



1

Peer Review in Academia

Eva Forsberg, Lars Geschwind, Sara Levander,
and Wieland Wermke

Introduction

Over the past few decades, peer review has become an object of great professional and managerial interest (Oancea, 2019) and, increasingly, academic scrutiny (Bornmann, 2011; Grimaldo et al., 2018). Nevertheless, calls for further research are numerous (Tennant & Ross-Hellauer, 2020). This volume is in answer to such interest and appeals. We aim to present a variety of peer-review practices in contemporary academic life as well as

E. Forsberg (✉) • S. Levander
Department of Education, Uppsala University, Uppsala, Sweden

L. Geschwind
Department of Learning in Engineering Sciences, KTH Royal Institute
of Technology, Stockholm, Sweden

W. Wermke
Department of Special Education, Stockholm University, Stockholm, Sweden

the principled foundation of peer review in scientific communication and authorship. This volume is unique in that it covers many different practices of peer review and their theoretical foundations, providing both an introduction into the very complex field and new empirical and conceptual accounts of peer review for the interested reader. The contributions are produced by internationally recognized scholars, almost all of whom participated in the conference ‘Scientific Communication and Gatekeeping in Academia in the 21st Century’, held in 2018 at Uppsala University, Sweden.¹ The overall objective of this volume is explorative; framings relevant to the specific contexts, practices and discourses examined are set by the authors of each chapter. However, some common conceptual points of departure may be laid down at the outset.

Peer review is a context-dependent, relational concept that is increasingly used to denote a vast number of evaluative activities engaged in by a wide variety of actors both inside and outside of academia. By peer review, we refer to peers’ assessments and valuations of the merits and performances of academics, higher education institutions, research organizations and higher education systems. Mostly, these activities are part of more encompassing social evaluation practices, such as reviews of manuscripts, grant proposals, tenure and promotion and quality evaluations of institutions and their research and educational programmes. Thus, scholarly peer review comprises evaluation practices within both the wider international scientific community and higher education systems. Depending on differences related to scientific communities and national cultures, these evaluations may include additional gatekeepers, internal as well as external to academia, and thus the role of the peer may vary.

The roots of peer review can be found in the assessment practices of reviewers and editors of scholarly journals in deciding on the acceptance of papers submitted for publishing. Traditionally, only peers (also known as referees) with recognized scholarly standing in a relevant field of research were acknowledged as experts (Merton, 1942/1973). Due to the differentiation and increased use of peer review, the notion of a peer

¹ Funded by Riksbankens Jubileumsfond (F17-1350:1). The keynotes of the conference are accessible on video at <https://media.medfarm.uu.se/play/kanal/417>. For more information on the conference, see www.konferens.edu.uu.se/scga2018-en.

employed in various evaluation practices may be extended. Who qualifies as an expert in different peer-review practices and with what implications are empirical issues.

Even though peer review is a familiar phenomenon in most scholarly evaluations, there is a paucity of studies on peer review within the research field of evaluation. Peer review has, however, been described as the most familiar collegial evaluation model, with academic research and higher education as its paradigm area of application and with an ability to capture and judge qualities as its main advantage (Vedung, 2002). Following Scriven (2003), we define evaluation as a practice ‘determining the merit, worth or significance of things’ (p. 15). Scriven (1980) identifies four steps involved in evaluation practices, which are also frequently used in peer review, either implicitly enacted and negotiated or explicitly stated (Ozeki, 2016). These steps concern (1) the criteria of merit, that is, the dimensions of an object being evaluated; (2) the standards of merit, that is, the level of performance in a given dimension; (3) the measuring of performance relative to standards; and (4) a value judgement of the overall worth.

Consequently, the notion of peer review refers to evaluative activities in academia conducted by equals that distribute merit, value and worth. In these processes of selection and legitimation, issues referring to criteria, standards, rating and ranking are significant. Often, peer reviews are embedded in wider evaluation practices of research, education and public outreach. To capture contemporary evaluations of academic work, we will include a number of different review practices, including some in which the term *peer* is employed in a more extended sense.

The Many Face(t)s of Peer-Review Practices

Depending on the site in which peer review is used, the actors involved differ, as do their roles. The same applies to potential guidelines, purposes, discourses, use of professional judgement and metrics, processes and outcome of the specific peer-review practice. These are all relative to the site in which the review is used and will briefly be commented upon below.

The Interplay of Primary and Secondary Peer Review

It is possible to make a distinction between primary and secondary peer reviews (British Academy, 2007). As stated, the primary role of peer review is to assess manuscripts for publishing, followed by the examination and judgement of grant applications. Typically, many other peer-review practices, so-called secondary peer review, involve summaries of outcomes of primary reviews. Thus, we might view primary and secondary reviews as folded into each other, where, for example, reviews of journal articles are prerequisite to later evaluation of the research quality of an institution, in recruitment and promotion, and so forth (Helgesson, 2016). Hence, the consequences of primary reviews can hardly be overstated.

Traditionally, both forms of primary peer review (assessment of manuscripts and grant applications) are *ex ante* evaluations; that is, they are conducted prior to the activity (e.g. publishing and research). With open science, open access journals and changes in the transparency of peer review, open and public peer reviews have partly opened the black box of reviews and the secrecy of the process and its actors (Sabaj Meruane et al., 2016). Accordingly, publishing may include both *ex ante* and *ex post* evaluations. These forms of evaluation can also be found among secondary reviews, with degree-awarding accreditation an example of the former and reviews of disciplines an example of the latter.

Sites and Reviewer Tasks and Roles

Without being exhaustive, we can list a number of sites where peer review is conducted as part of more comprehensive evaluations: *international, regional and national higher education agencies* conduct accreditation, quality audits and evaluations of higher education institutions; *funding agencies* distribute grants for projects and fellowships; *higher education institutions* evaluate their research, education and public outreach at different levels and assess applications for recruitment, tenure and promotion; *the scientific community* assesses manuscripts for publication, evaluates doctoral theses and conference papers and allocates awards. The

evaluation roles are concerned with the provision of human and financial resources, the evaluation of research products and the assessment of future strategies as a basis for policy and priorities. All of these activities are regularly performed by researchers and interlinked in an evaluation spiral in which the same research may be reviewed more than once (Langfeldt & Kyvik, 2015). If we consider valuation and assessment more generally, the list can be extended almost infinitely, with supervision and seminar discussions being typical activities in which valuation plays a central part. Hence, scholars are accustomed to being assessed and to evaluating others.

The role and the task of the reviewer differ also in relation to whether the act of reviewing is performed individually, in teams or in a blending of the two forms. In the evaluation of research grants, the latter is often the case, with reviewers first individually rating or ranking the applications, followed by panel discussions and joint rankings as bases for the final decision made by a committee. In peer review for publishing, there might be a desk rejection by the editor, but if not, two or more external reviewers assess a manuscript and recommend that the editor accept, revise or reject it. It is then up to the editor to decide what to do next and to make the final decision. The process and the expected roles of the involved editor, reviewer and authors may vary depending on whether it is a private publisher or a journal linked to a scientific association, for example. Whether the reviewer should be considered an advisor, an independent assessor, a juror or a judge depends on the context and the task set for the reviewer within the specific site and its policies and practices as well as on various praxes developed over time (Tennant & Ross-Hellauer, 2020).

Power-making in the Selection of Expertise

The selection process is at the heart of peer review. Through valuations and judgements, peers are participants in decisions on inclusion and exclusion: What project has the right qualities to be allocated funding? Which paper is good enough to be published? And who has the right track record to be promoted or offered a fellowship? When higher

education institutions and scholars increasingly depend on external funding, peer review becomes key in who gets an opportunity to conduct research and enter or continue a career trajectory as a researcher and, in many systems, a higher education teacher. In other words, peer review is a cornerstone of the academic career system (Merton, 1968; Boyer, 1990) and heavily influences what kinds of scientific knowledge will be furthered (Lamont, 2009; Aagaard et al., 2015).

The interaction involved in peer review may be remote, online or local, including face-to-face collaboration, and it may involve actors with different interests. Moreover, interaction may be extended to the whole evaluation enterprise. For example, evaluations of higher education institutions and their research and education often include members of national agencies, scholarly experts and external stakeholders. Scholarly experts may be internal or external to the higher education institutions and of lower, comparable or higher rank than the subjects of evaluation, and reviewers may be blind or known to those being evaluated and vice versa. Scholarly expertise may also refer to a variety of specialists, for example, to scholars with expertise in a specific research topic, in evaluation technology, in pedagogy or public outreach. A more elaborated list of features to be considered in the allocation of experts to various review practices can be found in a peer-review guide by the European Science Foundation (2011). At times the notion of *peer* is extended beyond the classical idea to one with demonstrated competence to make judgements within a particular research field. Who qualifies as a reviewer is contingent on who has the authority to regulate the activity in which the evaluation takes place and who is in the position to suggest and, not least, to select reviewers. This is a delicate issue, imbued with power, and one that we need to further explore, preferably through comparative studies involving different peer-review practices in varying contexts.

Acting as a peer reviewer has become a valuable asset in the scholarly track record. This makes participating as a reviewer important for junior researchers. Therefore, such participation not only is a question of being selected but also increasingly involves self-election. More opportunities are provided by ever more review activities and the prevalence of evaluation fatigue among senior researchers. The limited credit, recognition and rewards for reviewers may also contribute to limited enthusiasm amongst

seniors (Research Information Network CIC, 2015). Moreover, several tensions embedded in review practices can add to the complexity of the process and influence the readiness to review. The tensions involve potential conflicts between the role of the reviewer or evaluator and the researcher's role: time conflict (research or evaluate), peer expertise versus impartiality (especially qualified colleagues are often excluded under conflict-of-interest rules), neutral judge versus promoter of research interests (double expectation, deviant assessments versus unanimous conclusions, peer review versus quantitative indicators, and scientific autonomy versus social responsibility) (Langfeldt & Kyvik, 2015). Despite noted challenges, classical peer review is still the key mechanism by which professional autonomy and the guarding of research quality are achieved. Thus, it is argued that it is an academic duty and an obligation, in particular for senior scholars, to accept tasks as reviewers (Caputo, 2019). Nevertheless, the scholarly exchange value should be addressed in future discussions of gatekeeping in academia.

The Academic Genres of Peer Review

Peer reviews are rooted in more encompassing discourses, such as those concerning norms of science, involving notions of quality and excellence founded in different sites endogenous or exogenous to science. Texts subject to or employed or produced in peer-review practices represent a variety of academic genres, including review reports, editors' letters, applicants' proposals, submitted manuscripts, guidelines, applicant dossiers and curriculum vitae (CVs), testimonials, portfolios and so on. Different genres are interlinked in chains, creating systems of genres. A significant aspect of systems is intertextuality, or the fact that texts within a specific system refer to, anticipate and shape each other. The interdependence of texts is about how they relate to situational and formal expectations—in this case, of the specific peer-review practice. It is also about how one text makes references to another text; for example, review reports often refer to guidelines, calls, announcements or texts in application dossiers. The interdependence can also be seen in how the texts interact in academic communities (Chen & Hyon, 2005): who the

intended readers of a given text are, what the purpose of the text is, how the text is used in the review and decision process, and so on. Conclusively, the genre systems of peer review vary depending on epistemic traditions, national culture and regulations of higher education systems and institutions.

Given this diversity, we are dealing with a great number of genre systems involving different kinds of texts and interrelations embedded in power and hierarchies. A significant feature of peer-review texts as a category is the occluded genres, that is, genres that are more or less closed to the public (Swales, 1996). Depending on the context, the list of occluded genres varies. For example, the submission letters, submitted manuscripts, review reports and editor–author correspondence involved in the eventual publication of articles in academic journals are not made publicly available, while in the context of recruitment and promotion, occluded genres include application letters, testimonials and evaluation letters to committees. And for research grants, the research proposals, individual review reports and panel reports tend to remain entirely internal to the grant-making process. However, in some countries (e.g. in Sweden, due to the principle of openness, or *offentlighetsprincipen*), several of these types of texts may be publicly available.

The request for open science has also initiated changes to the occluded genres of peer review. After a systematic examination, Ross-Hellauer (2017) proposed ‘open peer review’ as an umbrella term for a variety of review models in line with open science, ‘including making reviewer and author identities open, publishing review reports and enabling greater participation in the peer review process’ (p. 1). From 2005 onwards, there has been a big upswing of these definitions. This correlates with the rise of the openness agenda, most visible in the review of journal articles and within STEM and interdisciplinary research.

Time and space are central categories in most peer-review genres and the systems to which they belong. While review practices often look to the past, imagined futures also form the background for valuation. The future orientation is definitely present in audits, in assessments of grant proposals and in reviews of candidates’ track records. The CV, a key text in many review practices, may be interpreted in terms of an applicant’s

career trajectory, thus emphasizing how temporality and spatiality interact within a narrative infrastructure, for example how scholars move between different academic institutions over time (Hammarfelt et al., 2020). Texts may also feed both backwards and forwards in the peer-review process. For example, guidelines and policy on grant evaluations and distribution may be negotiated and acted upon by both applicants and reviewers. Candidates may also address reviewers as significant others in anticipating the forthcoming reviewer report (Serrano Velarde, 2018). These expectations on the part of the applicant can include prior experiences and perceptions of specific review practices, processes and outcomes in specific circumstances.

Turning to the reviewer report, it is worth noting that they are often written in English, especially ones assessing manuscripts and frequently those on research proposals and recruitment applications as well. Commonly seen within the academic genre of peer review is the use of indirect speech, which can be linked to the review report's significance as related to the identity of the person being evaluated (Paltridge, 2017). Two key notions, politeness and face, have been used to describe the evaluative language of review reports and how reviewers interact with evaluatees. There are differences related to content and to whether a report is positive or negative overall in its evaluation. For example, reviewers of manuscripts invoke certain structures of knowledge, using different structures when suggesting to reject, revise or accept and when asking for changes. To maintain social relationships, reviewers draw on different politeness strategies to save an author's face. Strategies employed may include 'apologizing ('I am sorry to have to') and impersonalizing an issue ('It is generally not acceptable to')' (Paltridge, 2017, p. 91). Largely, requests for changes are made as directions, suggestions, clarifications and recommendations. Thus, for both evaluatees and researchers of peer reviews, particular genre competences are required to decode and act upon the reports. For beginning scholars unfamiliar with the world of peer review or for scholars from a different language or cultural background than the reviewer, it might be challenging to interpret, negotiate and act upon reviewer reports.

Criteria and the Professional Judgement of Quality

According to the classical idea of peer review, only a peer can properly recognize quality within a given field. Although, in both research and scholarly debate, shortcomings have been emphasized regarding the trustworthiness, efficacy, expense, burden and delay of peer review (Bornmann, 2013; Research Information Network CIC, 2015), many critics still find peer review as the least-worst system, in the absence of viable alternatives. Overall, scholars stand behind the idea of peer review even though they often have concerns regarding the different practices of peer review (Publons, 2018).

Calls for accountability and social relevance have been made, and there have been requests for formalization, standardization, transparency and openness (Tennant & Ross-Hellauer, 2020). While the idea of formalization of peer review refers to rules, including the development of policy and guidelines for different forms of peer review, standardization rather emphasizes the setting of standards through the employment of specific tools for evaluation (i.e. criteria and indicators used for assessment, rating or ranking and decision-making). An interesting question is whether standardization will impact the extent and the way peers are used in different sites of evaluation (Westerheijden et al., 2007). We may add, who will be considered a peer and what will the matching between the evaluator and the evaluation object or evaluatee look like?

It is widely acknowledged that criteria is an essential element of any procedure for judging merit (Scriven, 1980; Hug & Aeschbach, 2020). This is the case regardless of whether criteria are determined in advance or if they are explicitly expressed or implicitly manifested in the process of assessment. The notion of peer review has been supplemented in various ways, implicating changes to the practice and records of peer review. Increasingly, review reports combine classical peer review with metrics of different kinds. Accordingly, quantitative measures, taken as proxies for quality, have entered practices of peer review. Today, blended forms are rather common, especially in evaluations of higher education institutions, where narrative and metric summaries often supplement each other and inform a judgement.

In general, quantitative indicators (e.g. number of publications, journal impact factors, citations) are increasingly applied, even though their capacity to capture quality is questioned, especially within the social sciences, humanities and the arts. Among the main reasons given for the rapid growth of demands for metrics, one of the arguments we find is that classic peer review alone cannot meet the quest for accountability and transparency, and bibliometric evaluations may appear cheaper, more objective and legitimated. Moreover, metrics may give an impression of accessibility for policy and management (Gläser & Laudel, 2007; Söderlind & Geschwind, 2019). However, tensions between classical peer review and quantitative indicators have been identified and are hotly debated (Langfeldt & Kyvik, 2011). The dramatic expansion of the use of metrics has brought with it gaming and manipulation practices to enhance reputation and status, ‘including coercive citation, forced joint authorship, ghostwriting, h-index manipulation, and many others’ (Oravec, 2019, p. 859). Warnings are also issued against the use of bibliometric indicators at the individual level. A combination of peer narratives and metrics is, however, considered a possibility to improve an overall evaluation, given due awareness of the limitations of quantitative data as proxies for quality.

The literature on peer review has focused more on the weighting of criteria than on the meaning referees assign to the criteria they use (Lamont, 2009). Even though some criteria, such as originality, trustworthiness and relevance, are frequently used in the assessment of academic work and proposals, our knowledge of how reviewers ascribe value to, assess and negotiate them remains limited (Hug & Aeschbach, 2020). However, Joshua Guetzkow, Michèle Lamont and Grégoire Mallard (2004) show that panellists in the humanities, history and the social sciences define originality much more broadly than what is usually the case in the natural sciences.

Criteria, indicators and comparisons are unstable: they are situational and dependent on context and a referee’s personal experience of scientific work (Kaltenbrunner & de Rijcke, 2020). We are dealing here with assessments in situations of uncertainty and of entities not easily judged or compared. The concept of judgement devices has been used to capture how reviewers delegate the judgement of quality to proxies, reducing the

complexity of comparison. For example, the employment of central categories in a CV, which references both temporal and spatial aspects of scholars' trajectories, makes comparison possible (Hammarfelt, 2017). In a similar way, the theory of anchoring effects has been used to explore reviewers' abilities to discern, assess, compare and communicate what scientific quality is or may be (Roumbanis, 2017). Anchoring effects have their roots in heuristic principles used as shortcuts in everyday problem solving, especially when a judgement involves intuition. Reduction of complexity is visible also in how reviewers first collect criteria that consist of information that has an eliminatory function. Next, they search for positive signs of evidence in order to make a final judgement (Musselin, 2002). Dependent on context and situations, reviewers tend to select different criteria from a repertoire of criteria (Hug & Aeschbach, 2020).

On the one hand, the complexity of academic evaluations requires professional judgement: scholars sufficiently grounded in a field of research and higher education are entrusted with interpreting and negotiating criteria, indicators and merits. Still, the practice of peer review has to be safeguarded against the risk of conservatism as well as epistemic and social biases (Kaltenbrunner & de Rijcke, 2020). On the other hand, changes in the governance of higher education institutions and research, as well as marketization, managerialism, digitalization and calls for accountability, have increased the diversity of peer review and introduced new ways to capture and employ criteria and indicators. The long-term consequences of these changes need to be monitored, not least because of how they challenge the self-regulation and autonomy of the academic profession (Oancea, 2019).

How to understand, assess, measure and value quality in research, the career of a scholar or the performances of a higher education institution are complex issues. Turning to the notion of quality in a general sense will not solve the problem, since it has so many facets and has been perceived in so many different ways, including as fitness for purpose, as eligible, as excellent and as value for money (Westerheijden et al., 2007), all notions in need of contextualization and further elaboration to achieve some sense (see also Elken & Wollscheid, 2016).

When presenting a framework to study research quality, Langfeldt et al. (2020) identify three key dimensions: (1) quality notions

originating in research fields and in research policy spaces; (2) three attributes important for good research and drawn on existing studies, namely, originality/novelty, plausibility/reliability and value or usefulness; and (3) five sites where notions of research quality emerge, are contested and are institutionalized, comprising researchers, knowledge communities, research organizations, funding agencies and national policy arenas. This multidimensional framework and its components highlight issues that are especially relevant to studies of peer review. The sites identify arenas where peer review functions as a mechanism through which notions of research quality are negotiated and established. The consideration of notions of quality endogenous and exogenous to scientific communities and the various attributes of good research can also be directly linked to referees' distribution of merit, value and worth in peer-review practices under changing circumstances.

The Autonomy of a Profession and a Challenged Contract

Historical analyses link peer review to the distribution of authority and the negotiations and reformulations of the public status of science (Csiszar, 2016). At stake in renegotiations of the contract between science and society are the professional autonomy of scholars and their work. Peer review is contingent on the prevailing contract and is critical in maintaining the credibility and legitimacy of research and higher education (Bornmann, 2011). The professional autonomy of scholars raises the issue of self-regulation. Its legitimacy ultimately comes down to who decides what, particularly concerning issues of research quality and scientific communication (Clark, 1989).

Over the past 40 years, major changes have taken place in many OECD (Organisation for Economic Co-operation and Development) countries in the governance of public science and higher education, changes which have altered the relative authority of different groups and organizations (Whitley, 2011). The former ability of scientific elites to exercise endogenous control over science has, particularly since the 1960s,

become more contested and subject to public policy priorities. A more heterogeneous and complex higher education system has been followed by the exogeneity of governance mechanisms, formal rules and procedures, and the institutionalization of quality assurance procedures and performance monitoring. Expectations of excellence, competition for resources and reputation, and the coordination of research priorities and intellectual judgement have changed across disciplinary and national boundaries to varying degrees (Whitley, 2011). These developments can be seen as expressions of the evaluative state (Neave, 1998), the audit society (Power, 1997) and as part of an institutionalized evaluation machinery (Dahler Larsen, 2012).

Changes in the principles of governance are underpinned by persistent tensions around accountability, evaluation, measurement, demarcation, legitimation, agency and identity in research (Oancea, 2019). Besides the primary form of recognition through peer review, the weakened autonomy of academic fields has added new evaluative procedures and institutions. Academic evaluations, such as accreditations, audits and quality assurances, and evaluations of research performance and social impact now exist alongside more traditional forms (Hansen et al., 2019).

Higher education institutions worldwide have experienced the emergence and manifestations of the quality movement, which is part of inter-related processes such as massification, marketization and managerialism. Through organizations at international, national and institutional levels, a variety of technologies have been introduced to identify, measure and compare the performance of higher education institutions (Westerheijden et al., 2007). These developments have emphasized external standards and the use of bibliometrics and citation indexes, which have been criticized for rendering the evaluations more mechanical (Hamann & Beljean, 2017). Mostly, peer review, often in combination with self-evaluation, is also employed in the more recently introduced forms of evaluation (Musselin, 2013). Accordingly, peer review, in one form or another, is still a key mechanism monitoring the flow of scientific knowledge, ideas and people through the gates of the scientific community and higher education institutions (Lamont, 2009).

Autonomy may be defined as 'the quality or state of being self-governing' (Ballou, 1998, p. 105). Autonomy is thus the capacity of an

agent to determine their own actions through independent choice, in this case within a system of principles and laws to which the agent is dedicated. The academic profession governs itself by controlling its members. Academics control academics, peers control peers, in order to maintain the status and indeed the autonomy of the profession. Fundamentally, professionals are licensed to act within a valuable knowledge domain. By training, examination and acknowledgement, professionals are legitimated (at least politically) experts of their domain. The rationale of licence and the esotericism of professional knowledge raise the question of how professionals and their work can be evaluated and by which standards. There are rules of conduct and ethical norms, but these are ultimately owned and controlled by the academic profession. From this perspective, we can understand peer review as the structural element that holds academia together.

The increase of peer-review practices in academia can be compared with other professions that also must work harder than before to maintain their status and autonomy. In many cases, their competence and quality must be displayed much more visibly today. Pluralism and individualism in society have also resulted in a plurality of expertise and a decrease of mono-vocational functional systems. A mystique of academic knowledge (as in 'the research says') is not as acceptable in public opinion today as it once was. The term 'postmodern professionals' is suggested to describe experts who expend more effort in the dramaturgy of their competences than people in their positions might have in the past in order to generate trust in clients and in society (Pfadenhauer, 2003). Media makes professional competences, performances and failures much more visible and contributes to trust or mistrust in professions. In a pluralist society, extensive use of peer review may indeed function as a strategy to make apparent quality visible and secure the autonomy of the academic profession, which owns the practice of peer review and knows how to adjust it to its needs.

While most academic evaluations exist across scientific communities and disciplines, the criteria of evaluation can differ substantially between and within communities (Hamann & Beljean, 2017). Thus, research on peer review needs to take disciplinary and interdisciplinary similarities and differences seriously. Obviously, the impact of the intellectual and

social organization of the sciences (Whitley, 1984), the mode of research (Nowotny et al., 2001), the tribes and territories (Becher, 1989; Becher & Trowler, 2001; Trowler et al., 2014) and the epistemic cultures (Knorr Cetina, 1999) need to be better represented in future research. Then, examinations of peer review may contribute also to a fuller understanding of the contract between science and society and the challenges directed towards the professional autonomy of academics.

Why Study Peer Review?

As an ideal, peer review has been described as ‘the linchpin of science’ (Ziman, 1968, p. 148) and a key mechanism in the distribution of status and recognition (Merton, 1968) as well as part and parcel of collegiality and meritocracy (Cole & Cole, 1973). Above all, peer review is considered a gatekeeper regarding the quality of science both in various specialized knowledge communities and in research policy spaces (Langfeldt et al., 2020). Peer review is often taken as a hallmark of quality, expected to both guard and enhance quality. Early on, peer review, or refereeing, was linked to moral institutionalized imperatives. Perhaps most known are those formulated in the *Ethos of Science* by Merton (1942/1973): communism, universalism, disinterestedness and organized scepticism, or CUDOS. These norms and their counter-norms (individualism, particularism, interestedness and dogmatism) have frequently been the focus of peer-review studies. Norms on how scientific work is or should be carried out and how researchers should behave reflect the purpose of science, and ideas of how science should be governed, and are thus directly linked to the autonomy of the academic profession (Panofski, 2010). In short, research into peer review goes to the very heart of academia and its relation to society. This calls for scrutiny.

With changing circumstances, peer review is more often employed, and its purposes, forms and functions are increasingly diversified. Today, academic evaluations permeate every corner of the scientific enterprise, and the traditional form of peer review, rooted in scientific communication, has migrated. Thus, we have seen peer review evolve to be undertaken in all key aspects of academic life: research, teaching, service and

collaboration with society (Tennant & Ross-Hellauer, 2020). Increasingly, peer review is regarded as the standard, not only for published scholarship but also for academic evaluations in general. Ideally, peer review is considered to guarantee quality in research and education while upholding the norms of science and preserving the contract between science and society. The diversity and the migration of review practices and its consequences should be followed closely.

In the course of a career, scholars are recurrently involved as both reviewers and reviewees, and this is becoming more and more frequent. As stated in a report on peer review by the British Academy (2007), the principle of *judge not, that ye be not judged* is impossible to follow in academic life. On the contrary, the selection of work for publishing, the allocation of grants and fellowships, decisions on tenure and promotion, and quality evaluations all depend upon the exercise of judgement. ‘The distinctive feature of this academic judgement is that it is reciprocal. Its guiding motto is: judge only if you in turn are prepared to be judged’ (British Academy, 2007, p. vii).

Indeed, we lack comprehensive statistics on peer review and the involvement of scholars in its diverse practices. However, investigations like the Wiley study (Warne, 2016) and Publons’ (2018) *Global State of Peer Review* (2018), both focused on reviews of manuscripts, implicate the widespread and increasing use of peer review. In 2016, roughly 2.9 million peer-reviewed articles were indexed in Web of Science, and a total of 2.5 million manuscripts were rejected. Estimated reviews each year amount to 13.7 million. Together, the continuous rise of submissions and the increase in evaluations using peer reviews expose the system and its actors to ever more pressure.

Peer-review activities produce an incredible amount of talk and gossip in academia. In particular, academic appointments have contributed to the organizational ‘sagas’ described by Clark (1972). In systems where fierce competition for a limited number of chairs (professorships) is the norm, much is at stake. A single decision, one way or another, can make or break an academic career, and the same is true in relation to recurring judgements and decisions on tenure and promotion (Gunneriusson, 2002). Research on the emotional and socio-psychological consequences of peer rejection or low ratings and rankings is seldom conducted. While

rejection may function as either a threat or a challenge to scholarly identities, Horn (2016) argues that rejection is a source of stigmatization pervading the entire academic community. In a similar vein, scholars have to adjust to the maxim of ‘publish or perish’ and the demands of reviewers, even when these are against the scholars’ own convictions. Some researchers consider this a form of ‘intellectual prostitution’ (Frey, 2003), and reviewer fatigue is spreading through the scientific community. For example, it is widely recognized that editors sometimes have trouble finding reviewers. Obviously, peer review has become a concern to scholars of all kinds and to their identities and everyday practices and careers.

The mundane reality of peer-review practice is quite different from the ideology of peer review, and our knowledge is rather restricted and fragmented (Grimaldo et al., 2018). The roots of peer review can be traced through the seventeenth century and book censorship, the development of academic journals in the eighteenth century and the gatekeeping of scientific communication. As a regular activity, peer review is, however, a latecomer in the scientific community, and it is unevenly distributed across nations and disciplines (Biagioli, 2002). For example, publication practices, discourses and the lingua franca differ between knowledge communities. Traditional peer review is a more prominent feature of the natural sciences and medicine than of the humanities, the social sciences and the arts. This is also reflected in research on peer review. In a similar way, data show that US researchers supply by far the most reviews of manuscripts for journals, while China reviews substantially less. Nevertheless, review output is increasing in all regions and especially so in emerging regions (Publons, 2018).

Even though there are differences, peer review is a fundamental tool in the negotiation and establishment of a scholars’ merits and research, of higher education quality and of excellence. Peer review is also considered a tool to prevent misconduct, such as the fraudulent presentation of findings or plagiarism. Thus, peer review may fulfil functions of gatekeeping, maintenance and enhancement. Peer reviews can also be linked to struggles over which form of capital should be the gold standard and over gaining as much capital as possible (Maton, 2005). At stake is, on the one hand, scholastic capital, and on the other hand, academic capital linked to administrative power and control over resources (Bourdieu, 1996).

The introduction of ever new sites for peer review, changing qualifications of reviewers and calls for open science, as well as the increased use of metrics, increase the need for further research. Moreover, the cost and the amount of time spent on different kinds of reviews and their potential impact on the identity, recognition and status of scholars and higher education institutions make peer review especially worthy of systematic studies beyond professional narratives and anecdotes. Peer review has both advocates and critics, although the great majority of researchers are positive to the idea of peer review. Many critics find peer review costly, time consuming, conservative and socially and epistemically biased. In sum, there are numerous reasons to study peer review. It is almost impossible to overstate the central role of peer review in the academic enterprise, and the results of empirical evidence are inconclusive and the research field emergent and fragmented (Bornmann, 2011; Batagelj et al., 2017).

State of the Art of Research on Peer Review

There is a lack of consensus on what peer review is and on its purposes, practices, outcomes and impact on the academic enterprise (Tennant & Ross-Hellauer, 2020). The term *peer review* was relatively unknown before 1970. *Referee* was the more commonly applied notion, used primarily in relation to the evaluation of manuscripts and scientific communication (Batagelj et al., 2017). This lack of clarity has affected how the research field of peer review has been identified and described.

During the past few decades, a number of researchers have provided syntheses of research on peer review in the forms of quantitative meta- and network analyses as well as qualitative configurative analyses. Some are more general in character (Sabaj Meruane et al., 2016; Batagelj et al., 2017; Grimaldo et al., 2018), though the main focus is often research in the natural and medical sciences and peer review for publishing and, to some extent, for grant funding. Others are more concerned with either a specific practice of peer review or different critical topics. Below, we mainly use these recent systematic reviews to depict the research field of peer review, to identify the limits of our knowledge on the subject and to elaborate why we need to study it further.

Academic evaluations, like peer reviews, have been examined from a number of perspectives (Hamann & Beljean, 2017). From a functionalist approach, we can explore how well evaluative procedures serve their purposes—especially those of validity, reliability and fairness—and how well they handle various potential biases. The power-analytical perspective makes critical inquiries into dysfunctional effects of structural inequalities like nepotism and unequal opportunities for resource accumulation. The perspective on the performativity of evaluations and evaluative devices focuses on the organizational impact of the devices, on ranking and on the ways indicators incite strategic behaviour. The social-constructive perspective on evaluation emphasizes that ideas such as merits and originality are socially and historically context dependent. There is also a pragmatist perspective that stresses the situatedness of evaluative practices and interactions (e.g. how panellists reach consensus). More and more frequently used are analytical tools from the field of the sociology of valuation and evaluation, which emphasizes knowledge production as contextualization and the existence and impact of insecurities in the performative situations (Lamont, 2012; Mallard et al., 2009; Serrano Velarde, 2018). Some researchers highlight the variety of academic communities and the intradisciplinary, interdisciplinary and transdisciplinary aspects of research today as significant explanatory factors for evaluative practices (Hamann & Beljean, 2017). We may add changes in the governance of higher education institutions and research and the introduction of new evaluation practices as equally important (Whitley, 2011; Oancea, 2019).

In a network analysis of research on peer review from 1950 to 2016 Batagelj et al. (2017) identified 23,000 indexed records in Web of Science and, above all, a main corpus of 47 articles and books. These texts, which were cited in the most influential publications on peer review, focus on science, scholarship, systematic reviews, peers, peer reviews and quantitative and qualitative analysis. The most cited article allows for an expansion of this list to include the institutionalization of evaluation in science, open peer reviews, bias and the effects of peer review on the quality of research. Most items belonging to the corpus were published relatively early, with only a few published after the year 2000. However, overview papers were published more recently, mainly in the past decade.

The research field of peer review has been described as an emergent field marked by three development stages (Batagelj et al., 2017). The first stage, before 1983, includes seminal work mostly presented in social science and philosophy journals. Main topics include scientific productivity, bibliographies, knowledge, citation measures as measures of scientific accomplishment, scientific output and recognition, evaluations in science, referee systems, journal evaluations, the peer-evaluation system, review processes and peer-review practices. During the second stage, 1983–2002, biomedical journals were influential. Key topics focused on the effects of blinding on review quality, research into peer review, guidelines for peer reviewing, monitoring peer-review performance, open peer review, bias in the peer-review system, measuring the quality of editorial peer review, and the development of meta-analysis and systematic reviews approaches. Finally, in the third stage, 2003–2016, we find research on peer review mainly in specialized science studies journals such as *Scientometrics*. The most frequent topics include peer review of grant proposals, bias, referee selection and links between editors, referees and authors.

Another quantitative analysis (Grimaldo et al., 2018) of articles published in English from 1969 to 2015 and indexed in the citation database Scopus found very few publications before 1970, and fewer than around 100 per year until 2004. Then, from 2004 to 2015 the numbers increased rapidly, 12% per year on average. Half the records were journal articles, books, chapters and conference papers, and the rest were mostly editorial notes, commentaries, letters and literature reviews. Scholars from English-speaking countries, especially the United States, predominated, but authors from prominent European institutions were also found. A fragmented, potentially interdisciplinary research field dominated by medicine, sociology and behavioural sciences and with signs of uneven sharing of knowledge was identified. The research was typically pursued in small collaborative networks. Articles on peer reviews were published mostly by *JAMA*, *Behavioral and Brain Science* and *Scientometrics*. The most important topics were peer review in relation to quality assurance and improvement, publishing, research, open access, evaluation and assessment, bibliometrics and ethics. Among the authors of the top five most influential articles we find Merton, Zuckermann, Horrobin, Bornmann and

Siegelmann. Grimaldo et al.'s (2018) analysis revealed the presence of structural problems, such as difficulties in accessing data, partly due to confidentiality and lack of interest from editorial boards, administrative bodies and funding agencies. More positively, the analysis pointed to digitalization and open science as favourable tools for increases in research, cooperation and knowledge sharing.

In an overview (Sabaj Meruane et al., 2016) of empirical studies on peer-review processes, almost two thirds of the first-named authors had doctoral backgrounds in medicine, psychology, bibliometrics or scientometrics, and around one fifth in sociology of science or science and technology studies. There is definitely a lack of integration of other fields, such as those within the social sciences, the humanities and the arts and education in the study of peer-review processes. The following topics were empirically researched, in descending order: sociodemographic variables (83%), sociometric or scientometric data (47%), evaluation criteria (36%), bias (31%), rates of acceptance/rejection/revision (25%), predictive validity (24%), consensus among reviewers (17%) and discourse analysis of isolated or related texts (14%). The analysis indicates that 'the texts interchanged by the actors in the process are not prominent objects of study in the field' (Sabaj Meruane et al., 2016, p. 188). Further, the authors identified a number of gaps in the research: The field conceives of peer review more as a system than as a process. Moreover, bibliometric studies constitute an independent field of empirical research on peer review. Only a few studies combine analysis of indicators with content or functional analysis. In a similar way, research on science production, reward systems and evaluation patterns rarely includes actual texts that are interchanged in the peer-review process. Discourse analysis, in turn, rarely uses data other than the reviewer report and socio-demographics. Due to ethical issues and confidentiality, discourse studies and text analyses of reviewer reports are less frequent.

It might be risky to state that peer review is an under-studied object of research, considering the vast number of publications devoted to the topic. Nevertheless, it appears that the field of peer-review research has yet to be fully defined, and empirical research in the field has to be more comprehensively done. A common problem the authors consider important to examine is the consequences of the same actor being able to fulfil

different roles (e.g. author, reviewer, editor) in various single reviews. Above all, the field requires not only further but also more comprehensive approaches, and in addition, the black box of peer review needs to be fully open (Sabaj Meruane et al., 2016).

Among syntheses focusing on specific topics, those of trustworthiness and bias as well as the employment and negotiation of and the meaning ascribed to criteria in various evaluation practices or in different disciplines are relatively common. In a review of literature published on the topic of peer review, the state of research on journal, fellowship and grant peer review is analysed, focusing on three quality criteria: reliability, fairness and predictive validity (Bornmann, 2011). The interest was directed towards the norms of science, ensuring that results were not incidental, that certain groups or individuals were not favoured or disadvantaged, and that selection of publications and scholars were aligned to scientific performances. Predictive validity was far less studied in primary research than reliability and fairness. Another overview articulates and critiques conceptions and normative claims of bias (Lee et al., 2013). The authors raise questions about existing norms and conclude that peer review is social and that a diversity of norms and opinions among communities and referees may be desirable and beneficial. Bias is also studied in research on who gets tenure with respect to both meritocratic and non-meritocratic factors, such as ascription and social and academic capital (Lutter & Schröder, 2016). These authors show that network size, individual reputation and gender matter.

Epistemic differences point to the necessity of studying peer review within a variety of disciplines and transdisciplinary contexts. An interview study of panellists serving on fellowship grants within the social sciences and humanities shows that evaluators generally draw on four epistemological styles: constructivist, comprehensive, positivist and utilitarian (Mallard et al., 2009). Moreover, peer reviewers employ the epistemological style most appropriate to the field of the proposal under review. In the future, more attention has to be paid to procedural fairness, including from a comparative perspective. In another systematic review of criteria used to assess grant applications, it is suggested that forthcoming research should also focus on the applicant, include data from

non-Western countries and examine a broad spectrum of research fields (Hug & Aeschbach, 2020).

As shown in this introductory chapter, the research field devoted to peer review covers a great number of evaluation practices embedded in different contexts. As it is an emergent and fragmented field in need of integration, there are certainly many possible ways to make contributions to the research field of peer review. On the agenda we find issues related to the foundation of science: the ethos of science and the ideology of peer review, the production and dissemination of knowledge, professional self-regulation and open science. There are also questions concerning the development of theoretical framing and methodological tools adapted to the study of diverse review practices in shifting contexts and at various interacting levels. Not least, in response to calls for more comprehensive and integrated research, it is necessary to open the black boxes of peer review and analyse, in empirical studies, the different purposes, discourses, genres, relations and processes involved.

A single book cannot take on all the above-mentioned challenges ahead of us. However, following this brief introduction to the field, the volume brings together research on review practices often studied in isolation. We include studies ranging from the practice of assessing manuscripts submitted for publication to the more recent practice of open review. In addition, more encompassing and general issues are considered, as well as specificities of different peer-review practices. This is further developed below, where the structure of the volume and the contributions of each chapter are presented.

The Structure and Content of the Volume

The structure of the volume falls into three main parts. In the first part, Rudolf Stichweh and Raf Vanderstraeten continue the introduction begun in this chapter. They discuss the term peer review and the contexts of its emergence. In Chap. 2, Rudolf Stichweh explains the genesis of inequalities and hierarchies in modern science. He illuminates the forms and mechanisms of scientific communication on the basis of which the social structures of science are built: publications, co-authorships and

multiple authorships, citations as units of information and as social rewards, and peer review as an evaluation of publications (and of projects and careers). Stichweh demonstrates how, in all institutional dimensions of higher education, differences arise between successful and less successful participations. Success generates influence and social attractiveness (e.g. as a co-author). Influential and attractive participants are recruited into positions where they assess the achievements of others and thereby limit and control inclusion in publications, funding and careers.

Vanderstraeten, in Chap. 3, puts forward that with the expansion of educational research in the twentieth century, interested ‘amateurs’ have been driven out of the field, and the scientific community of peers has become the dominant point of orientation. Authorship and authority became more widely distributed; peer review was institutionalized to monitor the flow of ideas within scientific literature. Reference lists in journals demonstrated the adoption of cumulative ideals about science. Vanderstraeten’s historical analysis of education journals shows the social changes that contributed to the ascent of an ‘imagined’ community of expert peers in the course of the twentieth century.

Part II of this volume focuses mainly on how peer-review practices have emerged in many parts of higher education institutions. From being scholarly publication practices in early times, peer review appears to be internationally *the* most significant performative practice in higher education and research. In this part, the various scholars provide insight into such processes. Don F. Westerheijden, in Chap. 4, revisits the policy issue of the balance between peer review and performance indicators as the means to assess quality in higher education. He shows the paradoxes and unintended effects that emerge when peer review is the main method in the quality assurance procedures of higher education institutions as a whole. Westerheijden argues that attempted solutions of using self-assessments and performance indicators as well as specifically trained assessors increase complaints about bureaucracy from within the academic community.

In Chap. 5, Hanne Foss Hansen sheds light on how peer review as an evaluation concept has developed over time and discusses which roles peer review plays today. She presents a typology distinguishing between classical peer review, informed and standard-based peer review, modified

peer review and extended peer review. Peer review today can be found with all these faces. Peter Dahler Larsen argues in Chap. 6 that gatekeepers in institutional review processes who know the future and use this knowledge in a pre-emptive or precautionary way play a key role in the construction of reality, which comes out of Bibliometric Research Indicators, widely used internationally. By showing that human judgement sometimes enhances or multiplies the effects of 'evaluation machineries', this chapter contributes to an understanding of mechanisms that lead to constitutive effects of evaluation systems in research.

In Chap. 7, Agnes Ers and Kristina Tegler Jerselius explore a national framework for quality assurance in higher education and argue that such systems' forms are dynamic, since they change over time. Ers and Tegler Jerselius show how the method of peer review has evolved over time and in what way it has been affected by changes made in the system. *Gustaf Nelhans* engages in Chap. 8 with the performative nature of bibliometric indicators and explores how they influence scholarly practice at macro levels (in national funding systems), meso levels (within universities) and individual levels (in the university employees' practice). Nelhans puts forward that the common-sense 'representational model of bibliometric indicators' is questionable in practice, since it cannot capture the qualities of research in any unambiguous way.

In Chap. 9, Lars Geschwind and Kristina Edström discuss the loyalty of academic staff to their disciplines or scientific fields. They show how this loyalty is reflected in evaluation practices. They elaborate on the extent to which peer reviewers act as advocates for those they evaluate. By doing so, Geschwind and Edström problematize potential evaluator roles. In Chap. 10, Malcom Tight closes *Part II* of this book. Drawing on his extensive review experiences in various areas of higher education institutions, he assesses how 'fit for purpose' peer review is in twenty-first-century academe. He focuses on different practices of peer review in the contemporary higher education system and questions how well they work, how they might be improved and what the alternatives are.

Whereas *Part II* of this volume focuses on the relation and impact of higher education institutions considering education quality and research output, *Part III* illuminates different particular peer-review practices. Eva Forsberg, Sara Levander and Maja Elmgren examine in Chap. 11

peer-review practices in the promotion of what is called ‘excellent’ or ‘distinguished’ university teachers. While research merits have long been the prioritized criteria in the recognition of institutions and scholars, teaching is often downplayed. To counteract this tendency, various systems to upgrade the value of education and to promote teaching excellence have been introduced by higher education institutions on a global scale. The authors show that the intersection between promotion, peer review and excellent teaching affects not only the peer-review process but also the notion of the excellent or distinguished university teacher.

In Chap. 12, Tine S. Prøitz discusses how the role of scholarly peers in systematic review is analysed and presented. Peer evaluation is an essential element of quality assurance of the strictly defined methods of systematic review. The involvement of scholarly peers in the systematic review processes has similarities with traditional peer-review processes in academic publishing, but there are also important differences. In systematic review, peers are not only re-judging already reviewed and published research, but also gatekeeping the given standards, guidelines and procedures of the review method.

Liv Langfeldt presents in Chap. 13 processes of grant peer review. There are no clear norms for assessments, and there may be a large variation in what criteria reviewers emphasize and *how* they are emphasized. Langfeldt argues that rating scales and budget restrictions can be more important than review guidelines for the kind of criteria applied by the reviewers. The decision-making methods applied by the review panels when ranking proposals are found to have substantial effects on the outcome. Chapters 14 and 15 focus on peer-review practices in the recruitment of professors. First, Sara Levander, Eva Forsberg, Sverker Lindblad and Gustav Jansson Bjurhammer analyse the initial step of the typecasting process in the recruitment of full professors. They show that the field of professorial recruitment is characterized by heterogeneity and no longer has a basis in one single discipline. New relations between research, teaching and society have emerged. Moreover, the authority of the professorship has narrowed and the amount of responsibilities have increased. Then, Björn Hammarfeldt focuses on discipline—specific practices for evaluating publications oeuvres. He examines how ‘value’ is enacted with special attention to the kind of tools, judgements, indicators and metrics

that are used. Value is indeed enacted differently in the various disciplines.

In the last chapter of the book, Chap. 16, Tea Vellamo, Jonna Kosonen, Taru Siekkinen and Elias Pekkola investigate practices of tenure track recruitment. They show that criteria of this process can exceed notions of individual merits and include assessments of the strategic visions of universities and departments. The use of the tenure track model can be seen as a shift both for identity building related to a university's strategy and for using more managerial power in recruitment more generally.

We dedicate this book to our beloved colleague and friend, professor Rita Foss Lindblad, who was involved in the project but passed away in 2018.

References

- Aagaard, K., Bloch, C., & Schneider, J. W. (2015). Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator. *Research Evaluation*, 24(2), 106–117.
- Ballou, K. A. (1998). A concept analysis of autonomy. *Journal of Professional Nursing*, 14(2), 102–110.
- Batagelj, V., Ferligoj, A., & Squazzoni, F. (2017). The emergence of a field: A network analysis of research on peer review. *Scientometrics*, 113(1), 503–532. <https://doi.org/10.1007/s11192-017-2522-8>
- Becher, T. (1989). *Academic tribes and territories: Intellectual inquiry and the cultures of disciplines*. Society for Research into Higher Education.
- Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories. Intellectual inquiry and the culture of disciplines*. Open University Press.
- Biagioli, M. (2002). From book censorship to academic peer review. *Emergences Journal for the Study of Media & Composite Cultures*, 12(1), 11–45. <https://doi.org/10.1080/1045722022000003435>
- Bornmann, L. (2011). Scientific peer review. *Annual Review of Information, Science and Technology*, 45, 197–245. <https://doi.org/10.1002/aris.2011.1440450112>
- Bornmann, L. (2013). Evaluations by peer review in science. *Springer Science Reviews*, 1(1–4). <https://doi.org/10.1007/s40362-012-0002-3>
- Bourdieu, P. (1996). *Homo academicus*. Polity.

- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. The Carnegie Foundation for the Advancement of Teaching.
- British Academy. (2007). Peer review: The challenges for the humanities and social sciences. Retrieved December 1, 2020, from <https://www.thebritishacademy.ac.uk/documents/197/Peer-review-challenges-for-humanities-social-sciences.pdf>
- Caputo, R. K. (2019). Peer review: A vital gatekeeping function and obligation of professional scholarly practice. *Families in Society: The Journal of Contemporary Social Services*, 100(1), 6–16. <https://doi.org/10.1177/1044389418808155>
- Chen, R., & Hyon, S. (2005). Faculty evaluation as a genre system: Negotiating intertextuality and interpersonality. *Journal of Applied Linguistics*, 2(2), 153–184. <https://doi.org/10.1558/japl.v2i2.153>
- Clark, B. R. (1972). The organizational saga in higher education. *Administrative Science Quarterly*, 17, 178–184.
- Clark, B. R. (1989). The academic life: Small worlds, different worlds. *Educational Researcher*, 18(5), 4–8. <https://doi.org/10.2307/1176126>
- Cole, J. R., & Cole, S. (1973). *Social stratification in science*. University of Chicago Press.
- Csiszar, A. (2016). Peer review: Troubled from the start. *Nature*, 532(7599), 306–308. <https://doi.org/10.1038/532306a>
- Dahler Larsen, P. (2012). *The evaluation society*. Stanford University Press.
- Elken, M., & Wollscheid, S. (2016). *The relationship between research and education: Typologies and indicators. A literature review*. Nordic Institute for Innovative Studies in Research and Education (NIFU).
- European Science Foundation. (2011). *European peer review guide. Integrating policies and practices into coherent procedures*.
- Frey, B. S. (2003). Publishing as prostitution? Choosing between one's own ideas and academic success. *Public Choice*, 116(1/2), 205–223. <https://doi.org/10.1023/A:1024208701874>
- Gläser, J., & Laudel, G. (2007). The social construction of bibliometric evaluations. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences. The advent of research evaluation systems*. Springer.
- Grimaldo, F., Marušić, A., & Squazzoni, F. (2018). Fragments of peer review: A quantitative analysis of the literature (1969–2015). *PLOS ONE*, 13(2), e0193148. <https://doi.org/10.1371/journal.pone.0193148>

- Guetzkow, J., Lamont, M., & Mallard, G. (2004). What is originality in the humanities and the social sciences? *American Sociological Review* 2004, 69, 190. <https://doi.org/10.1177/000312240406900203>
- Gunneriusson, H. (2002). *Det historiska fältet: svensk historievetenskap från 1920-tal till 1957*. Uppsala: Acta Universitatis Upsaliensis.
- 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*. https://doi.org/10.1007/978-94-017-9553-1_295-1
- Hammarfelt, B. (2017). Recognition and reward in the academy: Valuing publication oeuvres in biomedicine, economics and history. *Aslib Journal of Information Management*, 69(5), 607–623. <https://doi.org/10.1108/AJIM-01-2017-0006>
- Hammarfelt, B., Rushforth, D., & de Rijcke, S. (2020). Temporality in academic evaluation: ‘Trajectorial thinking’ in the assessment of biomedical researchers. *Valuation Studies*, 7(1), 33–63. <https://doi.org/10.3384/VS.2001-5992.2020.7.1.33>
- Hansen, H. F., Aarrevaara, T., Geschwind, L., & Stensaker, B. (2019). Evaluation practices and impact: Overload? In R. Pinheiro, L. Geschwind, H. Foss Hansen, & K. Pulkkinen (Eds.), *Reforms, organizational change and performance in higher education: A comparative account from the Nordic countries*. Palgrave Macmillan.
- Helgesson, C.-F. (2016). Folded valuations? *Valuation Studies*, 4(2), 93–102. <https://doi.org/10.3384/VS.2001-5992.164293>
- Horn, S. A. (2016). The social and psychological costs of peer review: Stress and coping with manuscript rejection. *Journal of Management Inquiry*, 25(1), 11–26. <https://doi.org/10.1177/1056492615586597>
- Hug, S. E., & Aeschbach, M. (2020). Criteria for assessing grant applications: A systematic review. *Palgrave Communications*, 6(30). <https://doi.org/10.1057/s41599-020-0412-9>
- Kaltenbrunner, W., & de Rijcke, S. (2020). Filling in the gaps: The interpretation of curricula vitae in peer review. *Social Studies of Science*, 49(6), 863–883. <https://doi.org/10.1177/0306312719864164>
- Knorr Cetina, K. (1999). *Epistemic cultures*. Harvard University Press.
- Lamont, M. (2009). *How professors think. Inside the curious world of academic judgment*. Harvard University Press.
- Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual Review of Sociology*, 38(21), 201–221. <https://doi.org/10.1146/annurev-soc-070308-120022>

- Langfeldt, L., & Kyvik, S. (2011). Researchers as evaluators: Tasks, tensions and politics. *Higher Education*, 62(2), 199–212. <https://doi.org/10.1007/s10734-010-9382-y>
- Langfeldt, L., & Kyvik, S. (2015). Intrinsic tensions and future challenges of peer review. In *RJ Yearbook 2015/2016*. Riksbankens Jubileumsfond & Makadam Publishers.
- Langfeldt, L., Nedeva, M., Sörlin, S., & Thomas, D. A. (2020). Co-existing notions of research quality: A framework to study context-specific understandings of good research. *Minerva*, 58, 115–137. <https://doi.org/10.1007/s11024-019-09385-2>
- Lee, C. J., Sugimoto, G. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2–17. <https://doi.org/10.1002/asi.22784>
- Lutter, M., & Schröder, M. (2016). Who becomes a tenured professor, and why? Panel data evidence from German sociology, 1980–2013. *Research Policy*, 45, 999–1013. <https://doi.org/10.1016/j.respol.2016.01.019>
- Mallard, G., Lamont, M., & Guetskow, J. (2009). Fairness as appropriateness: Negotiating epistemological differences in peer review. *Science Technology Human Values*. <https://doi.org/10.1177/0162243908329381>
- Maton, K. (2005). A question of autonomy: Bourdieu's field approach and higher education policy. *Journal of Education Policy*, 20(6), 687–704. <https://doi.org/10.1080/02680930500238861>
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159(3810), 56–63. <https://doi.org/10.1126/science.159.3810.56>
- Merton R. K. (1973). *The sociology of science: Theoretical and empirical investigations* (Norman W. Storer, Ed.). University of Chicago Press. (Original work published 1942)
- Musselin, C. (2002). Diversity around the profile of the 'good' candidate within French and German universities. *Tertiary Education and Management*, 8(3), 243–258. <https://doi.org/10.1080/13583883.2002.9967082>
- Musselin, C. (2013). How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate. *Research Policy*, 42(5), 1165–1173. <https://doi.org/10.1016/j.respol.2013.02.002>
- Neave, G. (1998). The evaluative state reconsidered. *European Journal of Education*, 33(3), 265–284. <https://www.jstor.org/stable/1503583>
- Nowotny, H., Scott, P. B., & Gibbons, M. T. (2001). *Re-thinking science: Knowledge and the public in an age of uncertainty*. Polity Press.

- Oancea, A. (2019). Research governance and the future(s) of research assessment. *Palgrave Communications*, 5, 27. <https://doi.org/10.1057/s41599-018-0213-6>
- Oravec, A. (2019). Academic metrics and the community engagement of tertiary education institutions: Emerging issues in gaming, manipulation, and trust. *Tertiary Education and Management*. <https://doi.org/10.1007/s11233-019-09026-z>
- Ozeki, S. (2016). *Three Empirical Investigations into the Logic of Evaluation and Valuing Practices*. Dissertations. 2470. <https://scholarworks.wmich.edu/dissertations/2470>
- Paltridge, B. (2017). *The discourse of peer review. Reviewing submission to academic journals*. Macmillan Publishers.
- Panofski, A. L. (2010). In C. J. Calhoun (Ed.), *Robert K. Merton: Sociology of science and sociology as science*. Columbia University Press.
- Pfadenhauer, M. (2003). *Professionalität. Eine wissenssoziologische Rekonstruktion institutionalisierter Kompetenzdarstellungskompetenz* [Professionalism. A reconstruction of institutionalized proficiency in displaying competence]. Springer.
- Power, M. (1997). *The Audit Society. Rituals of verification*. Oxford University Press.
- Publons. (2018). Global state of peer review. Online.
- Research Information Network CIC. (2015). Scholarly communication and peer review. The current landscape and future trends. A report commissioned by the Wellcome Trust. Retrieved May 2015, from <https://wellcome.org/sites/default/files/scholarly-communication-and-peer-review-mar15.pdf>
- Ross-Hellauer, T. (2017). What is open peer review? A systematic review (version 2; peer review: 4 approved). *F1000Research*, 2017, 6(588). Last updated: 17 May 2019. Included in Science Policy Research Gateway. <https://doi.org/10.12688/f1000research.11369.2>
- Roumbanis, L. (2017). Academic judgments under uncertainty: A study of collective anchoring effects in Swedish research council panel groups. *Social Studies of Science*, 47(1), 95–116. <https://doi.org/10.1177/0306312716659789>
- Sabaj Meruane, O., González Vergara, C., & Pina-Stranger, Á. (2016). What we still don't know about peer review. *Journal of Scholarly Publishing*, 47(2), 180–212. <https://doi.org/10.3138/jsp.47.2.180>
- Scriven, M. (1980). *The logic of evaluation*. Edgepress.
- Scriven, M. (2003). Evaluation theory and metatheory. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation* (pp. 15–30). Kluwer Academic Publishers.

- Serrano Velarde, K. (2018). The way we ask for money... The emergence and institutionalization of grant writing practices in academia. *Minerva*, 56(1), 85–107. <https://doi.org/10.1007/s11024-018-9346-4>
- Söderlind, J., & Geschwind, L. (2019). Making sense of academic work: The influence of performance measurement in Swedish universities. *Policy Reviews in Higher Education*, 3(1), 75–93. <https://doi.org/10.1080/23322969.2018.1564354>
- Swales, J. M. (1996). Occluded genres in the academy. The case of the submission letter. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues*. ProQuest Ebook Central. <http://ebookcentral.proquest.com/lib/uu/detail.action?docID=680373>
- Tennant, J. P., & Ross-Hellauer, T. (2020). The limitations to our understanding of peer review. *Research Integrity and Peer Review*, 5(6). <https://doi.org/10.1186/s41073-020-00092-1>
- Trowler, P., Saunders, M., & Bamber, V. (Eds.). (2014). *Tribes and territories in the 21st century. Rethinking the significance of disciplines in higher education*. Routledge.
- Vedung, E. (2002). Utvärderingsmodeller [Evaluation models]. *Socialvetenskaplig tidskrift*, 9(2–3), 118–143.
- Warne, V. (2016). Rewarding reviewers—sense or sensibility? A Wiley study explained. *Learned Publishing*, 29, 41–50. <https://doi.org/10.1002/leap.1002>
- Westerheijden, D. F., Stensaker, B., & Joao Rosa, M. (Eds.). (2007). *Quality assurance in higher education. Trends in regulation, translation and transformation*. Springer.
- Whitley, R. (1984). *The intellectual and social organization of the sciences*. Clarendon Press.
- Whitley, R. (2011). Changing governance and authority relationships in the public sciences. *Minerva*, 49, 359–385. <https://doi.org/10.1007/s11024-011-9182-2>
- Ziman, J. M. (1968). *Public knowledge*. The University of Chicago Press.

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.

