



5

The Many Faces of Peer Review

Hanne Foss Hansen

Introduction

Evaluation is an ongoing activity in most parts of contemporary societies. Terms such as “the evaluative state” (Neave, 1998) and “the evaluation society” (Dahler-Larsen, 2012) have been used to describe this.

In academia evaluation organized as peer review dates back to the eighteenth century (Benos et al., 2007). Since then peer review practices have been discussed and developed. Today peer review processes take many forms. Almost all aspects of scientific enterprise rely on evaluation done by peers (Bornmann, 2011). Peers evaluate doctoral dissertations, applicants to academic positions, applications for promotion, applications for research grants, manuscripts submitted for publication, scholars proposed to receive awards and prizes as well as research organizations in the form of groups, departments, programs, institutes and universities (Langfeldt & Kyvik, 2011, 2015). In some countries peers even evaluate

H. F. Hansen (✉)

University of Copenhagen, Copenhagen, Denmark

e-mail: hfh@ifs.ku.dk

disciplines and interdisciplinary fields at the national level across universities and other research organizations. In some research areas they also critically evaluate newly published books in order to assess whether they contribute to new knowledge to the field. Furthermore, peers sometimes assess research in order to decide on which knowledge to be used as input to policy-making and regulation.

The scientific enterprise is in this way permeated with evaluation activities (Hamann & Beljean, 2017). The same goes for higher education. Here peer panels and other types of panels, including peers, experts on educational leadership, labor market representatives and students, evaluate the quality of educational programs and educational quality assurance systems (Hansen, 2009a, 2014). All in all, peers thus undertake many evaluator roles.

Alongside the development and expansion of peer review activities, also other governance structures in academia have been transformed. In the last decades, university systems in the Nordic countries have experienced continuous change. Funding arrangements, accountability measures as well as institutional management and organizational structures have been transformed (Geschwind et al., 2019; Hansen et al., 2019a). Universities have increasingly developed into what observers have termed “corporate actors” with organizational traits such as identity, hierarchy and rationality (de Boer et al., 2007), and “complete organizations” where an authoritative center coordinates and controls actions through the hierarchy (Brunsson & Sahlin-Andersson, 2000; Seeber et al., 2015).

Changes in the institutional context in which peer review activities are embedded seem to have influenced the role and use of peer review. Whereas peer review historically was an opportunity for peers to exercise academic power in a professional self-regulated system, it today often plays the role as giving managerial advice and ensuring accountability. Using the concepts of Johan P. Olsen (2007), the importance of peer review practices exercised in the context of the universities as self-governing communities of scholars has diminished while it has increased in the contexts of the universities as instruments for political agendas as well as service enterprises embedded in competitive markets.

This chapter sheds light on this development. Focus is on two research questions: (1) How has peer review as an evaluation concept been

developed over time? (2) What is the role of peer review today? Focus is both on research evaluation on the scientific side of universities and on the educational side. The analysis is mostly conceptual, aiming at mapping different types of peer review. The mapping of peer review is supplemented with examples from Denmark and the other Scandinavian countries.

The analysis is based on documentary material and self-experience from working in the university sector for more than 35 years. During that time I have been doing research in the development of the university sector and in evaluation practices (Borum & Hansen, 2000; Hansen, 2009a, 2009b; Hansen & Borum, 1999; Hansen et al., 2019a, 2019b). Further, I have been acting in the role as peer in Denmark and abroad in many contexts, among others, in the publication system, in assessment of dissertations, in assessing applicants for positions and promotion as well as in assessing project applications in research councils. Moreover, I have served as peer in a national discipline evaluation (Political Science in Norway, initiated by the Norwegian Research Council in 2002) and participated in panels in educational quality assurance (University of Tromsø, 2006; NLA University College, 2005; Norwegian Academy of Music, 2004; all under the auspices of the Norwegian Agency for Quality Assurance in Education [NOKUT]). I have also worked with evaluation authorities in both the research and educational fields in Denmark, Norway and Sweden as advisor and board member. I draw on these experiences in the analysis, knowing of course that other scholars may have other experiences from other scientific fields and other organizations.

The chapter is structured in seven sections. The next section presents classical peer review as an evaluation model. The following three sections discuss how peer review related to the evaluation of research has developed over time into other types of evaluation models, termed informed and standards-based peer review, modified peer review and extended peer review. Then, the next section is dedicated to the role of peer review in the higher educational field. The final section holds the conclusion and some personal reflections on the challenges which peer review practices meet these years.

Classical Peer Review as One of Several Evaluation Models

The concept of evaluation can briefly be defined as assessment or appraisal. In the specialized evaluation literature, the concept is defined as systematic assessment of the merit, worth and value (Scriven, 1991) of evaluands (= objects for evaluation). As mentioned above, there are multiple evaluands in relation to research and higher education. The literature on evaluation offers a range of evaluation models. Peer review can be characterized as a professional or collegial evaluation model (Vedung, 1997; Hansen, 2005) where the evaluation criteria are defined by the peers and not by other stakeholders, for example, users. The fundamental idea is that members of the profession are trusted to evaluate other members' activity and results. Evaluation may be summative, assessing against a standard or benchmark, for example, is the article worth publishing or is the applicant qualified for a professorship, or it may be formative, focusing on whether activities are in progress.

Peer review is the classical type of evaluation in research. In classical peer review evaluation recognized researchers read and assess other researchers' contributions, focusing either on project ideas, as in, for example, research councils; on manuscripts submitted for publication, as in relation to journals and publishing houses; or on CVs and publications handed in in relation to applications for appointment and promotion. The content and context of classical peer review is characterized in Table 5.1 (Hansen, 2009b). Classical peer review evaluation is the corner stone in gatekeeping in academia aiming at ensuring quality control and the best possible distribution of scarce resources.

Table 5.1 Classical peer review

Task	Assessment of research quality of products and individuals.
Process	Reading first-order material, assessing, nominating and sometimes ranking.
Peer panel composition	Homogeneous. Mono-professional.
Evaluation approach	Summative, clear decision focus.

Classical peer review is based on reading what can be termed first-order material, for example, publications, manuscripts and project proposals. Classical peer review is most often organized with a number of peers. Peers are either working in parallel, as in review processes of manuscripts submitted for publication, or as a panel, as in assessment of applicants for a position. The assessment process is relational in the sense that it is made in a context. A dissertation, for example, is assessed in relation to the research area to which it seeks to contribute, just as an applicant to a position is assessed in relation to the job description. Most often, the process includes a form of cross-control. The assessment of a manuscript is passed on from the reviewers to the editor(s) and the assessment of a dissertation from the assessment panel to, for example, the dean.

Classical peer review is still of great importance first and foremost in the publication system and in relation to assessing research applications. With country variations, classical peer review in the Nordic countries is however being transformed (Hansen et al., 2019b). In a Danish context classical peer review sometime back constituted the corner stone in the recruitment system. All applicants for a position were assessed by a peer panel reading their enclosed publications, nominating whether the individual applicant was qualified for the position and finally ranking the qualified applicants. The assessment document was passed on to the university management, which normally acted according to the proposed ranking. Only in cases of disagreement in panels there was more leeway for management. Further, the assessment document was distributed to all applicants for the position, a process often initiating and supporting a discussion in the research environment on the attributes of research quality. As I will return to in below, peer review has been transformed into applicant assessment processes, which, these years, are organized very differently.

Informed and Standards-Based Peer Review

In the Danish university system peer review in relation to recruitment and promotion has changed considerably. Peer influence has been reduced, managerial power increased. Applicants are assessed as either

qualified or not qualified, no longer ranked. Publicness is reduced, confidentiality is the norm. In still more contexts peers in assessment committees only look at the qualifications of the applicants shortlisted. Criteria for shortlisting are blurred.

Further, the reading of first-order material which are at the core of classical peer review has increasingly become supplemented by the use of metrics, and a range of other assessment criteria have become increasingly important. We can term this informed and standards-based peer review. The content and context of this type of peer review is characterized in Table 5.2.

In recent years still more easy accessible metrics have been developed. Some are global systems. Examples are Scopus and Google Scholar counting citations and presenting indexes such as, for example, the h-index as well as journal metrics such as the journal impact factors (JIF). Others are national systems such as the Danish bibliometric research indicator (BFI) developed upon inspiration from the Norwegian model and dividing journal and publishers into two, in some disciplines three, quality categories.

Although metrics are contested, it is my experience that they move into use in peer review processes, for example, in assessments of applicants for positions and promotion. Often, it is applicants themselves referring to metrics. Sometimes it is peers arguing that metrics are relevant. Sometimes it is managers arguing that panels are expected to or even demanded to include metrics. And this happens, even though bibliometric experts, for example, in the Leiden Manifesto (Hicks et al., 2015) argue that metrics should not be used at the level of individuals. It

Table 5.2 Informed and standards-based peer review

Task	Assessment of research quality of products and individuals.
Process	Reading first-order material and assessing, nominating and sometimes ranking by including metrics such as publication ranking, citation metrics, etc. Further giving priority to multiple evaluation criteria.
Peer panel composition	Homogeneous. Mono-professional.
Evaluation approach	Summative, clear decision focus.

seems that accessible metrics increasingly are used also for purposes they were not prepared to be used for.

Standardization seems to be another development tendency at least in Denmark. In relation to recruitment processes, departments increasingly specify the dimensions applicants have to be assessed upon, sometimes termed a scholarly qualification matrix, the SQM. For a professorship a SQM may, for example, outline four dimensions, all with several sub-dimensions, to be assessed: Research (internationally recognized, proven ability to engage in new areas of research, frequent publications, good track record, acquiring external grants, experience with leading roles in networks), education (ample experience and good results with supervision, solid experience and good results with course development, solid experience of PhD supervision), service to society (proven ability to engage with stakeholders outside the university,) and personal (evidence of active contributions to the administrative and managerial tasks, evidence of active mentorship, including co-authorships with junior colleagues). In addition to such performance goals there may be further criteria such as the ability to establish cross-disciplinary networks at faculty level, good results with study program management and evidence of how own research has had non-academic impact. Compared to classical peer review related to recruitment, criteria these years have become very multidimensional.

This development of peer review becoming more standardized and metric based may reflect an aim to make assessments more transparent and fair, but it transforms peer review practices from being discretion-norm-based practices to standards-based practices. Further, it reflects an increasing individualization. In the thinking of the SQM, departments are not entities which on a group or collective basis have to fulfill multidimensional performance goals. Rather, every individual in the department has to fulfill every imaginable performance goal.

Another area where peer review becomes more standardized is in assessments of applications for research projects, centers and so on. In some contexts, research councils demand peers to use multiple specified criteria and sometimes also to put grades on all criteria. Again, this development may reflect an aim of transparency and fairness, but it enables decision-makers to make mechanical decisions.

Modified Peer Review

In addition to scientific floor-level evaluands such as research products and scholars, more evaluands such as research groups, departments and universities have been introduced. Evaluation is no longer only going on at the micro level. In the meta- and macro-level evaluation cases, the reading of first-order material (publications) either is less important or it has been replaced by the reading of other types of materials such as self-evaluation reports, a range of available metrics and, in some systems, impact cases. Also site visits and interviews have become important evaluative information. We can term this modified peer review. The content and context of modified peer review is characterized in Table 5.3.

Modified peer review is implemented both on the individual department, faculty or university level, and, in some countries, in national evaluation systems. In all three Scandinavian countries modified peer review has been on the agenda for some years.

The Research Council of Norway in the last 20 years has carried through several rounds of research quality evaluation based on modified peer review (<https://www.forskningradet.no/Statistikk-og-evaluering/evaluering/>). Most evaluations have had the focus on disciplines, one panel evaluating all departments in the discipline across the universities. In recent years, focus has been broadened. In 2011 an evaluation of biology, medicine and health research was carried out, in 2017 one of the humanities and in 2018 one of the social sciences. The broad evaluations have been organized with several panels. In the case of the biology,

Table 5.3 Modified peer review

Task	Assessment of research quality at the organizational levels (groups, departments, disciplines, universities).
Process	Reading second-order material, doing site visits, presenting assessments by reporting. Sometimes rating.
Peer panel composition	Heterogeneous. Panel members having different fields of specialization as to cover the organizational level in question.
Evaluation approach	Summative.

medicine and health evaluation, seven independent panels looked into subfields and a principal evaluation committee integrated the findings, conclusions and recommendations from the seven panels in a joint evaluation report. In the context of the national evaluation system, Norwegian universities have been reluctant to initiate modified peer review evaluation at the university level.

The opposite situation is found in Denmark and Sweden, where the universities have been the primary agenda setters in relation to modified peer review evaluation. In the Scandinavian context Copenhagen Business School was a frontrunner in implementing modified peer review at university level in the 1990s (Hansen & Borum, 1999; Borum & Hansen, 2000). After internal discussions followed by voluntary trials, an evaluation program over a number of years passed all departments through evaluation processes. The individual departments had significant room for maneuver in regard to how to organize and which material to produce to the disposal for peer review panels. The possibilities for local adaption reduced conflicts but also resulted in variations in the value of the processes.

In Sweden, Uppsala University has been a frontrunner in implementing modified peer review. In 2006/2007 Uppsala University organized a process, called the KoF07, taking 75 departments and units through peer review conducted by 24 peer panels with a total of 176 panel members. The aim of the evaluation was to find and display the “gold nuggets” in the university’s basic production units, both those which could already provide evidence of success and those that appeared to have significant potential for the future. Later, other Swedish universities followed in the footsteps of Uppsala, and Uppsala University itself repeated the evaluation exercise in 2011.

Other universities in Scandinavia has also been inspired by Uppsala and other European universities. At my own university, University of Copenhagen, modified peer review of departments some years ago was taken up at the Faculty of Social Science. Later, a common concept for the whole university was worked out and implemented across all faculties, the concept being considerably more standardized than the prior one used at the business school.

Like in Uppsala University, the Copenhagen approach was first and foremost a summative approach. In addition to inspiration from Uppsala, the University of Copenhagen also looked to the UK experiences, where modified peer review evaluation has been used as a method to grade research institutions and subsequently use the grades as a basis for distributing resources. The University of Copenhagen wanted to use the UK grades in the form of numbers in its approach. This proposal met resistance in the organization. Instead of asking the panels to grade through numbers, the panels were asked to use the prose version of the UK grades.

The Copenhagen evaluations have not been used for distributing resources but have served as input into a strategy and planning process. Panel reports were used as leadership information in the management hierarchy. Department heads had to work out plans as to how to act upon the assessment and advice in the panel reports. Deans had to work out a faculty report on the basis of the individual department reports, and the vice-chancellor, a report to the board on the basis of the reports from the deans. In this way the summative evaluation approach turned out to be used in a more formative learning-oriented process. This seems also to be the case at the department level, where department heads characterize the self-evaluation part of the process as more valuable than the panel reports.

Contrary to Uppsala University, where evaluation reports are public and easily accessible (see: <https://uu.se/en/about-uu/quality/evaluation/evaluation-of-research/>), the University of Copenhagen chose not to go public, but treat evaluation reports as internal documents. Only the short vice-chancellor report to the board is publicly accessible. The decision to not go public was anchored in discussions about whether publicness would turn self-evaluations into beautification as well as restrict the panels in presenting honest critique.

The Scandinavian comparison shows that different actors may set the scene and the agenda for research quality evaluation based on modified peer review. The comparison between Uppsala University and the University of Copenhagen further shows that the modified peer review evaluation concept is spacious. It is possible to adapt it to local organizational values and agendas.

Extended Peer Review

Peer review at organizational levels may also be implemented in a formative approach. We can term this extended peer review. The content and context of extended peer review is characterized in Table 5.4. In extended peer review focus is not on the quality of research results and contributions to knowledge production, but instead on whether the way of organizing, strategies and management processes support research quality development.

Recently, Uppsala University organized a third round of peer review evaluation. In the last round, called the KoF17, focus was less on the quality of research and more on whether the research environments were well-functioning, with a special emphasis on conditions for and processes contributing to research quality and renewal. In this round the evaluation approach was thus formative, whereas the approaches in the two first rounds were summative. In this third round 130 external peers organized into 19 panels participated. Before peers came on site visits, the university conducted a survey of research staff aiming at investigating their view on the quality of the research environments. As a consequence of the evaluation, several development initiatives have been launched among these initiatives related to the development of clear career paths. The Uppsala example illustrates how peer review evaluation at organizational levels across time can take different directions.

Modified and extended peer-review-based research evaluation initiated by universities aims at securing and developing quality in research and

Table 5.4 Extended peer review

Task	Assessment of aspects related to research quality at organizational levels (e.g. organization, structure, management, strategy).
Process	Reading second-order material, doing site visits and interviews with stakeholders, presenting assessments by reporting.
Peer panel composition	Heterogeneous. Multi-professional.
Evaluation approach	Formative.

research organizations. It also serves other purposes. One important purpose is to act as a shield toward national authorities and political initiatives. By taking ownership to quality assurance of research the universities try to protect their autonomy. In this perspective it is interesting that the Association of Swedish Higher Education Institutions, SUHF, which is an interest organization for Swedish universities and university colleges, has worked out a joint framework for quality assurance and quality development in research (<https://suhf.se/gemensamt-ramverk-for-larosatenas-kvalitetssakring-och-kvalitetsutveckling-av-forskning/>). The idea is to support quality work at individual institutions. Further, it is an attempt to influence the Swedish Higher Education Authority's (UKÄ), which has recently been asked by the Swedish government to develop a national system for the scrutiny of higher education institutions' (HEIs) quality assurance of research.

Peer Review in the Educational Field

The use of peer review is also important in the field of education, where quality assurance systems and accreditation procedures in the wave of the Bologna process have become part of the daily life in higher education institutions. In this sphere the professional and collegial peer review evaluation model has become mixed with other actor models, including the user evaluation model inviting, among others, labor market representatives and students into the evaluation process. In this sphere there are also several evaluands. The quality of educational programs (their curriculum content, pedagogical principles and, sometimes, even student learning outcomes) may be in focus and/or institutional quality assurance procedures and systems.

Sweden was in the late 1960s and early 1970s the frontrunner country in Scandinavia in relation to developing a national system for quality assurance in higher education. Pedagogical research and development projects delivered the ideational raw material for system development (Gröjer, 2004). Across time, different agencies have developed and used different concepts, moving back and forth between giving priority mostly to assessing quality in programs and assessing quality work at institutions.

These years the Swedish Higher Education Authority, UKÄ, has regained focus on scrutinizing quality work at institutions (<https://english.uka.se/quality-assurance/quality-assurance-of-higher-education.html>). Reviews focus on how well institutional quality systems help to improve the quality of programs. Six assessment areas are in focus: governance and organization, preconditions, design, implementation and outcomes, student and doctoral student perspective, working life and collaboration, and, finally, gender equality.

The review process includes several elements, among these: asking the institutions to work out self-evaluations, inviting student associations to give input to how they experience their influence on institutional quality work, panels doing site visits, posing questions to the self-evaluation and examining one or two specific quality assurance processes, so-called audit trails. Panels consist of student and doctoral student representatives, employer and working life representatives, and experts/peers from the higher education sector. The overall judgment is given on a three-point scale: approved, approved with reservation and quality assurance processes under review. The panel makes a preliminary judgment, which serves as the basis for UKÄ's decision. All review processes have follow-ups, but the form differs according to the overall judgment given.

In Denmark educational evaluation has also been organized differently across time. Ad hoc initiatives saw the light of the day in the late 1980s. Inspiration came from, among other countries, the Netherlands. In 1992 the Danish Centre for Evaluation of Higher Education was established. In 1999, the center was reorganized into the Danish Evaluation Institute. In 2007 a large re-organization was undertaken and an accreditation agency called ACE Denmark was introduced (Hansen, 2009a, 2014). In 2013 one more reform followed and the name of the agency was changed to the Danish Accreditation Institution (<https://akkr.dk/akkreditering/>).

Along with all the organizational changes, evaluation approaches have also changed. In the early years, evaluations were formative and learning oriented. When accreditation was introduced, the approach became summative. Every higher educational program at bachelor's and master's level, old and new ones, had to go through an accreditation process in order to become approved.

With the 2013 reform the system has gradually switched from accreditation of individual programs to accreditation of entire education institutions. The model for institutional accreditation gives the individual institution a free hand to organize its own quality assurance system as long as it lives up to the five criteria for quality and relevance laid down in the ministerial order. Two of the five criteria concern the quality assurance system, while the remaining three criteria concern the quality and relevance of the educational programs.

In all the shifting evaluation regimes, expert panels have been set up to perform the evaluation process. In institutional accreditations, panels typically have five members, including peers, one labor market representative and one student. The panel visits the institution to interview management, teachers, quality assurance employees and other relevant stakeholders. The accreditation institution works out a report gathering the panel's assessment.

It is emphasized to panels that they have no decision power. This rests with the Accreditation Council (<https://akkrediteringsraadet.dk>), comprised of nine members, including experts on higher education, labor market representatives and one student. On the basis of the panel report and other sources of information, the council takes decisions on whether to approve, conditionally approve or reject institutional accreditation.

There are examples of panels experiencing that their assessments are somewhat overruled. Peer discretion may be restricted by a more authoritative council culture giving priority to fairness and equal treatment across institutions. The question is whether overruling experiences in the long run have consequences for peer recruitment.

As a consequence of a decision of an institution being conditionally approved or rejected, the institution is faced with demands for accreditation of individual educational programs. The evaluation system is in this way an arena for negotiation about university autonomy parallel to the dynamics related to research evaluation.

In Norway, Norgesnettrådet was established in 1997, one of the tasks being to draw up guidelines for quality work within higher education. Some years later in 2002, the Norwegian Agency for Quality Assurance in Education (NOKUT) was established (<https://www.nokut.no/en/>).

NOKUT was given the task to review the quality assurance systems at all higher education institutions and this has been done in several rounds.

Contrary to Denmark, the universities in Norway have authority to establish educational programs. Other types of higher educational institutions have to apply for accreditation of programs. Rules for this differ across institutional types. Institutions may also apply for accreditation in another institutional category. In this way institutions can follow a path to become university, move up the hierarchy so to speak, thereby obtaining increased autonomy. And several former university colleges have become universities across time. The Norwegian system is in this respect very different from the Danish system, with the Swedish system in between. Like in Denmark and Sweden, panels including peers are important in the assessment processes. There are different rules for panel composition according to the assessment task.

NOKUT itself was evaluated in 2007/2008 by a panel composed by researchers assigned to the task by the Norwegian Ministry of Higher Education. As part of the evaluation a survey was sent to institutional leaders, administrative staff, students and academic staff investigating their experience of the impact of different forms of external evaluation. One of the findings of the study was that although NOKUT worked with different types of assessments with different purposes, impacts were perceived as quite similar regardless of the evaluation approach (Stensaker et al., 2011). One explanation discussed was that “formal procedures, rules and regulations are ‘softened’ during practice, underlining the classical distinctions between ‘talk’ and ‘action’, and between formal rules and more pragmatic practices” (Stensaker et al., 2011, p. 475). The authors further state that this possibility may occur due to the fact that all evaluations include peers who seem to contribute to change the (formal) focus of the process. Peers thus appear to have some discretion to translate the authoritative point of departure to a professional practice. Another interesting finding was that the institutional leadership and the administration were the groups identifying most positive effects of the schemes. Positive effects seem not to trickle down to academic staff.

Comparing across countries, both similarities and differences can be observed. In all three countries quality assurance systems are in place and steadily further developed. Also, peers are important members in the

panels carrying through the assessments. Further, hitherto the national systems have been specialized in evaluating the educational side of universities. This may be changing in the Swedish context, where UKÄ has been given the task of doing pilots in evaluation of institutional quality assurance of research. In relation to how agencies are organized, there are considerable differences. Structures are different and there seems to be differences in the relations between panels and the formal decision-makers, with the Danish Accreditation Council going further than the decision-makers in Norway and Sweden in restricting the discretion of panels.

Conclusion and Discussion

Focus in this chapter has been on two research questions: (1) How has peer review as an evaluation concept been developed across time? (2) What is the role of peer review today? The analysis has shown that peer review is a concept in continuous development. Where peer review evaluation formerly was carried through in the context of, to a large extent, a self-managing community of scholars, it is today carried through in a context of managerialism. Peer review evaluation has become a mediation tool between society and universities as well as a management tool within universities.

Peer review today has many faces. Classical peer review is still important first and foremost in relation to decision-making in the publication system and in relation to evaluation of dissertations. In other areas, for example, in relation to recruitment and promotion, peer review has become more standardized as sets of evaluation criteria and metrics increasingly have come into use. Also peer review practices in relation to applications of research grants in research councils seem to have become more standardized. This type of peer review was termed informed and standards-based peer review.

In recruitment and promotion contexts, managerialism encircles the peer evaluation process. On the one side, this may strengthen peer review evaluation by reducing biases and securing fairness, but on the other side, it restricts peer discretion and may be experienced by peers as not

legitimate managerial intrusion. Managerialism also embraces peer review when peer review is used to evaluate the quality of research at organizational levels such as research groups, departments and universities. This type of peer review was termed modified peer review.

But managerialism not only, to an increasing extent, encircles peer review, but has also been brought into focus for peer review. This is seen in what here has been termed extended peer review, where peer panels are asked to scrutinize the organization, management and strategy in research groups, departments, faculties and universities. Further, it is seen in quality assurance practices related to the educational side of institutions. In some contexts, it is the institutional top management which uses peer review to scrutinize lower-level management. In other contexts, it is external authorities, which, as a result of political agendas, have been given the job to have a keen eye on institutional management.

While it is meaningful conceptually to distinguish between modified and extended peer review, the examples looked into show that these types may be mixed in practice. The examples of summative modified peer review from the University of Copenhagen and Uppsala University (KoF07 and KoF11) included formative advice on how to organize and thus included an element of extended peer review. And in the example of formative extended peer review from Uppsala University (KoF17), some peers reported that they found it hard to evaluate organization and management without assessing the quality of research. Peers thus seem to be more comfortable in the summative approach, with a focus on research quality, and less comfortable in the formative, with a focus on whether organizational and management practices support the development of research quality.

Peer review obviously plays different roles in different contexts, and the different types of peer review are expected to play different roles. Classical and informed, standards-based peer review play important roles in decision-making. Modified peer review plays a role as a rewarding as well as a naming and shaming technology in some situations, followed by change and improvement initiatives, and extended peer review plays a role as a learning process also in some situations, followed by change and improvement initiatives.

However, there is more to it. Modified and extended peer review processes as well as peer review related to educational evaluation constitute arenas for struggles about autonomy and legitimacy. If universities are able to be accountable and maintain order in their own house, they may keep the authorities at a distance, thereby protecting institutional autonomy. University-initiated modified and extended peer review processes thus aim at securing that the authorities experience the activities as legitimate, and authority-initiated modified and extended peer review processes aim at securing that the political leadership and citizens experience the activities as legitimate. In this way peer review processes link to policy processes and questions about how to distribute resources in university systems.

Still, one can wonder what the balance is between costs and values both in modified and in extended peer review. As the examples have shown, these processes often involve many peers as well as a considerable amount of university staff, and thus occupy many working hours, which could have been used for doing research. Likewise, quality assurance in education often demands considerable paperwork, occupying considerable administrative resources and probably building up a new administrative layer specialized in evaluation. Future studies should look into this. Further, future studies should pay attention to the linkages between managerialism and peer review practices, and put focus on the consequences of managerialism encircling peer review practices.

References

- Benos, D. J., Bashari, E., Chaves, J. M., Gaggar, A., Kapoor, N., LaFrance, M., Mans, R., Mayhew, D., McGowan, S., Polter, A., Qadri, Y., Sarfare, S., Schultz, K., Splittgerber, R., Stephensom, J., Tower, C., Walton, R. G., & Zotov, A. (2007). The ups and downs of peer review. *Advances in Physiology Education, 31*, 145–152.
- Bornmann, L. (2011). Scientific peer review. *Information, Science and Technology, 45*(1), 199–225.
- Borum, F., & Hansen, H. F. (2000). The local construction and enactment of standards for research evaluation: The case of the Copenhagen Business School. *Evaluation, 6*(3), 281–299.

- Brunsson, N., & Sahlin-Andersson, K. (2000). Constructing organizations: The example of public sector reform. *Organizations Studies*, 21(4), 721–746.
- Dahler-Larsen, P. (2012). *The evaluation society*. Stanford University Press.
- De Boer, H. F., Enders, J., & Schimank, U. (2007). On the way towards new public management? The governance of university systems in England, the Netherlands, Austria and Germany. In D. Jansen (Ed.), *New forms of governance in research organizations: Disciplinary approaches, interfaces and integration* (pp. 137–152). Springer.
- Geschwind, L., Hansen, H. F., Pinheiro, R., & Pulkkinen, K. (2019). Governing performance in the Nordic universities: Where are we heading and what have we learned? In R. Pinheiro et al. (Eds.), *Reforms, organizational change and performance in higher education*. Palgrave Macmillan/Springer Nature. https://doi.org/10.1007/978-3-030-11738-2_9
- Gröjer, A. (2004). *Den utvärdera(n)de staten. Utvärderingens institutionalisering på den högre utbildningens område*. Statsvetenskapliga Institutionen, Stockholms Universitet.
- Hamann, J., & Beljean, S. (2017). Academic evaluation in higher education. In J. C. Shin & P. Teixeira (Eds.), *Encyclopedia of international higher education systems and institutions*. Springer Science+Business Media.
- Hansen, H. F. (2005). Choosing evaluation models. *Evaluation*, 11, 447–462.
- Hansen, H. F. (2009a). Educational evaluation in Scandinavian countries: Converging or diverging practices? *Scandinavian Journal of Educational Research*, 53(1), 71–87.
- Hansen, H. F. (2009b). *Research evaluation: Methods, practice, and experience*. Danish Agency for Science Technology and Innovation.
- Hansen, H. F. (2014). ‘Quality agencies’: The development of regulating and mediating organizations in Scandinavian higher education. In M. Chou & Å. Gornitzka (Eds.), *Building the knowledge economy in Europe. New constellations in European research and higher education governance*. Edward Elgar Publishing Limited.
- Hansen, H. F., Aarrevaara, T., Geschwind, L., & Stensaker, B. (2019b). Evaluation practices and impact: Overload? In R. Pinheiro et al. (Eds.), *Reforms, organizational change and performance in higher education*. Palgrave Macmillan/Springer Nature. https://doi.org/10.1007/978-3-030-11738-2_8
- Hansen, H. F., & Borum, F. (1999). The construction and standardization of evaluation: The case of the Danish university sector. *Evaluation*, 5(3), 303–329.
- Hansen, H. F., Geschwind, L., Kivisto, J., Pekkola, E., Pinheiro, R., & Pulkkinen, K. (2019a). Balancing accountability and trust: University reforms in the Nordic countries. *Higher Education*, 78(3), 557–573.

- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(22), 429–431.
- Langfeldt, L., & Kyvik, S. (2011). Researchers as evaluators: Tasks, tension and politics. *Higher Education*, 62, 199–212.
- Langfeldt, L., & Kyvik, S. (2015). Inre spänningar och framtida utmaningar för *peer review*. In *RJ:s årsbok 2015/2016* (pp. 153–169). Riksbankens Jubileumsfond & Makadam Förlag.
- Neave, G. (1998). The evaluative state reconsidered. *European Journal of Education*, 33(3), 265–284.
- Olsen, J. P. (2007). The institutional dynamics of the European university. In P. Maassen & J. P. Olsen (Eds.), *University dynamics and European integration* (pp. 25–54). Springer.
- Scriven, M. (1991). *Evaluation thesaurus*. Sage.
- Seeber, M., Lepori, B., Montauti, M., Enders, J., de Boer, H., Weyer, E., Bleiklie, I., et al. (2015). European Universities as Complete Organizations? Understanding identity, hierarchy and rationality in public organizations. *Public Management Review*, 17(10), 1444–1474.
- Stensaker, B., Langfeldt, L., Harvey, L., Huisman, J., & Westerheijden, D. (2011). An in-depth study on the impact of external quality assurance. *Assessment & Evaluation in Higher Education*, 36(4), 465–478.
- Vedung, E. (1997). *Public policy and program evaluation*. Transaction Publishers.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

