ASSESSMENT OF RESEARCH IN GERMANY ILLUSTRATED BY THE EXAMPLE OF THE 84 INSTITUTES OF THE LEIBNIZ SOCIETY

di Dietrich Wegener

Most of the more than 80 institutes of the LO have been evaluated three times by the GSC and by the senate of the LO respectively. The aim of these evaluations, the procedures developed and the results achieved are described. The steps being essential to arrive at an effective evaluation are discussed.

1. Introduction

The research assessment in Germany like in other countries has a long tradition based on an ex ante system realized by the recruitment procedure of the scientific management personal. The evaluation of the Academy Institutes of the former DDR by the German Science Council (GSC) with its strict selection procedures, based on peer review of their antecedent performance, has enforced a paradigm shift. Nowadays *ex post* assessments of Research Institutes has become an important steering element in science policy. This holds to a different extend for the Universities and Polytechniques, the Max-Planck and the Fraunhofer society, the Helmholtz association [1] and the Leibniz organization (WGL); the latter nowadays constitutes the organizational framework of the so called Blue List Institutes. In this paper I will describe the procedure applied in the assessment of the 84 Institutes of the WGL I am most familiar with. Moreover, the steps of the evaluation procedure are well defined, have turned out to be very effective in the sense that the results of the evaluation process are accepted by the Institutes themselves and by the funding ministries. In addition most of the Institutes have been assessed since 1981 three times, hence experience exists to which extend improvements and changes suggested have been implemented and how effective they were to improve the performance of an institute. The field of activities covered by the Blue List Institutes is broad; it extends from humanities, social science to science and engineering. In addition service institutes for research as well as research museums are members of the WGL. The arguments given in this paper primarily apply to research institutes, but experience shows that to a large extend they also hold for the other Institutes [2].The first two assessments have been performed in the years 1981-1994 and 1995-2000 [3] respectively by the GSC. The current one has started in 2001, it is supervised by the senate of the Leibniz organization and will be finished in 2009. In this case care has been taken to minimize the organization's influence on the results of the assessment. While for the 1981-1994 evaluation admistrative criteria as independence of the research program, the super regional impact of the results achieved and their relevance for national science policy were the dominant criteria to extend the promotion of an institute, the scientific quality of its research is now the primary necessary criterion since 1995. The evaluation aims for a higher flexibility of the system, in the worst case the end of government aid is foreseen, the released money used to finance other Institutes of the Leibniz organization positively evaluated; moreover the possibilities to increase of the percentage of fixed term positions of an Institute are analyzed to improve its flexibility. In addition the evaluation urges more influence of the scientific advisory committees, the strengthening of the cooperation with the Universities and the increased engagement in setting up international networks.

A short description of the work, its composition and the basic procedures of the GSC will help to understand better the assessment procedure. It was founded 1957 as the central advisory committee of science policy in Germany. 32 members of the academic community belong to it, each holding one personal vote. 24 of the members, leading scientists of the major research disciplines, are appointed by the German president from a list of candidates suggested by the major science organizations; 8 personal members represent public life and industry. 32 votes are held by the representatives of the Federal and the States ministries, while 16 votes are assigned to the Federal government, the rest is shared by the representatives of the 16 State governments. Decisions of the German Science Council are taken in a three step process. First the Academic Commission considers a report, usually in a very

detailed and critical analysis of the document and votes upon it. The report is prepared by a subcommittee, which is composed of members of the Academic Commission of the council, admistrators of the Federal and the State governments and external experts. In a second step the Adminstrative Commission analyzes the report and votes upon it. Finally the decision is taken by a vote of the General assembly; a 2/3 majority of the 64 votes is necessary for the acceptance of a recommendation. Since politics participates in the decision finding from the beginning the realization of the proposed actions is eased. Note, however, that politics does not dominate the procedure.

2. Evaluation by the German Science Council 1995-2000

The evaluation process consists of 4 steps indicated in fig.1. The recommendation has to be prepared for the Bund-Länder-Konferenz (BLK) which is the decision taking body. Members of this body are high ranking representatives of the Federal and the State ministries of research. The BLK has asked the GSC 1994 to evaluate within a 5 years period the 82 Institutes of the Blue List (Blaue Liste), which all in the meantime joined the Leibniz organization. In addition the GSC was asked to investigate the quality and the importance of the research of 8 further Institutes with the aim to accept them for the funding system valid for the Blue List in case that the funding of other Institutes terminates.

The GSC set up the so called Blue List committee (BLC) to steer the evaluation of the Blue List Institutes and to draft the recommendation report. The evaluation report itself was prepared by the evaluation group (EG) which included, beside two members from the BLC and one of the GSC, experts working in the field of the institute scrutinized. Also two members of the BLC representing the federal and state governments participated in the evaluation. The average number of participants varied between 6 and 20 depending on the size of the institute and the research fields to be covered. The group was chaired by an academic member of the BLC. A comprehensive list of questions providing qualitative and quantitative information of the research work was responded in advance by the institution under scrutiny. Typical examples of information supplied in advance by the Institute are collected below. Note that this information was only used as the basis of the peer review taking place during the visit of the institute by the EG.

- **General information**
  - Central research fields of the institution, most important results in the last 5 years
  - Arguments why the research could not be performed at a University
  - Specific strength of the Institute and weaknesses observed and highlights of research results during the evaluation period
  - Foreseen development of the Institute and its research expected on medium term

- **Structural characteristics of institute**
  - Organigram
  - Service for research provided by the Institute
  - Information concerning composition and work of the scientific advisory committee, minutes of the meetings of the last 3 years have to be provided.
  - In which way quality management is realized.

- **Means available**
  - Basic institutional funding, third party money from DFG, EU, industry, foundations
  - Space allocation, IT technique
  - Detailed information concerning the personal available (researchers, technical and administrative staff, separate for permanent and fixed term staff)
  - Specific problems to fill staff positions

- **Cooperation**
  - Promotion of young scientists
  - National and international cooperation
  - Information concerning visitors from outside (numbers for recent years, time they stayed, scholarships)
  - Teaching at Universities

- **Results achieved**
  - List of publications in refereed journals and monographs, conference contributions, 10 publications of the last 7 years with the highest impact on the field
  - Knowledge transfer: peer assessment for public funding agencies, policy advice
  - Technology transfer: applied and assigned patents, licenses granted, cost and royalties, spin offs
  - Organization of international and national conferences
  - Invited talks at international and national conferences

– Number of thesis supervised
– Appointments of staff at other institutions
– Prizes and honors

Clearly not all of the criteria can be applied with the same weight to each Blue List Institute.

The publications habits in science differ from those in engineering as well as from those in the humanities and social sciences.

The peer review allowed to arrive at a fair consideration of the criteria of the different research fields. It was interesting to note that also the publication habits in a field can change with time: in 1996 the reviewers stressed in case of an institute in social science the importance of monographs while in 2004 priority was assigned to papers in refereed international journals.

Looking back to more than 100 evaluation reports I had to appraise in the last 12 years as generalist I found the following information especially instructive. Basing my judgment on them and comparing it with the finding of the experts I observed a high degree of agreement. Institutes with a well defined coherent research program, integrated in international networks usually were able to raise appreciable funds in competition and published in high ranking journals. Their staff frequently was offered appointments by other institutions and was invited for talks at major international conferences. Quite often they were able to quote a result which has influenced their research field in a way forcing the international community to follow their line: they were able to define a fashion and not only to follow it.

Of the 82 Institutes evaluated 8 showed such an unsatisfying performance that the GSC proposed in its recommendation to stop the common funding of the Institute by the Federal and the State governments.

Also in this case some common features were observed. In most cases many small groups worked on different problems and no coherent research program was visible. They had many permanent positions specifically also such funded by third party money. Hence they were forced to orientate their research program according to the guidelines defined by a multitude of funding agencies. A coherent focusing of the Institute's research program turned out to be impossible in these cases. Moreover some of these Institutes missed the latest development of the field, pottering about their old research problems and using outdated techniques; in addition the control by the board of trustees and the scientific advisory committee respectively failed. Furthermore, there seems to exist a kind of scientific policy dyslexia: the chairman of the board of trustees of an institution, undersecretary of state, after taking note of the very critical evaluation report came to the following conclusion "the institute has been certified as outstanding". This was not the only case of fatal involvement and misjudgment of politics observed in the course of the evolutions!

The evaluation report prepared by the EG was based on the written information provided by the Institute and the information collected during the two days visit of the Institute. The same procedure of the visits was followed to ensure that all institutes were treated in the same way in order to avoid legal objections at later stages of the evaluation process. The visit started with an internal discussion of the evaluation group plus guests. In this phase open questions were formulated and critical items were raised such that the members could try to collect information for their final judgment. Next followed a meeting of the EG with the scientific and the administrative director and the group leaders of the institute. In addition the chairman of the scientific advisory board participated. After a short presentation of the Institute by the director the EG took the chance to clarify open questions which were raised during the internal discussion and by the presentation of the director. It followed an extensive visit of the Institute which allowed discussions with the researchers and technicians at their workplace. Depending on the the number of research fields of the Institute the EG splitted up into subgroups. Usually details of the research results achieved were presented by posters, hence an in depth analysis of the quality of work by the experts of the EG was possible. Moreover the administration, the library and the IT-group were visited.

An important and from my experience an extremely instructive part was the discussion with the staff of the Institute excluding the directors and the group leaders. Often problems otherwise hidden became evident during these discourses. Next followed a meeting of the EG and guests where first conclusions on the quality of the research and the organization of the Institute were given. Furthermore the guests from the supervising Federal and State Ministry as well as the chairman of the scientific advisory committee of the Institute were asked to summarize their impressions. A one hour interview followed with representatives of cooperating Universities and in some cases also with collaborating non University Institutes. This discussion

was of special importance since a high priority aim of the evaluation process was the strengthening of the connections between the Blue List Institutes and Universities. A final discussion of the EG with the Institutes management followed in order to clarify problems still open. According to my own experience it was preferable to invite only the director for this discussion since under this circumstances he could point out internal problems which otherwise would have been left hidden.

In the final session the EG without guests each member summarized his judgments on the scientific work of the Institute, its strengths and its weaknesses. Moreover the EG analyzed to which extend actions proposed by the previous evaluation group were realized. All members were requested to provide their assessment in written form to the office of the GSC. These written summaries and the discussion results documented in the minutes of the visit constituted the basis of the evaluation report. A draft of the report was formulated by the office of the GSC, distributed together with the minutes to the EG members, who got the possibility to apply corrections on both. A revised draft was prepared next to be agreed by the members. Usually this procedure converged fast. Only in two cases the report was not accepted unanimously and minority votes were formulated. The evaluation report once agreed upon by the members of the EG was not changeable during the following steps of the procedure. This turned out to be of essential importance because otherwise in case of negative judgments clearly the Federal and State Ministry concerned were tempted to improve the picture of the Institute drawn by the EG. They never were successful in case of the evaluation performed in the years 1995-2000 by the GSC.

In preparation of the meeting of the BLC the final evaluation was distributed to the members of the BLC and to the responsible Federal and State ministry. In a three step process the recommendation was prepared based on the unchanged evaluation report. In a first step the BLC analyzed the evaluation report comparing the assessment with that of other Institutes in the research field. This allowed to minimize the influence of possible biases. In case the EG arrived at a positive judgment of the research quality of the Institute, its super regional impact and importance for national scientific policy was analyzed. In the second step a hearing of admistrators from the Federal and the State government took place. In this step also objections of the Institute could be raised. These discussions in some cases were quite emotional from the side of the officials. In these cases the sessions had to be chaired with great caution to avoid reasons for legal steps. Finally the results were summarized in a short recommendation which combined with the evaluation report of the EG. This report was presented to the GSC which voted on it as described in section 1. Under any circumstances for each Institute an explicit argument had to be given why the work of the Institute could not be performed at a University.

The overall statistics of the 5 years work was extremely positive. All recommendations were accepted by the General assembly of the GSC and they have been realized in nearly all cases by the BLK as the decision taking body. In summary of the 82 Institute for 70 further funding was proposed, for 8 Institutes the termination and for 4 Institutes major changes of the organization were suggested. 5 of the negative evaluated Institutes were finally closed, for 3 a major rearrangement of the research program and the closure of unproductive research groups were agreed upon. The present running evaluation of the Leibniz Institutes showed already for two of these reorganized Institutes a strong positive development. In addition 8 Institutes had been proposed by the state governments for integration into the Leibniz organization, three of them failed to show the standards demanded. One further positive aspect of the comprehensive review should be mentioned. Many institutes have considered carefully the critical points mentioned in their evaluation report which in some cases required major changes of the institutes organization and research program. In many cases a strong improvement of the research output was observed in the evaluation by the Leibniz organization presently under way. The director's support by the scientific advisory committee to realize necessary changes turned out to be of high importance.

3. Evaluation by the Leibniz organization 2001 - 2009

As a conclusion of its work the BLC performed a system evaluation of the Blue List Institutes based on its experience of the 5 years work [4]. Amongst other things it was proposed that the Leibniz organization takes over the responsibility of the evaluation. A clear separation between the administrative office supporting this process from the Leibniz organization was

demanded. This was realized by assigning the responsibility of the GSC to the senate of the Leibniz organization. The senators are high ranking members of science organizations, administrators and leading scientists. The BLC is replaced by the SAE which have nearly the same composition. The only difference being that the status of two WGL vice-presidents and of the general secretary of the BLK changed from the status of a guest to a member without vote. It should be stressed that the WGL members of the SAE always strongly supported the critical and objective evaluation of the Institutes. Furthermore the composition of the EG has changed slightly, the chairman of the scientific advisory committee of the Institute is now participating as a guest, moreover an observer of the Leibniz organization attends to watch over the procedures applied. Only in one case he rightly had to criticize the course of the visit.

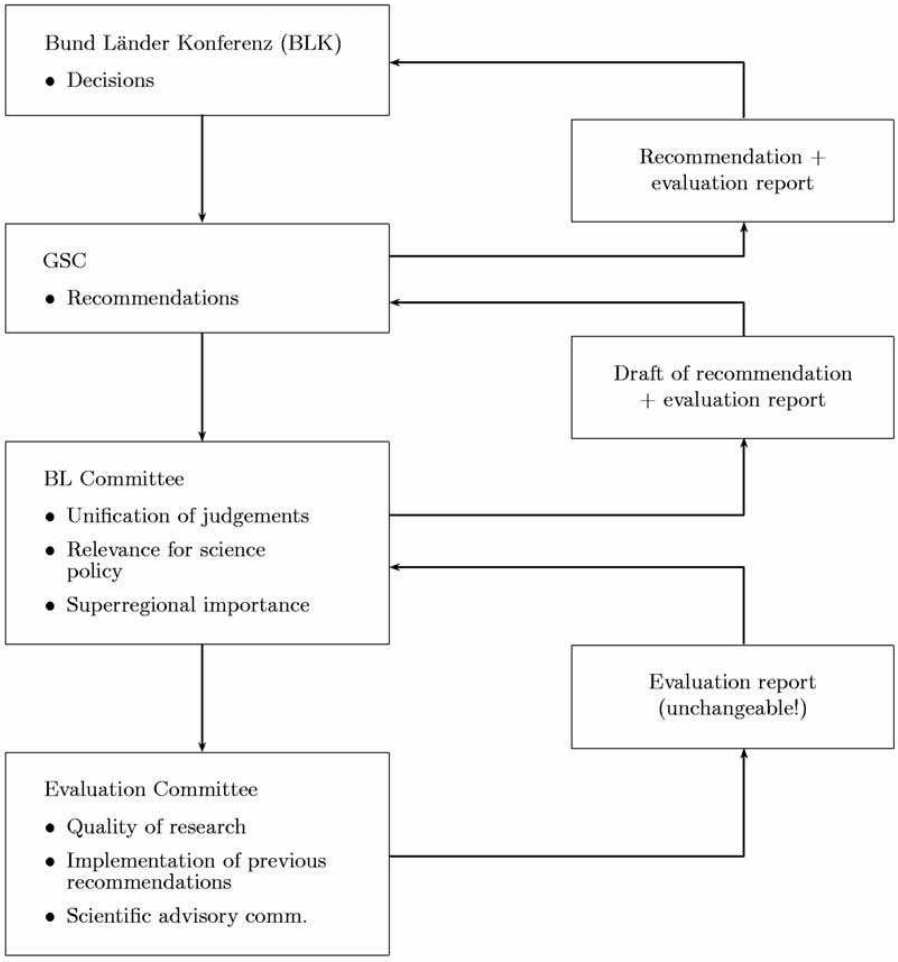
These changes of the composition of the SAE and the EG turn out to be advantageous. In addition at different stages of the procedure formal rules of appeal have been introduced; in this connection it should be mentioned that that for the 82 institutes evaluated by the GSC in no case the application of such rules turned out to be necessary. According to my experience the evaluation process under the supervision of the GSC and the Leibniz senate respectively are of the same quality; both are to a large extend objective, interferences from outside are rare and can be rejected. Nevertheless organizations which want to introduce a similar extensive evaluation process should follow the GSC procedure. I always appreciated that the recommendations were first discussed by the Academic Commission of the GSC, composed mainly of active scientists, since in this case the argument that research quality should have have highest priority was never disputed. The second weak point can be traced back to the fact that the members of the SAE office, which draft the evaluation report on the basis of the minutes of the visit and the written comments of the peers, are employees of the Leibniz organization and naturally have a bias in favor of that institution. In very few cases I had to expend a lot of time to ensure critical observations of the peers to be considered properly in the evaluation reports and in one case I failed.

4. Summary

The quality of the Leibniz organization institutes is presently evaluated for third time. The evaluation procedure follows closely the one applied by the German Science Council to the WGL institutes in the years 1995 - 2000. It is based on an *ex post* peer review. The criteria of continuing support of the institutes by the Federal and State government are high quality of the research output as a necessary condition. If this is fulfilled in addition the Institute has to be of super regional impact and of interest for national science policy. The evaluation led to changes of the science system in Germany: of the 82 institutes evaluated by the GSC 5 did finally not fulfill the conditions and had to leave the Leibniz organization, new positively evaluated Institutes replaced them. Many Institutes considered carefully the critical remarks of the peers and improved in an impressive way their research output. The importance of the strong engagement of many scientific advisory committees in this process of improvement should be stressed. The experience shows that evaluation reports should not hide weaknesses of the Institute but instead spell them out clearly. This is helpful for the director of an Institute and is a necessary prerequisite for the administration to overcome legal problems in case an Institute has to be closed.

When the evaluation process started 1995 it was not clear if at the end the work would be effective and politics would consider the advice given. A few circumstances were extremely helpful. The evaluation of the academy of sciences of the former DDR and the strong selection requirements applied defined the benchmark to be applied to a successful research institute. Politicians were willing to accept it. In addition it was clear from the beginning that scientists only would take over the heavy load of the evaluation process, if their advice was considered seriously i.e. that negative judgments should have appropriate consequences. After 2 years of complicated negotiations solutions were found which satisfied both sides. The support of the attitude of the Academic Commission of the GSC by the Federal government was of great help in this context. In addition the presidents of the Leibniz organization backed up the aim of the GSC for high quality of the institutes without restriction. Being engaged in the evaluation process of the Leibniz organization for more than 12 years I can say that it was not a lost time; my colleagues and I were able to contribute to the improvement of the effectiveness of an important part of the German science system.

Figure 1: Flow diagram: Research Assessment by GSC



DIETRICH WEGENER

Dietrich Wegener is Professor emeritus of experimental physics at Dortmund University. He has published more than 400 refereed papers in particle, nuclear and medical physics. In 1994—2000 he was member of the GSC. He chaired the Blue List Committee of the GSC in 1995—2000 and the senate committee for evaluation of the LO in 2001—2006. Moreover he participated in the system analysis of the HGF and the LO performed by the GSC

Contacts:

Institute of Physics - University of Dortmund *D 44221 Dortmund*
Tel. 0049 231 755 3555 *Email: wegener@physik.uni-dortmund.de*

References

[1] Max Planck Forum 6 (2002), Science between Evaluation and Innovation: A Conference on Peer Review
[2] D. Wegener, Museumskunde 67 2/02 (2002), p.112, Forschung an Museen - Beobachtungen bei der Evaluation der Blauen Liste (WGL)
[3] Wissenschaftsrat: Stellungnahmen zu den Instituten der Blauen Liste Band I-XI, Köln 1996-2001
[4] Wissenschaftsrat: Systemevaluation der Blauen Liste - Stellungnahme des Wissssenschaftsrates zum Abschluss der Bewertung der Blauen Liste Köln 2001