

## EVALUATION OF SCIENCE AS CONSULTANCY?

**Silke Gülker, Dagmar Simon, Marc Torka**

Éditions de la Maison des sciences de l'homme | « [Quaderni](#) »

2012/1 n° 77 | pages 41 à 54

ISSN 0987-1381

Article disponible en ligne à l'adresse :

---

<http://www.cairn.info/revue-quaderni-2012-1-page-41.htm>

---

Pour citer cet article :

---

Silke Gülker *et al.*, « Evaluation of science as consultancy? », *Quaderni* 2012/1 (n° 77), p. 41-54.

---

Distribution électronique Cairn.info pour Éditions de la Maison des sciences de l'homme.

© Éditions de la Maison des sciences de l'homme. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

---

Silke Gülker, Dagmar Simon et Marc Torka

## Evaluation of science as consultancy?

---

### Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

**revues.org**

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

### Référence électronique

Silke Gülker, Dagmar Simon et Marc Torka, « Evaluation of science as consultancy? », *Quaderni* [En ligne], 77 | Hiver 2011-2012, mis en ligne le 05 janvier 2014, consulté le 27 décembre 2012. URL : <http://quaderni.revues.org/547>

Éditeur : Les éditions de la Maison des sciences de l'Homme

<http://quaderni.revues.org>

<http://www.revues.org>

Document accessible en ligne sur : <http://quaderni.revues.org/547>

Ce document est le fac-similé de l'édition papier.

Cet article a été téléchargé sur le portail Cairn (<http://www.cairn.info>).



Distribution électronique Cairn pour Les éditions de la Maison des sciences de l'Homme et pour Revues.org (Centre pour l'édition électronique ouverte)

Tous droits réservés

## evaluation of science as consultancy ?

Silke  
Gülker

*Research fellow*

Dagmar  
Simon

*Head of Research Group*

Marc  
Torka

*Research fellow*

*Social Science Research Center,  
Research Group "Science  
Policy Studies", Berlin*

### **Restrictiveness of discourse about evaluations of science**

Within the media, and in political and sociological discourse, evaluations of science are considered to be both the expression and engine of a comprehensive deprofessionalization of the academic profession. The introduction of regularly conducted performance evaluations by policy makers is just the latest in a series of crisis diagnoses that has a long history (Altbach 1980 ; Clark 1989 ; Enders 1999 ; Musselin 2007). The reasons for this are many and varied. They range from a generalized loss of trust in the capacity of the sciences to organize themselves (Weingart 2005), to the need to allocate scarce research funds on the basis of evidence ("*evidence-based policy*"), to the triumph of new public management, a "*Weltanschauung*" whose origins lie outside the realm of science (Power 1997). "*Its message : replace the old regime, dominated by a state-regulated profession, with a new regime, dominated by a market- and state-driven organization*" (Schimank 2005). The deprofessionalization thesis encompasses three of the academic profession's existential dimensions :

- "*a reduction in academic self-governance*" (Schimank 2005 : 365), because exogenous stakeholders gain in influence through evaluation processes that involve the application of non-scientific criteria ;
- "*a decline in collegiality*" (Martin and Whitley 2010 : 73), because evaluations strengthen "*epistemic elites*" that assert their own interests at the expense of their colleagues' ;
- "*a retreat from case-specific evaluation*" (Oevermann 2005 : 47), because formalized,



quantified, and standardized indicators increasingly become relevant for making decisions in evaluations.

But, this multidimensional deprofessionalization thesis gives too little consideration to the concrete ways in which the evaluated and evaluating scientists act, interpret and formulate judgments within the evaluation process. These parties are still the principal actors in evaluations of science, and thus have considerable freedom in interpreting and shaping these processes. What in fact emerges, when the microlevels of evaluation processes are considered, is that scientists do not react passively to evaluation processes (Leisyte et al. 2010), instead bringing their own values, standards, and expectations to bear on the process and thereby structuring it. Through this, they attempt to relate this externally initiated process to their own value system.

Our microanalytical study, conducted in the context of an international comparative research project<sup>1</sup>, shows that scientists do not view evaluations solely as external performance monitoring, but also interpret it as a form of collegial consultancy. Therefore, evaluations always involve both the quality control (assessment) of *and* improvement of quality (consultancy) at scientific institutes. This perspective, which is very important for scientists, contradicts the deprofessionalization thesis and underlines the second function of evaluation, which is present in every process to a greater or lesser extent. This aims to improve the quality of research institutes in addition to monitoring performance.

Quality improvements and performance moni-

ring do, however, necessitate different modes of action and are thus in conflict with one another. Whereas a productive learning process aiming to improve quality presupposes the “*collegial consultancy*” mode, efficient performance monitoring relies on “*sanctioning evaluation*”.

Taking the German Leibniz Association’s (chapter 3) evaluation process as an example, we investigate the question of how the tension between a consultancy focus and a potentially consequential performance assessment is negotiated within an evaluation process. Or, to put it another way : (How) is consultancy possible within an evaluation context ?

In selecting the Leibniz Association’s evaluation process, we have selected a type of evaluation that is used in many countries. At the heart of the process is an interactive site visit to the scientific institutes being evaluated by a team of evaluators. The evaluators ultimately issue recommendations to science policy makers and the respective institutes (see also Röbbcke and Simon 2001). It is simultaneously also a form of audit with potentially far-reaching consequences that include the closure of the institute. Tensions are therefore written into the process.

The structural tension between evaluation and consultancy will firstly be theoretically explored, and then reconstructed in the Leibniz Association’s evaluation process guidelines, the evaluated institutes’ preparations, the evaluators’ ways of interpreting, acting, and reaching judgments, and finally in the ways in which the institutes’ react to the evaluation results. Finally, we will present two arguments for the opening



of research and discourse about evaluations of science : the role of scientists and their value systems within evaluation processes once again be strongly considered, which is why an analysis of the microstructures of evaluation is necessary.

### **The structural tension between consultancy and assessment<sup>2</sup>**

Why does a structural tension emerge, when evaluations wish to both assess and advise ? Evaluation is indeed a cognitive process, which is carried out a countless number of times daily. Evaluations in the form of assessments of science are, however, a special case that one might compare to a school exam. These are institutionally mandated processes that do not take place every day and have consequences for the future. Therefore, the situation must be made explicit in the form of criteria that are as clear as possible. This situation opposes two asymmetrical roles to each other. On one hand, the party being evaluated has not requested the evaluation, and cannot avoid the evaluation or ignore its results. On the other, there is an evaluator who has been accorded the authority to make decisions. This authority is based on presumably superior knowledge, clear evaluation criteria, and a distanced, or even neutral, external perspective. The evaluator is trusted to differentiate between good and bad performance (Buchholz 2008). This authority is also institutionally safeguarded, which means that the evaluator is invested with power. The evaluations have institutionally anticipated consequences and are not open to discussion ; just as school grades determine whether a student may proceed to the next class or what career opportunities they might have, institutional evaluations determine the fu-

ture of a research institute. The basic conditions that must be fulfilled for this communicative situation are that the subject of the evaluation is complete and clearly describable, and that there are defined, usually scalable evaluation criteria.

An ideal typical consultancy situation differs in many respects from an assessment. Consultancy does also presuppose a decision-making problem, and is not possible without knowledge and assessment of the problem. But, a consultancy situation is triggered by a decision-making problem on the part of the advice-seeking party, when he/she asks: “*What do you think ?*”. What follows is a natural rather than institutionally forced differentiation between the roles of the perplexed seeker of advice and the advisor, who is (presumably) equipped with helpful additional knowledge. The person may have his/her own criteria system, however, in order to be helpful this must be brought to bear on the advice seeker’s problems. Consultancy situations are apparent through a distinction between advice and its acceptance, between words and deeds (Fuchs 2004). The decision whether to accept advice within a consultancy situation remains with the advice seeker, whereas in an assessment situation it is made by the evaluator.

Voluntary participation, trust, and openness can be considered the central preconditions for a consultancy situation. The advisor/consultant should feel free to provide the advice that a particular situation calls for – he/she cannot be forced to provide any one specific piece of advice. The inverse is also true – a party cannot be forced to seek advice. The aim of a consultancy situation is that it produces something that the seeker of



advice can act on. For this to happen, what is discussed must be integrated into the advice seekers own value system, a degree of conviction that cannot be externally imposed. This requires complete trust that the advisor is acting in the best interests of the advice seeker – “*disinterestedness and good will*” is how Schützeichel and Brüsemeister (Schützeichel and Brüsemeister 2004 : 277) refer to it.

In a pure assessment, the point, from the perspective of the party being assessed, is to convince the reviewers by any means necessary, as a lot rides on the conclusion that is drawn. In the end, the grades signal performance and will not be questioned afterwards. It is therefore rational to employ strategies that could improve these grades, including deceit and fraud. This is not the case in a consultancy situation. Both parties must approach the situation in the spirit of mutual openness. The situation requires the conviction that openness, and not deceit or fraud, helps the advice seeker.

When evaluation and consultancy are interwoven in one process this is naturally associated with tension. There are nevertheless many situations where this combination is institutionalized<sup>3</sup> –, and the evaluation of science is one of them. Institutional research evaluations are politically initiated and the processes are to a large extent politically determined. The points of departure and goals are to inform decisions on funding allocation. The evaluator’s decision affects the way the institute is perceived, and in extreme cases it may have implications for the institute’s continued existence. This makes the observation that colleagues will encounter each other “within” evaluation

processes and use communicative forms of collegial consultancy all the more interesting. In the analysis that follows, we are therefore concerned with how this tension between evaluation and consultancy becomes evident, how it is dealt with, and what consequences emerge from it.

### **The evaluation procedure of the German Leibniz Association**

In the German scientific establishment, there are four scientific research organizations that, alongside the universities, play an important role in conducting research. These are the Max Planck Society, the Fraunhofer Gesellschaft, Helmholtz Association and the Leibniz Association. In the *Leibniz Association* there are 87 non-university research institutes with a variety of disciplinary focuses. Evaluations for these institutes have existed for a long period of time due to the particular way they are funded. These institutes are jointly funded by both the federal government and state governments and the purpose and necessity of this funding is regularly examined. This type of evaluation was developed by the German Science Council and was adopted by the Leibniz Association and other evaluation agencies. The fact that the institutes are evaluated *regularly*, every seven years, distinguishes the Leibniz Association’s procedures from others.

A visit to the institute in question, the so-called site visits, are at the heart of the Leibniz Association’s evaluation process. In preparation for this, the institute prepares a report in advance about itself. The visits are a day and a half in length and consist of a presentation by the institute directors to the entire team of evaluators, departmental



visits, discussions with employees without the presence of the management, a meeting with the administrative heads, and a discussion with selected partners.

This type of process, which focuses particularly on direct interaction between evaluators and evaluatees, is common at an international level. The *Standard Evaluation Protocol* in the Netherlands and the university evaluations by the French Evaluation Agency for Research and Higher Education (AERES) function in a similar way. Other systems do not allow any interaction between the evaluating and the evaluated researcher. For instance, the Research Assessment Exercise (RAE) in the UK and the Quality Review of the Australian National University (ANU) analyzes documentation in a system that is principally based on grading publications.<sup>4</sup>

So, in comparison with the other evaluation methods in use, this method should particularly emphasize the tensions between consultancy and evaluation. Direct interaction with the aim of formulating recommendations should encourage the development of a consultancy situation. The evaluators are peers that have been assembled in an evaluation team based on the institute's particular profile and reach judgments based on the specifics of each case : *"Peer reviewers are responsible for selection, evaluation and, if necessary, amendment of the criteria, depending on the related scientific community and mission of the institution"* (Leibniz-Association 2007). Interaction has great importance within this process.

At the same time, this method is also a procedure intended to gather information for making fun-

ding allocation decisions based on a *"standard list of criteria"* (Leibniz-Association 2007). A report is composed about the site visit. The work of the evaluation committee concludes with the submission of a report. The institute then has an opportunity to respond in a written statement. Both documents are submitted to the Senate of the Leibniz Association, which then makes a recommendation to the Joint Science Conference (Gemeinsame Wissenschaftskonferenz), which has overall responsibility in these matters. This method, therefore, investigates whether an institute meets standardized criteria and is deserving funding.

The method makes it possible to interpret evaluations as both collegial consultancy and performance assessment. The question is, how do the participants, both the evaluated and the evaluators, interpret the situation – how do they deal with the tension between consultancy and assessment ?

### **How an institute prepares for evaluations**

In preparing for an evaluation, an institute has two important tasks to complete : the documents must be prepared in time, and preparations for the site visit must be made as far as content and organization is concerned. There are countless decisions, both large and small, associated with these tasks, from content questions such as which research units should be the central focus, or what should be portrayed as the institute's specialism, down to organizational questions about what poster should be displayed in what room and what food should be prepared. The process of answering these questions is shaped by the ten-



sion between consultancy and assessment associated with this process.

What emerges above all else is that all of the participants in an institute are aware that it is an evaluation, or even an examination, and that a lot is at stake for the institute in question. As a result of this, strategic behavior is also evident, such as pleasing the evaluators and ensuring as host that you are showing your best side. This includes cleaning up, because one doesn't want a "railway station atmosphere, having pleasant surroundings does create a different atmosphere" (Bock<sup>5</sup> 469<sup>6</sup>) or effective event management : "because you only have to think of what you have to do when you are preparing for a wedding [laughs], to make sure everything works out." (Dittmer 401).

So, strategic behavior does play a role in the preparations for a site visit. But it is also clear that the institution's representatives expect more from the evaluation than just an examination on a given day. The production of preparatory documents in fact becomes part of an all-encompassing process of self-understanding. In all analyzed institutes the preparation process requires at least a year. It involves both compiling a summary of output, and summarizing and examining the institute's structure. As a consequence the evaluation procedure is seen much more as a kind of forced organizational consultancy. Thus, one deputy institute director rated the evaluation process as positive nonetheless, precisely because it does not just focus on past performance but also looks at future prospects. In order to develop these prospects, a two-day workshop with selected staff members was organized at the behest of the scientific advisory board. This workshop "was

*definitely a preparation for the future and for the evaluation"* (Müller 91). These internal processes of self-understanding go and above and beyond strategic preparations, because the identity of the institute is itself up for debate : "*That was exciting, I think the institute learned a lot about itself also*" (ibid. : 130). In this case, a new direction for the institute even emerged from these intense discussions (connecting policy-oriented and pure research), which was established in the relevant area. The evaluators could then consider it as an object for consultancy. "*In this case it really was all about the direction that the institute should take, and there was one camp that was saying, [...] we've done enough pure research, [...] but now our mission is advising policy, come hell or high water. [...] And the other camp was saying, if we go in this direction [...] we won't be playing to the strengths that made the [evaluated institute] great. [...] And it was just a long discussion that went on for months and at some point the boss made his decision and said ok, that's it now.*" (ibid. : 187)

This example demonstrates quite clearly that preparations for evaluations are not just about giving the institute a new coat of paint, in the sense of well-formulated texts with a professional layout. If one assumes that in an evaluation the only thing that matters are the results, then this tactic would be completely rational. But the evaluation was "let into" the institute because the situation was also interpreted as a form of consultancy. The team of evaluators is also a group of colleagues, from whom one expects to receive constructive feedback. Hence, underlying assumptions are rigorously questioned in the preparatory phase, much more carefully than if



it was exclusively about drawing up documents. And decisions were made. The one-day event that is the evaluation forces the making of decisions about questions that have probably been a latent undercurrent to the institute's work for some time. The consequences of these decisions, which will at first only appear on paper, depend on how the evaluation proceeds. The question of how the evaluators interpret the situation – as consultancy or evaluation – is crucial here.

### Reviewers in action

During site visits, evaluators are charged with assessing the institute's performance and making recommendations as to whether the institute is still worthy of funding and how its work can be improved. So, they advise both science-policy authorities and the evaluated institute. This constellation gives rise to a tense set of relations, in which loyalties become blurred. How do evaluators position themselves in this kind of situation, which has been characterized by Uwe Schimank (2004) as being a drastic transition from collegial "tact" to "treachery" because "*the advisor advises the advised how to deal with third parties that are also the advisor's professional colleagues*"?

Based on accounts by evaluators on why they participate in evaluations in the first place, we can assume that they feel a primary duty to the scientific community and to a lesser extent to science-policy authorities. They see the evaluation as part of scientific self-monitoring that one should do in any case. It is, in fact, "[...] *a service that we should provide because it is part of the job of science [...] to do something positive serving science.*" (Nunzinger 5). An expectation

on the part of evaluated parties to receive support with as little criticism as possible is unsuitable for this, as is delivering devastating criticism without any suggested solutions. Constructive criticism is instead the most effective means of completing their job as consultants: "*That is the important thing about evaluations, what's the use of an evaluation if you don't get any criticism at the end of it all, it would really be pointless then if you do not get any help for your ongoing work – support might be a better way of putting it*" (Dallmeier 1056).

The search for critical points serves to diagnose problems and formulate constructive recommendations, and structures the actions of the evaluator in the evaluation process. Identifying critical points within the context of evaluations is however a particularly challenging process. After all, they receive documents, participate in a tightly scheduled program of events and interact with staff members who have rehearsed everything in advance and eliminated any obstacles to get support from the evaluators. The evaluators encounter a structurally inconvenient situation, given their role as consultants, as performance and not problems are presented. Therefore, evaluators themselves are required to uncover problems in what is a time-consuming multistage process:

1. *Uncovering problems individually*: Evaluators get their first impression about the institute as a whole from the documents submitted. They refer back to "*very specific standards in their head*" (Troemmel 353) and a general knowledge of the field and its typical challenges and problems (Barlösius 2008). Using understandings of nor-

mality – such as the particularities of the scientific topic, typical career problems, the institutional embedding of the institute and expected publication performance or qualification times – the heterogeneous information sources are “*read very selectively*” (Troemmel 54). Deviations, missing explications, or textual inconsistencies provide an opportunity for questions, criticisms, and the search for possible improvements. Problems, therefore, must first be hermeneutically reconstructed, because they are not explicated in the documents.

2. *Collective problem stabilization* : Before the active site visit begins, these initial impressions are collected by the evaluation committee and are tested within the team for their robustness. A very general communicative mechanism is central to this : “*if nobody takes it up [the evaluators comment] then it is dead*” (Troemmel 226). Individual comments may be supported by others or contested and be strengthened or weakened and thus successful stand the test or not. Using these kinds of discursive processes, a consensual impression is developed of whether the evaluation will be more or less problematic. Given that the closure of an institute is only a last resort, one element of the evaluators’ dual role tends to be stronger : “*This ended up going much more strongly in the direction of consultant than auditor over the course of the [evaluation]*” (Deichmann 430). The tension between both poles of the evaluator role does not, however, simply vanish : “*So it’s a combination, because advising is too optional, as if you can do with my advice what you want*” (Notter 117).

3. *Interactive problem testing* : Direct interac-

tion with members of the institute during the site visit ultimately serves to further inspect whether the impression that the team of evaluators collectively got from the written document can be tallied with practice in the institute. The standard is : “*The papers [must] confirm what the people say.*” Precise observation of the ways in which institute acts and answers give the evaluators a less window-dressed reference point for making their judgments because “*some nonverbal indicators*” (Troemmel 449) come into play. In doing this, evaluators go above and beyond a simple performance assessment, because they see communication about problems and a realistic self-assessment as beneficial, even though strategic behavior and highlighting performance might be expected : “*When they [...] admit a problem that is basically also an indicator that they are going in the right direction and that they have a realistic self-image*” (Fissler 864). Being open about problems and having a realistic self-image are the basic conditions that must be fulfilled if evaluators are to be able to advise. These preconditions may not, however, be fulfilled, if evaluators first have to check whether these self-descriptions correspond with reality, and dig up problems. A consultancy focus on the part of evaluators is therefore in conflict with exam-like performance assessment in evaluations.

4. *Public communication of problems* : Only at the end of the site visit is it decided which observations will be adopted by the evaluation committee and be made public as intersubjectively shared collective judgments, in the form of “*recommendations*”. In principle, a more intense version of the above-described discursive mechanism is repeated, in that suggestions from other evalua-



tors must be accepted and thus confirmed. This is because, for one, the report is now intended for science-policy authorities that could draw unwanted conclusions from tough criticism. This is also because binding action orders will be drawn up on the basis of these “*recommendations*” as soon as they are set down in writing, and their implementation will be checked in the next evaluation. Therefore, evaluators place great value formulating their results. The tone is praising and rough edges are smoothed so that “*certain things are only included in a very indirect form*” (Kunst 227). The tension between collegial consultancy and assessment is thus continued when recommendations are formulated. Clear language would be more beneficial for giving advice, were it not for the fact that direct sanctions are associated with it.

Based on the evaluators’ ways of behaving and interpreting, it can be seen that this type of evaluation is far removed from a impersonal, interpretation-free explicitly procedure-driven “*mechanical objectivity*” (Daston and Galison 2007 ; Porter 1995). Uncovering problems and providing solutions and advice to the evaluated institutes are social processes that rely on interactive rules (Lamont 2009 ; Lamont et al. 2009), discursive mechanisms and interpretive competencies. But how, then, do the institutes interpret the evaluation process and results ?

### **How an institute responds**

The evaluators therefore understand themselves overwhelmingly as consultants to the evaluated institutes. The institutes also emphasize this role during the “*site visits*” : “[It] was because the

*evaluators were consistently very constructive, of course there were some that had their own agenda or their own particular interests, but despite this I have to say, it was extremely constructive”* (Dagendorf 196).

The way they deal with the evaluation report shows that the institutes interpret the results both as assessment and as consultancy, and, in keeping with this, integrate this into their action strategies. As one would expect in the assessment mode, the results and recommendations must be interpreted as binding action orders to the institute, which, if not followed, may result in serious consequences up to and including the closure of the institute. As consultancy instruments, the evaluation results are interpreted as an information basis for appropriate action, the strategy for which lies largely in the hands of institute actors.

As the preparations have already shown, their way of dealing with the results of evaluations is not purely strategic in nature. That would imply that the evaluations were (at best) only being used as a proof of quality that could be used in negotiations with ministries to get financial assistance. That would also mean accepting and implementing recommendations without exception. The reactions of the institutes indicate a different approach : evaluation results are used as an impetus for change, confirmation of development processes that have already been internally initiated and/or as validation from an external authority of their own plans. “*Yes, that was a confirmation of the plans*” (Ulbricht 12). “*If we hadn’t gotten that from outside, [...] there wasn’t an opportunity to push it through*” (Xaver-Unger 1625).



The institutes use the evaluations as a form of organizational consultancy, even if they did not commission it themselves. The recommendations are reflected upon earnestly and implemented, but not necessarily on a one-to-one basis, rather links to internal structural or content development processes that are already ongoing are sought : *“They are generally not must-dos, they are possibilities that we can examine and see to what extent they can be implemented”* (Dagendorf 43).

In summary we can determine that not only the evaluators, but also the institute representatives accord evaluations a meaning that goes beyond a purely externally forced monitoring situation. The case specificity, that is the focus on tasks, goals, problems, and future prospects for the research institute in question is what gives these evaluations added value : *“[...] when it was not about just number of publications or third-party funding attracted, but about how issues such as knowledge transfer, supporting new researchers, policy advice etc. work, my impression is that the evaluators also succeeded very well in not just judging [...] but also in justifying these very well”* (Dagendorf 28).

The institutes not only really look for the evaluators’ advice and take it up in internal discussions about the future of the institute ; they also use their observations, advice and recommendations as a validating authority for their own plans for institutional change.

### **Conclusion and outlook**

Based on the fact that evaluations of science are externally initiated and organized systems of

justification, the dominant discourse often draws far-reaching conclusions about the dominance of external quality criteria and standards (defined by science policy), and a deprofessionalization of the scientific profession. When the microlevel is considered it is apparent that evaluations are not at all exclusively considered as *“external control”* but rather also as collegial feedback. The evaluators’ self-understandings are not those of auditors ; they see their task as a service to the profession. The institutes’ preparations go far beyond strategic calculation and the connections at the level of content are sought in the results.

These findings should also provide impetus for a new orientation in research on evaluations of science. The ways in which scientists engage with evaluations of science should be more strongly considered than thus far. At present, there is a tendency in research on evaluations of science to hastily associate new institutional frameworks with comprehensive changes in the functional logic of science. In order to better understand the mechanisms by which science policy instruments affect the organization and production of science, an analysis of microstructures should be considered to close gaps in the sociology of science. Only in this way will it become apparent that, for example, the principle of collegiality cannot simply be replaced by either the mechanisms of competition, the influence of an evaluator elite or by formalized evaluation procedures.

That does not by any means imply that evaluations have no influence on the type and methods of scientific evaluation, its structures and organization forms, and scientists’ professional self-understandings. Our analysis cannot make



any firm statements about the possible long-term consequences. Researching the effects of evaluations of science on the specific ways that both evaluating and evaluated scientists react would be a useful research perspective in this sense, and would thus make a new contribution to the discussions about the deprofessionalization of the academic profession and the impact and “triumph” of new public management in science.

## R · É · F · É · R · E · N · C · E · S

ALTBACH, Philip G. (1980) : “The Crisis of the Professoriate”. *The ANNALS of the American Academy of Political and Social Science* 448 : 1-14.

BARLÖSIUS, Eva (2008) : “Urteilsgewissheit und wissenschaftliches Kapital”. In : *Wissenschaft unter Beobachtung. Effekte und Defekte von Evaluationen*, edited by Hildegard Matthies and Dagmar Simon. Wiesbaden : VS Verlag für Sozialwissenschaften : 248-264.

BUCHHOLZ, Kai (2008) : *Professionalisierung der wissenschaftlichen Politikberatung ? Interaktions- und Professionssoziologische Perspektiven*. Bielefeld.

CLARK, Burton R. (1989) : “The Academic Life : Samll Worlds, Different Worlds”. *Education Researcher* : 4-8.

DANIEL, Hans-Dieter (2001) : *Wissenschaftsevaluation. Neuere Entwicklungen und heutiger Stand der Forschungs- und Hochschulevaluation in ausgewählten Ländern*. Bern, im Internet : [http ://www.swtr.ch/Publikationen/2001/CEST\\_2001\\_2.pdf](http://www.swtr.ch/Publikationen/2001/CEST_2001_2.pdf), letzter Zugriff 8. August 2011 : CEST Center for Science and Technology Studies.

DASTON, Lorraine, and GALISON, Peter (2007) : *Objectivity*. New York : Zone Books.

ENDERS, Jürgen (1999) : “Crisis ? What crisis ? The academic professions in the ‘knowledge’ society”. *Higher Education* 38 : 71-81.

FUCHS, Peter (2004) : “Die magische Welt der Beratung”. In : *Die beratene Gesellschaft. Zur gesellschaftlichen Bedeutung von Beratung*, edited by Rainer Schützeichel and Thomas Brüsenmeister. Wiesbaden : 239-257.

- GLÄSER, Jochen, and LAUDEL, Grit (2005) : *The Impact of Evaluations on the Content of Australian University Research*. In : *TASA Conference*. University of Tasmania.
- (2007) : “Evaluation without Evaluators : The Impact of Funding Formulae on Australian University Research”. In : *The Changing Governance of the Sciences : The Advent of Research Evaluation Systems*, edited by Richard Whitley and Gläser Jochen. Dordrecht : 127-151.
- HORNBOSTEL, Stefan (2010) : “(Forschungs-) Evaluation”. In : *Handbuch Wissenschaftspolitik*, edited by Dagmar Simon, Andreas Knie and Stefan Hornbostel. Wiesbaden : VS, Verl. für Sozialwiss. :
- KROMREY, Helmut (2003) : “Qualität und Evaluation im System Hochschule”. In : *Evaluationsforschung*, edited by Reinhard Stockmann. Opladen : Leske+Budrich : 233-258.
- LAMONT, Michèle (2009) : *How Professors Think : Inside the Curious World of Academic Judgment*. Cambridge : Harvard University Press.
- LAMONT, Michèle, Mallard, Grégoire, and Guetzkow, Joshua (2009) : “Fairness as Appropriateness : Negotiating Epistemological Differences in Peer Review”. *Science, Technology and Human Values* 34 (5) : 573-606.
- LEIBNIZ-ASSOCIATION. 2007. “Evaluation Criteria for Institutions of the Leibniz-Gemeinschaft.” Leibniz-Association.
- LEISYTE, Liudvika, Enders, Jürgen, and Boer, Harry de (2010) : “Mediating Problem Choice : Academic Researchers’ Responses to Changes in their Institutional Environment”. In : *Reconfiguring Knowledge Production. Changing Authority Relationships in the Science and their Consequences for Intellectual Innovation*, edited by Richard Whitley, Jochen Gläser and Lars Engwall. Oxford : Oxford University Press : 266-290.
- LOUVEL, Séverine, and Lange, Stefan (2010) : “L’évaluation de la recherche : l’exemple de trois pays européens”. *Sciences de la société* 79 : 11-26.
- MAGNIN, Chantal (2004) : “Consultation and Control. A Typical Dilemma for the Activating State”. *Schweizerische Zeitschrift für Soziologie* 30 : 339-361.
- MARTIN, Ben, and WHITLEY, Richard (2010) : “The UK Research Assessment Exercise : A Case of Regulatory Capture ?”. In : *Reconfiguring Knowledge Production. Changing Authority Relationships in the Science and their Consequences for Intellectual Innovation*, edited by Richard Whitley, Jochen Gläser and Lars Engwall. Oxford : Oxford University Press : 51-80.
- MUSSELIN, Christine (2007) : *The Transformation of Academic Work : Facts and Analysis. Research & Occasional Paper Series : Center of Studies in Higher Education*. California : University of California.
- OEVERMANN, Ulrich (2005) : “Wissenschaft als Beruf. Die Professionalisierung wissenschaftlichen Handelns und die gegenwärtige Universitätsentwicklung”. *die Hochschule* : 14-51.
- PORTER, Theodore M. (1995) : *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton, New Jersey : Princeton University Press.
- POWER, Michael (1997) : *The Audit Society : Rituals of Verification* Oxford : Oxford University Press.
- RÖBBECKE, Martina, and SIMON, Dagmar (2001) : *Reflexive Evaluation. Ziele, Verfahren und Instrumente der Bewertung von Forschungsinstituten*. Berlin : edition sigma.
- SCHIMANK, Uwe (2004) : “Leistungsbeur-



teilung von Kollegen und Politikberatung am Beispiel von Evaluationen im Hochschulsystem”. In : *Die beratende Gesellschaft : Zur gesellschaftlichen Bedeutung von Beratung*, edited by Rainer Schützel and Thomas Brüsemeister. Wiesbaden : VS-Verlag : 39-56 ; (2005) : “‘New Public Management’ and the Academic Profession : Reflections on the German Situation”. *Minerva* 43 : 361-376.

SCHÜTZEICHEL, Rainer, and BRÜSEMEIER, Thomas eds.) (2004) : *Die beratene Gesellschaft. Zur gesellschaftlichen Bedeutung von Beratung*. Wiesbaden : VS Verlag für Sozialwissenschaften.

WEINGART, Peter (2005) : “Das Ritual der Evaluierung und die Verführung der Zahlen”. In : *Die Wissenschaft der Öffentlichkeit. Essays zum Verhältnis von Wissenschaft, Medien, Öffentlichkeit*, edited by Peter Weingart. Weilerswist : Velbrück Wissenschaft : 102-122.

## N · O · T · E · S

1. The project employed the *Research Assessment Exercise* in Great Britain, the *Standard Evaluation Protocol* in the Netherlands and the *Leibniz Association evaluation procedure* in Germany as research subjects. Between 2007 and 2008, approximately 100 interviews with reviewers and representatives of the evaluated institutions were conducted in order to determine the internal processes of evaluations.
2. Consultancy and assessment are often assigned to different categories in research on evaluations: formative versus summative (Kromrey 2003). The problem of the mixture of these two elements is, however, rarely discussed (Hornbostel 2010).
3. One well known example is consultancy in the public employment service. If the job seeker is totally open and honest – as it would be necessary for a adequate help –, he or she would probably risk a cut of benefits (Magnin 2004).
4. For more on the different principles underlying evaluations of science see Hornbostel (Hornbostel 2010), for international comparisons Daniel (Daniel 2001), (Louvel and Lange 2010), on the RAE Martin and Whitley (Martin and Whitley 2010), on the Australian system Gläser and Laudel (Gläser and Laudel 2005; Gläser and Laudel 2007).
5. All names are changed in order to anonymize. The numbers refer to the beginning line of the interview’s transcription.



## R · É · S · U · M · É

Le discours actuel sur les évaluations institutionnelles de la recherche s'est construit autour d'une thèse dominante : en tant que contrôle de la performance initié par l'extérieur, les évaluations sont considérées comme l'expression, et le moteur, d'une déprofessionnalisation globale de la profession académique. Les évaluateurs comme les chercheurs évalués supposent qu'ils devront ajuster leurs systèmes de valeurs et d'évaluation à des critères externes. Notre analyse empirique des procédures d'évaluation à un niveau microsociologique se distance de cette thèse, et montre que les valeurs centrales de la profession académique persistent, et continuent à structurer la prise de décision et les actions de celle-ci. Une de ces valeurs stipule que les évaluations ne sont pas seulement au service du contrôle de la performance et de la sanction, mais qu'elles apportent les conseils de collègues. Mais comment le conseil est-il possible dans un contexte de contrôle de la performance, dont les conséquences sont potentiellement graves ? Nous répondons à cette question en analysant les décisions et les actions d'évaluateurs, et de chercheurs évalués, dans le cadre de la procédure d'évaluation des instituts de recherche allemands de l'Association Leibniz.

### **Abstract**

The current discourse on institutional research evaluations has been shaped by a dominant thesis. As an externally initiated form of performance monitoring, evaluations are considered to be both the expression of and drivers of a comprehensive deprofessionalization of the academic profession. Both evaluating and evaluated researchers assume they will have to adjust their value and evaluation systems in keeping with external benchmarks. Our empirical analysis of

evaluation processes at a microlevel contrasts with this thesis, and shows that the academic profession's central values persist, and continue to structure the ways in which the profession makes decisions and acts. One of these values is that evaluations should not just serve to monitor performance and sanction, but instead provide collegial consultancy and feedback. But how is consultancy possible in a performance-monitoring context whose consequences are potentially serious? We address this question by analyzing the decisions and actions of evaluating and evaluated researchers in the context of the German Leibniz Association's evaluation procedure.