (8.5%).

A better indicator of productivity is, however, to view production volume in the context of capacity. The EU Key Figures published by CORDIS (Community Research & Development Information Service, 2005) achieve this by relating the number of scientific publications to

a country's population. On this indicator, Germany's ranking drops to position nine, behind France with eight and the UK with five. However, smaller countries, with Sweden, Denmark, Finland and the Netherlands in the top four positions, dominate this list. To a certain extent this reflects these countries' specialisation in a select few disciplines, instead of covering all disciplines without such specialisation, as with Germany.

These comparative figures, therefore, show the German science landscape as very large, diverse, but (in terms of publications, at least) less output-orientated than other countries. If Germany wants to improve its productivity and therefore its comparative standing, methods and procedures of research evaluation will be important if we are to map out the local research terrain, to incentivise higher performances and to control the effects of such incentive programmes.

The scale of Germany's research landscape

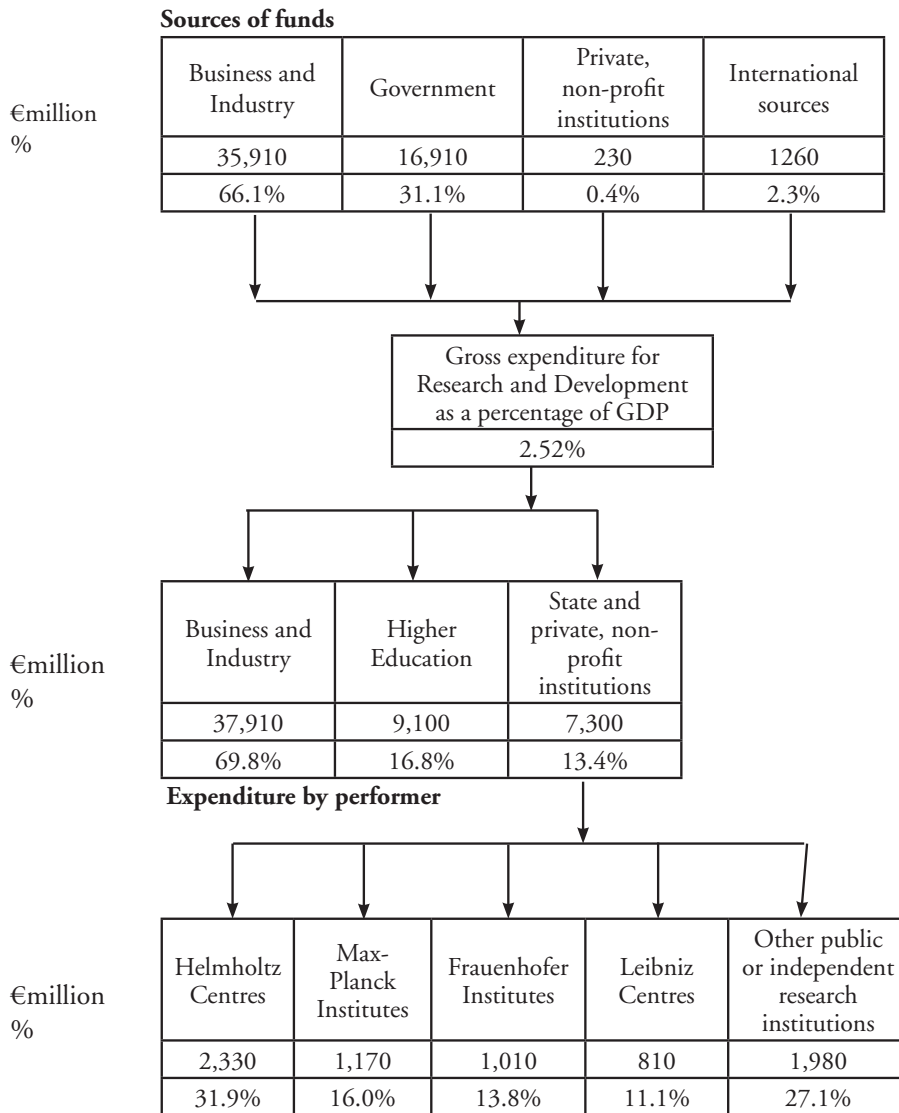
	no. of researchers (FTE) by institutional sector, LYA	in % by sector			no. of researchers (FTE) in HE, LYA	total number of graduates of ISCED6 (advanced researchers, 2003)	proportion of graduates of ISCED6 graduates to all graduates 5 & 6 (%), 2003	world share of scientific publications (%), 2003	no. of scientific pub. per million pop., 2003
		business enterprise	government	higher education					
EU-25	1,178,237	49.0	13.4	36.5	430,057	88,100	2.72	38.31	639
US	1,261,227	80.5	3.8	14.7	185,400	46,000	1.96	31.10	809
Japan	675,330	67.9	5.0	25.5	172,209	14,500	1.39	9.55	569
Germany	264,721	58.1	14.7	27.2	72,004	23,000	7.55	8.40	772
France	186,420	51.1	12.9	34.1	63,569	8,400	1.44	6.08	773
UK	157,662	57.9	9.1	31.1	49,033	14,900	2.48	8.49	1,086
Italy	71,242	39.3	19.0	39.7	28,283	6,400	2.20	4.62	611
Sweden	45,995	60.6	4.9	34.5	15,868	3,600	7.30	1.94	1,642
Netherlands	43,539	46.9	15.6	36.4	15,848	2,600	2.91	2.51	1,177
Finland	41,724	56.6	11.3	31.2	13,018	1,800	4.66	0.96	1,397
Denmark	25,130	59.7	9.3	30.5	7,665	900	2.11	1.03	1,457
Austria	24,124	66.3	4.1	28.9	6,972	2,200	7.53	1.02	959

Source: EUROSTAT databank and CORDIS 2005 (LYA = Last year available, FTE = full-time equivalent)

The structure of Germany's research landscape

One way to characterise the structure of the research landscape in Germany is to view it from the perspective of expenditure on research and development (R&D). Since a major cost factor of any R&D work is staffing, this perspective also indicates the contours of the research infrastructure in terms of human resources (see Figure 2).

Figure 2: Research and development expenditure by source and performer (2003)



Source: BMBF, 2005: Tables 3 and 21a, own calculations

In the year 2003 approximately 54 billion euros were invested in R&D in Germany, two-thirds of which was funded by business and industry. In sum, this investment in R&D amounted to roughly 2.5% of gross domestic product (GDP). This is broadly comparable to Germany's counterparts (France 2.3%, UK 1.9%, but Sweden 4.3%).

The performers of R&D are also shown in Figure 2. It can be seen that business and industry is not only the biggest investor, but also the biggest performer of R&D. However, this research focuses on commercial application. This discussion focuses on institutions of higher education (HEI) that largely receive public funding to carry out public, in the first instance non-commercial, research.

The higher education sector in Germany is made up of roughly 230 full universities and so-called universities of applied science (*Fachhochschulen*). The universities are major players in terms of research and development, particularly because of the diversity and regionality of their activities. HEIs have been able to increase their expenditure on R&D in the decade between 1992 and 2002 from 6.6 to 9.1 billion euros (+38%). One-third of this expenditure is financed through institutional funding by the individual *Länder* for “their” HEIs.

Two-thirds of R&D expenditure at HEIs is provided through third-party funding contracts from industry, government and – especially for universities – from the German Research Foundation (*Deutsche Forschungsgemeinschaft* – DFG) (BMBF, 2005: Table 20b). The DFG is funded jointly by the Federal Government and the *Länder* at a ratio of 58:42 (Waugaman *et al.*, 2004: 50ff.). Around 40% of all third-party funding received by HEIs comes from the DFG. The largest proportion of DFG grants (c. 60%) is for the general promotion of research. These funds are allocated either as grants for projects proposed by individuals or a group of academics. Collaborative research programmes involving multiple projects in a broad subject area (*Sonderforschungsbereiche*) account for a further c. 28% of DFG expenditure. To promote prospective academics, the DFG also funds research training groups (*Graduiertenkollegs*).

The lower half of Figure 2 shows that besides institutions of higher education, four non-university research organisations play a significant role in the German research system, accounting for roughly 10% of all R&D expenditure. Each of these research organisations has a different profile, both in terms of mission and in terms of sources of institutional and contract funding (Waugaman *et al.*, 2004: 56ff.).

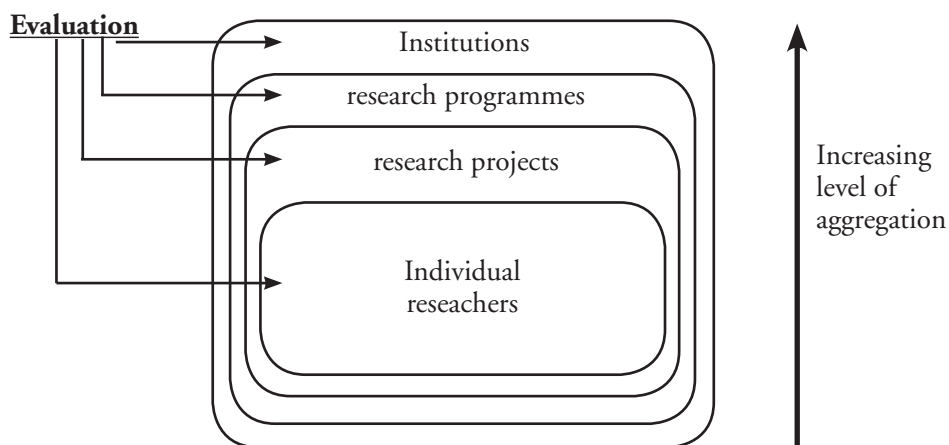
The Federal Government and the *Länder* fund both the Helmholtz Centres and the Fraunhofer Institutes at a ratio of 90:10, whereas the other two organisations are supported by equal grants from these different levels of government. This institutional funding accounts for between 60% and 75% of the organisations’ total income, with the exception of the Fraunhofer Institutes, which carry out a lot of applied research and earn over two-thirds of their income through contracts.

The research landscape can, in sum, be characterised as complex, both with regard to institutions – universities, *Fachhochschulen*, and the individual institutions of the four non-university research organisations – and their funding. In each case institutions are funded by a mix of Federal and *Länder* institutional grants, which are supplemented by third-party funding from various sources, particularly the DFG. It becomes clear that on system level there are 18 major actors with strong interests in implementing research evaluation in order to make the effectiveness and efficiency of research activities more transparent: the Federal state, the 16 individual regional states (the *Länder*) and the German Research Foundation (DFG).

Research evaluation in Germany

There are a number of specific starting points for research evaluation, beginning with individual researchers and ending with system evaluations of groups of institutions (see Figure 3). The choice between these points is influenced by two major factors: the purpose of the evaluation and the organisation evaluating research activities.

Figure 3: Levels of aggregation of research activities and starting points of evaluations



Source: Adapted from Kuhlmann and Heinze, 2004

A number of multi-dimensional models for the characterisation of evaluation procedures have been developed in the literature, which are framed by at least four different purposes of evaluation (cf. Ewell, 1999: 193ff; Harvey and Askling, 2003: 73ff; Kuhlmann and Heinze, 2004: 65f; Orr, 2005b: 112ff): to provide accountability for activities, control for threshold standards, improve performance and judge the efficiency or effectiveness of processes. However, it may be most useful to differentiate between two more fundamental purposes in this discussion:

- Evaluations with an *allocative function*: Assessment of academics' research efforts and their resulting publications has been carried out more or less since the birth of modern science (De Groff *et al.*, 1998: 119). In most cases the assessment has led to clear consequences: evaluations by academic peers are used to decide on the assignment of professorships or on the merit of publishing an academic paper. In this way the assessment leads to the allocation of academic status. A further example of such science-immanent procedures is the assessment of a project proposal in order to judge the value of funding the implementation of the proposal.

Within the New Public Management programme, a further type of allocative evaluation has developed (Leszczensky *et al.*, 2004). An assessment of past performance (so-called ex-post assessment) at institutional level is used to inform or even determine public allocation of institutional budgets. New Public Management aims to afford institutions the greatest possible autonomy to make decisions on inputs and processes and judges only the merit of the outputs that result from these autonomous decisions. By this, it re-forms an academic tool, which is based on a low aggregation level (researcher, project), to an administrative tool, focused on a much higher level of aggregation – groups of institutions (see Figure 3). The higher up the level of aggregation a procedure is, the more important it becomes that the procedure is transparent and the evaluative criteria are standardised.

- Evaluations with a *cartographic function*: One purpose of research evaluation can be the comprehensive description of a research landscape to make the scale, scope

and quality of such activities more visible. In the past, such studies at this level of aggregation have often been implemented on an *ad hoc* basis. Over the last 20 years in Europe it has become increasingly common to carry out such studies on a regular basis with information, which goes beyond pure input statistics. In both cases (*ad hoc* and regular), the evaluation can result in recommendations for the future, which flow into decision-making processes, but do not themselves determine specific decisions. Depending on the depth of the information required, only standard quantitative indicators (input and output variables) may be used or these may be supplemented by non-standardised information, including contextual information and opinions of the persons who are being judged.

Since regular cartographic evaluations are essentially new to the scientific domain, they can be very controversial. Non-standard procedures that lead to no specific consequences are often open to criticism on the basis of the time and effort put into them and their ambivalent value.

The dichotomy suggested above is not always clearly found in practice, since some procedures attempt to combine both functions. However, in general, procedures will emphasise one of these functions over the other. In Germany the following actors and procedures are to be found in the field of research evaluation:

- The German Research Foundation (DFG) uses a number of different formats for research evaluation. In line with its function as a funding body for research, the DFG has a long established procedure for the evaluation of proposals for funding, which is based on peer review of both past performance of academics and their institutions and the merit of their proposals. However, it has recently been recognised that the effectiveness of the projects and programmes funded by the DFG should also be assessed. A new body set up at the start of 2006 and called the Institute for Research Information and Quality Assurance (IFQ) will take up this task. Additionally, at the end of the 1990s the DFG recognised a need to produce a cartographic analysis in the form of a ranking. This periodic ranking of research activities funded through public third-party funding (1997, 2000, 2003) analyses funding streams at institutional, regional and discipline level. In 2003, the report focused on the network character of research and ranked *inter alia* research institutions according to the number of collaborative projects in which they were involved.

A further development of the DFG evaluative tasks is the implementation of the so-called Excellence Initiative, which the DFG jointly leads together with the Science Council (*Wissenschaftsrat*). This evaluation is essentially allocative. It is based on components of the DFG procedures, but extended to include institutions' strategic development plans. The Excellence Initiative will be implemented for the first time in the autumn of 2006. Between 2006 and 2011 the programme will allocate €1.9 billion shared between both individual and collaborative programmes from universities and non-university research organisations, based on their excellence in graduate (doctoral) teaching and world-class research. For those universities successful in both areas, a third stream of funding to promote institutional strategies for top-level research will be

available. It remains to be seen, what exact data and quality criteria will be used by the peers in the various assessment committees, although this is essentially a combination of existing procedures, all based on proposals written by applicant institutions.

- The Science Council is an agent of the state, funded by both the Federal Government and the *Länder*. It has the specific task of making recommendations on the structure and performance, planning and development of scientific institutions and this also involves cartographic evaluations based on institutional self-evaluations and peer review. However, these evaluations are carried out irregularly and are initiated on an *ad hoc* basis by the Federal Government and/or the *Länder*.
- Each of the associations responsible for the non-university research institutes began to carry out its own evaluation of institutes by, at the latest, the start of 2000. These occur on a regular basis and are based in the main on peer reviews by academics from outside the respective associations, including foreigners. They are generally carried out both to inform decision-making within the associations and to report to external grant-givers on the performance of the associations and their respective institutes. For example, the Senate of the Leibniz Society, which has only external members, carries out regular evaluations of the Leibniz Centres and uses these to make recommendations to the Federal Government and the *Länder* on the funding of the centres and on their profiles.
- At regional state level a number of *ad hoc* reports have been commissioned in the recent past to evaluate the current structure of higher education performance (including research) and to make recommendations for the future (e.g. in Hamburg, Schleswig-Holstein, North-Rhine Westphalia and Bavaria). Besides these one-off exercises, *Länder* such as Baden-Württemberg and Lower Saxony have introduced regional evaluation bodies, whose purpose is to inform both individual institutions and the state on comparative performances. Research evaluation is carried out in Lower Saxony by the Academic and Research Commission. This institution has been until now unique in Germany, although the establishment of such a body was recently recommended for Bavaria (Mittelstraß, 2005) and has been discussed in other states.

In sum, the picture is of a plethora of procedures, which at the aggregation level of regional state or the whole of the Federal Republic rarely have direct allocative functions, but are used to inform policy and decision-making processes. The Academic and Research Commission is a particularly interesting actor within this framework, because its research evaluation procedure occurs regularly and aggregates to institutional levels. It encompasses research at both universities and non-university research institutes financed by the regional government.

Case study: Research evaluation by the Academic and Research Commission Lower Saxony (ARC)

Context

Lower Saxony is one of the largest states in Germany and is comparable in land area and population to small European states. Its land area is, for example, similar to that of Belgium

and its population to that of Austria. The region has 13 public HEIs of university character, five public universities of applied science (*Fachhochschulen*) and roughly 30 non-university research institutions (financed by public funds at Federal, state or regional level).

Lower Saxony has assumed a pioneering role with the realisation of separate quality assurance systems for studying and teaching and for research, designed to assess long-term and region-wide effects. This has been achieved by:

- founding the Central Evaluation and Accreditation Agency (*Zentrale Evaluations – und Akkreditierungsagentur ZEvA*) in 1995 following a decision by the State University Conference. Alongside its function as an accreditation agency for new courses, it is also responsible for the assessment of the teaching and learning landscape in Lower Saxony.
- founding the Academic and Research Commission (ARC) in 1997. This body assumes an advisory role for the regional state parliament and includes amongst its other responsibilities quality assurance through research assessment. The Commission was preceded by intensive negotiations between the Ministry and State University Conference, as its installation entailed a certain realignment of responsibilities and influence for the respective governing bodies.

This meant the realisation of a conscious division between teaching and learning on the one hand, and research on the other, otherwise united in accordance with the Humboldt Principle. The objective was to achieve accurate and specific results for each sphere of activity, which necessitated, for example, different criteria of evaluation.

The core task of the ARC is to advise the regional state parliament and its academic facilities on questions pertaining to higher education policy and to develop recommendations for the structural development of the universities and research institutions of Lower Saxony. The assessment competency held by the ARC is thus clearly distinct from the decision-making competency held by the Ministry and the HEIs. The quality assessment procedure has so far been the most significant instrument in making such recommendations. Furthermore, the ARC is committed to developing forward-looking concepts for the regional research landscape.

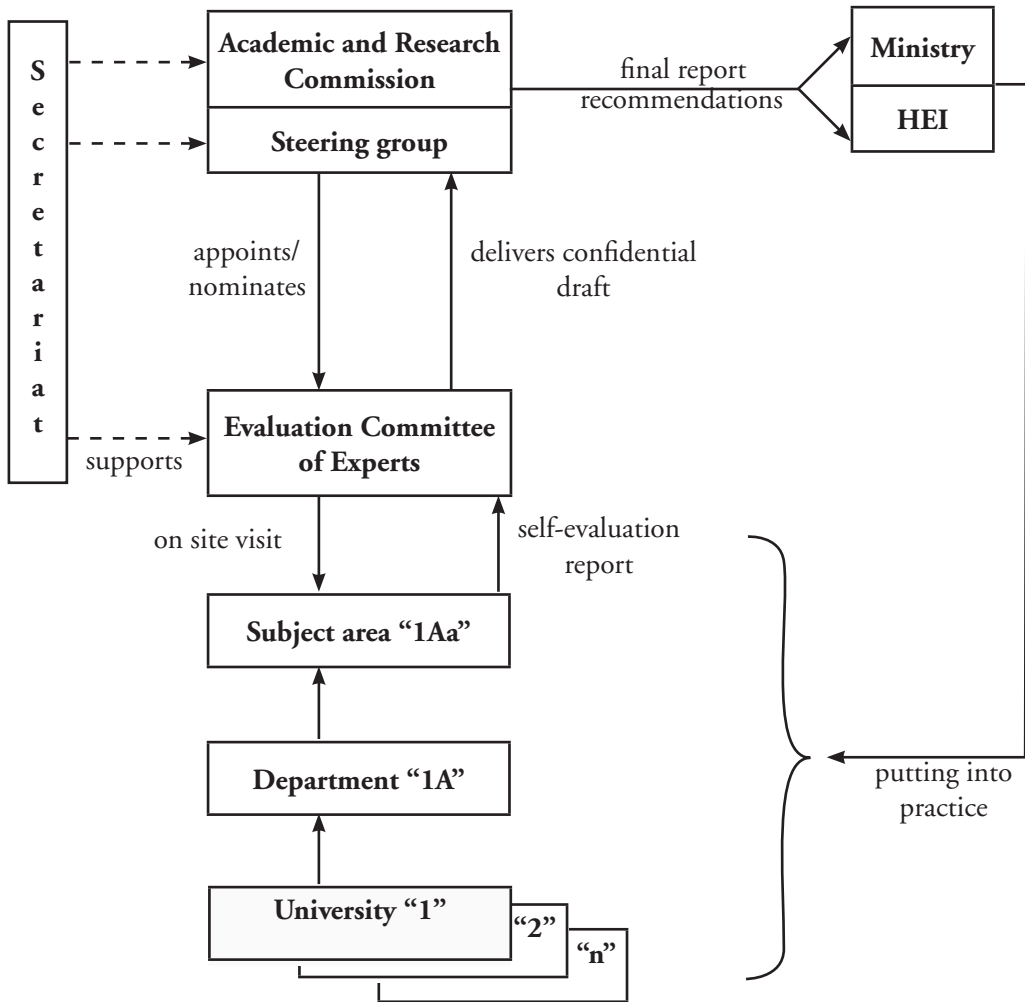
Structure of the procedure

Within the framework of quality assessment, systematic, cross-regional and comparative research evaluation has been carried out since 1999 at universities and other publicly funded non-university research institutions. The aim is to achieve valid statements regarding the quality of research within institutions and subject areas in Lower Saxony and to summarise the results cartographically.

In drawing up the basic structure of the procedure, the authorised bodies agreed on a multi-dimensional, mainly qualitative procedure, to be applied in the form of peer reviews by subject area. To this end, the ARC appoints panels of experts external to the Commission.

The procedure of informed peer review is run in accordance with internationally recognised standards, albeit with regional variation (see Figure 4). A self-assessment report from the institution to be evaluated, portraying the achievements of the last five years as well as future plans and perspectives, is followed by a visit to the institution by the panel of experts. Talks thus take place at the institution with the HEI directors, the faculty directors and finally with each individual research unit (these usually consist of single professorial chairs). The procedure can, therefore, be characterised as an in-depth evaluation from institutional down to individual research level.

Figure 4: Structure of ARC evaluation procedure



Source: Adapted from Orr, 2003:32

The experts on the panel draw up a confidential draft assessment based on their impressions from the institution's report and their site visit. The directors of the HEI assessed may then issue a statement on the draft. The experts are, in turn, obliged to respond to criticism or recommendations contained in the HEI's statement. The ARC is presented with the final report from the expert panel as well as the statements of the assessed institutions. It decides on

a definitive version of the report, which is then in confidential form, i.e. containing the names of individuals, passed on to the Ministry and the directors of the assessed HEIs for their further use. A version of the report from which persons' names have been removed is published and made generally accessible (e.g. on a dedicated website).

The reports contain, on the one hand, qualitative assessments of the achievements of individual research units and of the subject area within the HEI as well as on a regional state level, and, on the other hand, the reports link these assessments to recommendations, equally related to these levels. These recommendations may touch upon such questions as the denominations of professorial chairs about to be vacated at research unit level; on the next level, the HEI may receive suggestions to restructure a subject area. From time to time, relocations from one site to another may be recommended in order to reinforce the HEI's strengths, concentrate research potential and to raise achievement capacities and thus national as well as international prestige in the long term.

These – sometimes extensive – recommendations can, in turn, be implemented only at a variety of levels and, particularly if the recommendations take on greater proportions, through the interaction of several agents. They require discourse and negotiation, not only within the HEIs themselves, but also between the HEIs and the Ministry. Instruments of implementation may include review committees and structural committees, for instance, coordinated by the ARC. In the last few years, the most important means of steering between the Ministry and HEIs have been joint two-yearly target agreements (*Zielvereinbarungen*).

Criteria

The ARC has laid down criteria, comparable to other national and international evaluations, on which the procedure is to be based. These criteria are intended to ensure recognition for the procedure and to allow a comparison of the results. On the one hand, it is a question of quality and relevance of the research results; on the other hand, effectiveness and efficiency within the research process take precedence. Each panel of experts appointed by the ARC is, therefore, required to answer the following questions, ranging from the level of individual research units to the subject area within the HEI and its achievement capacity on a regional level:

- What is the contribution of this research towards the performance of the relevant discipline within the HEI as well as on a regional, national and international level?
- Are funds implemented to achieve the intended effects whilst protecting standards of quality? How do these results compare to those of other locations?

Alongside the general criteria common to all procedures, particular significance is attributed to the definition of subject-specific criteria and the formulation of a subject-specific research concept. The experts in the panel are at liberty to assign varying significance and expectations of success to individual indicators in the relevant subject areas, particularly quantitative indicators such as third-party funding and international publications. The agreed assessment guidelines are thus applied uniformly across the regional state within the framework of procedures for the particular disciplines.

Results

By spring 2006, 27 research evaluation procedures and nine evaluations of non-university research institutions have been completed. Furthermore, follow-up reports documenting the status of implementation in the HEIs are available on the first assessment procedures for eight subjects. The first assessment cycle will be completed in 2007, accompanied by an analysis of the procedure (i.e. review of the evaluations). First conclusions may, therefore, be drawn from the procedure in Lower Saxony.

Acceptance of procedure

The multi-dimensional, mainly qualitative procedure of research evaluation has met with acceptance both from the Ministry and from the HEIs. This acceptance can be traced back to several factors:

- The research evaluation procedures and their results have demonstrated that assessments of achievement must be organised in a manner appropriate for academic institutions and must make use of recognised specialist academic expertise. The appointment of renowned experts is indispensable for the acceptance of assessments and recommendations. Up until now, the ARC has been successful in this regard.
- The assessment of performance must be flexible enough, even within standardised procedures, to take subject-specific factors into account; acceptance within the scientific community will otherwise be limited. This point is of particular importance to the procedures of the ARC. Subject-specific criteria are described transparently in a chapter especially for this purpose in each report.
- The assessment of performance must display high standards of precision and accuracy. The recommendations drawn from the assessment must be sound and realistic as well as forward-looking. In the ARC reports, the precision of the assessments is reflected at a variety of levels (research unit, subject area on site, subject area on a regional level) in objective and valid recommendations for implementation at the various levels of responsibility.
- Recommendations must involve a follow-up. Part of the standardised procedure demands that the ARC request an interim report on the status of implementation after three to four years. Experiences from eight interim reports indicate that, in most cases, progress has been made concerning the implementation of the recommendations.
- Research evaluations should be clearly distinguishable from other quality assurance procedures through their aims and instruments and they should make their own specific contribution to quality assurance. As a multi-dimensional procedure, the method of peer review can lead to in-depth research assessment, in contrast to the results of rankings and other performance comparisons restricted to only a few indicators and/or to a single level of analysis.

Limitations of the procedure

The multi-dimensional procedure, however, also meets with limitations:

- Quality assurance measures entail a cost in terms of time and effort for the evaluated institutions and distract academics from their core activities of teaching and research. This burden should be limited and can be reduced to some extent through a greater professionalism in the internal management of a university, in which some universities have made more progress than others.
- The increase in quality assurance measures is rendering access to high-ranking experts more difficult. It is, therefore, necessary to deal efficiently with the capacities of expert panels members for evaluation processes.
- A further issue is the extent to which the long-established institutional structures within subject areas, which are the basis of the research assessments, reflect the current structure of research activity. This may affect individual researchers working at the margins between subject areas or whole disciplines, which have a multi-disciplinary or interdisciplinary structure. In Lower Saxony, challenges have been in the subject areas of Education Science and Environmental Sciences. Efforts must be made to adequately reflect the area being evaluated whilst maintaining the practicability of the evaluation.

Implementing recommendations

The advice and recommendations of the expert panel are used to aid decision-making processes within the HEIs, as well as in the structural planning of the regional state.

The chances of success in implementing recommendations are dependent upon a range of factors, which can mostly be categorised into either the degree of complexity of the recommendations or their costs. If recommendations have not been implemented in the past, it is primarily due to further deterioration in the basic conditions for researchers (such as budget cuts) or due to the fact that multiple layers of interest have exacerbated the process. Because of the highly restricted human resources at HEIs, impasses are swift to arise, especially if aspects of teaching or other service functions demand simultaneous consideration.

Planning security for HEIs over a longer period of time, such as recently implemented in Lower Saxony with the five-year “pact” between all HEIs and the Ministry and the multi-year target agreements for individual universities, can therefore contribute considerably to a constructive treatment of the recommendations of the expert panels. Recent assessments of the procedure have clearly demonstrated that these new instruments of steering and coordinating higher education policy will enable recommendations to be considered even more consistently than in the past.

Perspectives

Despite these positive results the ARC may not conduct a second cycle of evaluation in this form and with this structure. This is primarily due to considerations of time and effort and also to avoid the dangers of routine and adaptation.

The analysis of this procedure and the multi-regional debate on procedures of quality assurance clearly show that the burden for the HEIs as well as for the panels of experts has grown to an extent that is no longer acceptable. The objective of achieving scientific adequacy has led to a burdensome procedure with less than clear outcomes. Furthermore, the HEIs, including the non-university research institutions, are increasingly taking on the responsibility for the quality of their performance and developing their own evaluation procedures, with the result that the external assessment procedures carried out up until now have been an important initial stimulus, but need not be continued with the same depth. Instead a change of focus is being considered, with less emphasis on the cartographic and more on the allocative function of the procedure. In this case, the purpose would be, for example, to shape decisions on the research profile of a specific institution or decisions on where to set up centres, institutes, graduate programmes or graduate schools.

Prospects for research evaluations in Germany

The research evaluation landscape in Germany must be classified as fragmented, with many actors using different procedures and covering different institutions or regions. The ARC can be seen as a role model for regional evaluation procedures and with the further implementation of the New Public Management programme in Germany, it can be assumed that the other big German states (i.e. Bavaria, Baden-Württemberg, North-Rhine Westphalia) will introduce their own ARC. However, the regional character of Germany's higher education system is likely to become further pronounced under Federal reforms of the present government and so differences between states will remain.

Initiatives to counter this fragmentation are likely on the Federal level with reporting systems based on higher aggregation levels and the mix of quantitative and qualitative information. Indeed the Science Council has just begun subject reviews covering the whole of Germany and the DFG, German Science Foundation, and the Science Council played a significant role in the establishment of the Institute for Research Information and Quality Assurance (IFQ). A further stimulus may also be expected as a by-product of the Excellence Initiative, which strives to evaluate in a comparative way the quality of research performance and future research strategies.

References:

- Bundesministerium für Bildung und Forschung (BMBF) (2005). *Forschung und Innovation in Deutschland [Research and Innovation in Germany]*. Retrieved April 6, 2006 from the World Wide Web: http://www.bmbf.de/pub/forschung_und_innovation_05-07.pdf
- Community Research & Development Information Service (CORDIS) (2005). *European Research Area Key Figures 2005*. Retrieved April 6, 2006 from the World Wide Web: ftp://ftp.cordis.lu/pub/indicators/docs/2004_1857_en_web.pdf
- De Groff, J., Neave, G., & Svec, J. (1998). *Democracy and Governance in Higher Education* (Council of Europe) series: Legislating for higher education in Europe. Den Haag: le luwer Law International.
- European Commission (2002). *Education and Training in Europe – Diverse systems, shared goals for 2010*. Retrieved April 6, 2006 from the World Wide Web: http://europa.eu.int/comm/dgs/education_culture/publ/pdf/educ-training/en.pdf
- Ewell, P.T. (1999). Linking Performance Measures to Resource Allocation: Exploring unmapped terrain. *Quality in Higher Education*, 5 (3), 191–209.
- Harvey, L. & Askling, B. (2003). Quality in Higher Education. In R. Begg (Ed.), *The Dialogue between Higher Education Research and Practice* (pp. 69–83). Dordrecht: Kluwer Academic Publishers.

- Kuhlmann, S. & Heinze, T. (2004). Evaluation von Forschungsleistungen in Deutschland: Erzeuger und Bedarf; Teil 1: Konzeptionelle Grundlagen [*Evaluation of Research Performance in Germany – Producers and needs; Part 1: Concepts*]. *Wissenschaftsrecht*, Volume number? (37), 53–69.
- Leszczensky, M., Schwarzenberger, A., Orr, D., & Weitz, B. (2004). Staatliche Hochschulsteuerung durch Budgetierung und Qualitätssicherung: Ausgewählte OECD-Länder im Vergleich [*State Higher Education Steering through Budgeting and Quality Assurance – Selected OECD countries in comparison*]. Hanover: HIS.
- Mittelstraß, J. [Chairman] (2005). Expertenkommission Wissenschaftsland Bayern 2020 [*Expert Commission Science State Bavaria 2020*]. Bayerisches Staatsministerium für Wissenschaft, Forschung und Kunst, München. Retrieved April 6, 2006 from the World Wide Web: http://www.stmwfk.bayern.de/downloads/wl2020_2_analysen.pdf
- Orr, D. (2003). Verfahren der Forschungsbewertung im Kontext neuer Steuerungsverfahren im Hochschulwesen: Analyse von vier Verfahren aus Niedersachsen, Großbritannien, den Niederlanden und Irland [*Procedures of Research Evaluation in the Context of New Steering Mechanisms in Higher Education: Analysis of four procedures from Germany (Lower Saxony), United Kingdom, Netherlands and Ireland*]. Hanover: HIS.
- Orr, D. (2005a). Can performance-based funding and quality assurance solve the state vs. market conundrum? *Higher Education Policy*, Volume number? (18), 31–50.
- Orr, D. (2005b). Hochschulsteuerung und Autonomie englischer Universitäten – Hochschulfinanzierung und Qualitätssicherung aus einer Verfahrensperspektive [*Steering Higher Education and Autonomy of English Universities – Higher education funding and quality assurance as organisational procedures*]. Münster: Waxmann.
- Orr, D., Jaeger, M., & Schwarzenberger, A. (In press). Performance-based funding as an instrument of competition in German higher education. *Journal of Higher Education Policy and Management*.
- Waugaman, P.G., Friedrich, H-R., Tornatzky, L.G., & Schmidt, H-U. (2004). *Competitiveness in Research – A comparative pilot study*. Hanover: HIS.

Note

1. The ISCED (International Standard Classification of Education) categories distinguish between different courses of study at different levels and typically with different types of content. ISCED 5 refers to the first stage of tertiary education and ISCED 6 to advanced research qualifications.