LIV LANGFELDT & SVEIN KYVIK

Intrinsic tensions and future challenges of peer review

Researchers spend a great deal of time assessing research, an activity with several intrinsic tensions and potential role conflicts. Future challenges to peer review include the question of how demands for transparency and usefulness to society can possibly be met while, at the same time, academic independence is sustained.

❦

ONE IMPORTANT TASK included in the researcher role is evaluating other researchers' work. The intention is to promote good, at the expense of poor, research. Evaluators give or deny their colleagues access to research grants, publication and advancement through the collegial process of peer review. The problem, according to the British physicist and theorist of science John Ziman, is that this calls for large amounts of the most valuable resource in research: the most competent researchers' time.[1] Research evaluation has also increased in both importance and scope in recent years, partly owing to the fact that a larger share of resources is being distributed in the form of project grants. Moreover, requirements of higher quality in university research, greater relevance to society and improved efficiency (value for money) have resulted in a wave of evaluations focusing on the organisational level.

Different evaluation roles and tasks

The researcher's role as evaluator comprises at least nine distinct tasks:

1. assessing doctoral theses
2. selecting new staff and promoting staff

3. distributing research grants
4. assessing manuscripts submitted for publication
5. reviewing books and writing review articles
6. assessing candidates for academic awards
7. evaluating research organisations
8. assessing research as the basis for policy and decisions
9. assessing future research strategies and priorities.

The first three tasks are about providing human and financial resources for the research system; the next four concern the products of research; and the last two relate to future strategies. Some of these evaluation tasks are included in university employees' work duties (assessing theses and applicants for university positions), while others are performed as fee-based extra services (evaluation work for other organisations) and still others are carried out as unpaid work in the researcher's own subject area (assessing journal articles).

The expectation that academic staff will assist in performing these tasks is embedded in norms, and seldom in formal rules or regulations. The role as evaluator is voluntary; but it offers a chance to exert academic power and is important for the distribution of prestige and credibility in the research community. Extensive literature about peer review exists, but it deals mainly with scrutiny of applications for research grants and of manuscripts submitted to academic journals. Studies of the other assessment roles are more limited. The key aspects of the nine different roles are presented below.

1. Examiner. In most countries doctoral students account for a substantial proportion of total university research, and reviewing doctoral theses is an important assessment role. First, this is a way of certifying the PhD candidate as a researcher; second, it identifies promising and talented researchers. Scrutinising theses is also an important way of staying up to date with new research and promising young researchers. An analysis of thesis examinations suggests that the examiners assume a role similar to that of readers of any new text,[2] but assessing theses is nevertheless one of the more demanding evaluation tasks. Theses are, in general, long; and thorough reviews, often both oral and written, are expected.

2. Staff selector. Assessing applicants for academic positions includes, for example, deciding whether an applicant fulfils the requirements to be appointed as an associate or full professor. In this evaluator role, research-

ers can influence a particular department's future academic profile, and also the gender composition of its staff.[3] These are processes that may cause conflicts.[4]

3. Distributor of research grants. A third important role is that of serving on committees set up to distribute research grants, or acting as an individual referee of applications for these grants. Since a rising proportion of resources for research is distributed through research councils, foundations and other organisations that support research, where peer review forms the basis of these awards, the importance of this role has grown. The downside of this is the increasing share of rejections, which are frustrating both for the applicants concerned and for the evaluation panels. Consequently, the use of peer review for distribution of research grants has been called into question and there have long been proposals to replace it, also for reasons beyond the argument that the assessment takes up valuable research time.[5] On the one hand it is difficult to judge research that has not yet been done, and there is not always much point in doing so. On the other, partial assessments may have a far-reaching influence on the type of research that is funded and who fails to receive funding. If only certain schools of thought are represented in an evaluation panel, the consequence may be that other approaches are discouraged.[6] Ensuring impartial decisions in multidisciplinary assessment committees is another challenge.[7] Peer review of project applications is thus a controversial form of assessment and the one that could most easily be replaced, for example by direct allocation to departments on the basis of previous performance.

4. Referee and editor. While the referees assess the quality and relevance of manuscripts submitted to journals for publication, the editor takes the final decision on the basis of advice from, normally, two or more people.[8] A closely related but less formalised task is to assess abstracts and papers for academic conferences. There is an extensive literature on peer review for publication, and there are also special conferences in this area.[9] Reviewer bias is a central issue in this discussion.

5. Review writer. This function includes writing book reviews and review articles about the current state of knowledge in a defined research field. Review writers assign praise and blame to the researchers concerned, and can also indicate their opinions by neglecting articles, books and other academic contributions.

6. Prize awarder. Awarding academic prizes and other distinctions is

regarded as a prestigious duty and can take up a great deal of time. Evaluation work of this kind can also provide scope to exert power and join, at a high level, in knowledge policy. This is evident in, for example, studies of the work underlying the academic Nobel Prizes.[10]

7. Evaluator of research organisations. Evaluating research groups, departments, programmes, institutes and even whole universities has become a great deal more extensive over the past quarter-century, in response to the needs of the 'evaluative state'.[11] This relates not only to assessments of quality; it may also include concrete recommendations on how to improve the research and organise the unit evaluated. There are major differences between countries regarding the purpose of these evaluations and how they are carried out.[12]

8. Policy adviser. Another important task is to review research to be used as the basis for policy-making and regulation. For example, extensive review work is performed by researchers employed at public agencies in areas like health and the environment. Academic committees of various kinds also serve public bodies with summaries of this kind, and researchers serve on international bodies like the IPCC (UN Intergovernmental Panel on Climate Change). The combination of peer review and practical policy poses, however, a number of challenges, such as possible disagreement among researchers and potential bias.[13]

9. Foresight viewer. In the 1980s and '90s, public research funding bodies were encouraged by their governments to become more strategic in their funding policy by such means as creating large-scale R&D programmes focusing on particularly promising new research areas. John Irvine and Ben R. Martin, the British science-policy researchers, used the term 'foresight activities' for the mechanisms and procedures used to identify areas of basic research with major future potential.[14]

All in all, researchers seem constantly engaged in evaluating and being evaluated: they belong to an evaluation spiral in which the same research is reviewed repeatedly. A project may undergo its first evaluation (1) when the researchers apply for one or more research grants, and later (2) when abstracts for conferences, (3) journal articles, (4) doctoral theses, if any, and (5) book manuscripts from the project are submitted. In addition, (6) publications from the project undergo scrutiny when the authors apply for academic positions, (7) when they are nominated for prizes and awards, and (8) when their departments or (9) the research programme that funded their projects are evaluated. One or more (10) reviews of

books from the project may be published in academic journals, and (11) publications from the project may come to be evaluated in review articles. If the project yields results of political or economic importance, or contributes interesting openings for new research, it may also become part of (12) policy-making processes and foresight studies. Finally, the results from the project are fed back into stage 1, when researchers' previous achievements are assessed after they have submitted new applications for research grants, and so the evaluation spiral continues. In addition, ongoing evaluation is an integral part of the research process. Thus, besides all the formal peer review, continuous scrutiny of approaches, methods and results takes place.

There are numerous ways in which peer review affects research and the research community, and consequently the role of evaluator may, in several different respects, come into conflict with the researcher role.

Tension 1:
Time for research versus time for evaluation

The most obvious tension between the roles of evaluator and researcher is the time conflict. The more time researchers spend on evaluation, the less is available for their research.

Little is known about how much time is devoted to formal evaluation tasks. A study we carried out in 2001 among all permanent academic employees at Norwegian universities investigated four of the nine evaluation roles: scrutiny of doctoral theses; assessment of applicants for

vacant positions and promotion to more senior ones, review of manuscripts for publication, and assessment of research organisations.[15] Altogether, some 80% of the academic staff were engaged in at least one of these four evaluator roles in the course of a year, and these individuals spent an average of nearly 17 working days on these tasks. The task that most were engaged in was review of manuscripts. The evaluator role took most time for professors (19 working days). Since the survey did not cover all the nine evaluation roles, the total number of days is higher than the figures observed. If we add to the figures above a few days for scrutiny of research applications, the result obtained for an average researcher (one of the 80% who took on any evaluation tasks) is approximately 20 working days a year (some 25 for professors[16]), corresponding to between four and five working weeks, spent on evaluation work.

The question of whether the evaluator roles have changed over time can be partially answered by our questionnaire data. The proportion of employees who joined in reviewing doctoral theses and selecting staff rose from 46% to 58% between 1991 and 2000. The average number of days spent on these tasks was the same on the two occasions, indicating an increase in the total time spent on these two different tasks during the 1990s. For the other tasks, too, there is reason to believe that there has been an increase in total pressure on the staff. Referee work has hardly decreased in the past few decades, partly because a growing share of resources for university research is distributed in response to applications, which has the effect of getting more researchers engaged in the work of assessment. But the reviewer role that has probably increased its time requirement most is evaluation of research organisations. This is because of the introduction of new public management, with its greater emphasis on retrospective assessments. Another change emerging from the questionnaire data is a much more international profile of the evaluator role. In 1991, 9% of academic staff at Norwegian universities took part in evaluation work abroad; the proportion rose to 22% in 2000 and 38% in 2012.[17] Summing up, there is thus every reason to assume that both the importance of and the time required for the role of evaluator have increased.

It may reasonably be assumed that the time conflict makes it impossible for the researchers most in demand to undertake all the evaluation tasks they are invited to take on. They then limit their inputs to what they perceive as the most important assignments. The probable result is

that the highest-ranking academics will assume the most prestigious and powerful evaluation tasks[18], leaving the less prestigious ones to more junior researchers.

Tension 2:
Peer expertise versus impartiality

Some severe tensions between the roles of evaluator and researcher derive from conflicts of interest. The researchers who are the most competent to evaluate particular research are often close colleagues or competitors of those who have done the research[19], but close colleagues are disqualified under conflict-of-interest rules.

Conflicts of interest may be an argument for taking precautions to ensure the autonomy of scholarship, while in other cases it can be an argument for alternatives to peer review, i.e. inducing the research community to refrain some of its autonomy to avoid 'setting the fox to mind the geese'. For example, if a researcher with commercial interests in the drugs industry is asked to scrutinise manuscripts about effects of drugs, it would jeopardise the autonomy and credibility of science. It may therefore be asserted that only 'pure' academics should perform such tasks. On the other hand, taking part in the review of proposals for a research programme may disqualify the evaluators and their research group from applying for grants from the programme. Such losses of rights are in conflict with researchers' academic interests and may make them more hesitant about taking on such review assignments.

These tensions can be tackled by restricting autonomy of the individual research fields or limiting the use of peer review. The former could, for example, involve using foreign reviewers or broadly composed panels without true peers in the research area concerned. Researchers from the home country then lose the chance to distribute research resources within their own area. An example of the latter case is distributing more resources through channels that do not require peer review, for example by transferring public funds directly to universities, based on result indicators or administrative decisions, and also through commissioned research and tender competitions.

Tension 3:
Neutral judge versus promoting research interests

There are also tensions connected with double expectations concerning evaluators' neutrality. On the one hand, the evaluators are expected to be neutral judges who carry out an impartial and thorough scrutiny. On the other, the evaluators expect to be able to influence what is considered to be good research and how key resources in their area are distributed. These double expectations leave a negotiable space for the requirements and significance of a 'neutral assessor'. Such a person obviously cannot pursue personal interests, but what about conflicting views based on the evaluators' research fields, research interests or 'schools of thought'? In some respects, the notion of an academically neutral evaluator is mean-ingless. Researchers have different academic backgrounds and divergent opinions about, for example, method and theory, and the resulting differences in assessments may reasonably be regarded as legitimate differences, not as bias.[20] In other words, evaluators are not expected to be academically neutral; instead, they are supposed to make assessments based on scholarly discretion. The tensions that arise originate from the lack of any clear boundaries for scholarly discretion, i.e. no clear boundary between what evaluators may legitimately attempt, and what they should refrain from attempting, to influence. For example, it may be difficult to distinguish promotion of academically appropriate methods and research perspectives from assessments that support the evaluator's own research interests.

Tensions between expectations of a neutral assessor and of being able to influence the outcome may be particularly disturbing when there are no clear norms for selecting evaluators and when the processes lack transparency. Some of the literature on examination of doctoral theses is about this very lack of national standards for PhD examinations[21], and gives examples of procedures that make distortions and other disruptions possible. For example, lack of standards and openness about how examiners are selected may entail detrimental power relations between supervisors and examiners.[22]

Tension 4:
Deviant assessments versus unanimous conclusions

Different scholars have divergent assessments and priorities. At the same time, processes of peer review are intended to bring about unanimous conclusions on distribution of scarce resources and prestige.

There are two main ways of handling these tensions. One is for the expert reviewers to meet in person and reach a compromise through discussion. The other is for non-experts to take decisions based on a number of individual peer assessments. Journal editors' decisions on whether to publish papers submitted are an example of the latter. An example of the first method is decision-making in grant review panels. Studies show that an important mechanism for attaining unity in such groups is to respect the assessment of the panel member with the most established proof of competence on the application in question.[23] The way in which the process is organised can also affect the outcome.[24]

In the assessment of manuscripts and grant applications, and in most other peer review allocating scarce resources in the scientific community, deviant assessments are unproblematic and even considered an important part of the dynamics of research. On the other hand, in the assessment of research for use in public policy divergent assessments often cause problems. Stakes are high and review processes that leave scope for bias or random results are not tolerated. The potential for bias is more pronounced in reviews relating to public regulations and policy than in ordinary academic peer review, since the reviewers' research perspectives can influence how risks are assessed.[25]

Tension 5:
Peer review versus quantitative indicators

Alongside the evaluation spiral described above and the general rise in the use of peer review, the use of quantitative indicators is also increasing. Quantitative indicators are compelling since they offer simple answers at a low cost.[26] Even when quantitative indicators are aggregate conclusions from previous peer assessments (i.e. bibliometric measures), there are tensions between these indicators on the one hand and peer review on the other, and they may yield different conclusions. Studies comparing the outcomes of peer review and bibliometrics have found some correla-

tion, but in many cases they are much less closely correlated than one would expect.[27] Although both methods are ultimately based on peer review, directly or indirectly, they rest on very different logics. Peer review is based on subtle and tacit judgements and detailed knowledge of the research frontier in the area under assessment[28], whereas the bibliometric method counts articles and citations in indexed journals. Bibliometrics can provide important information about a researcher's previous activities, networks and collaboration patterns, but are disputed as evidence for academic quality. In some cases, such as evaluations of major research programmes, peer review and bibliometrics are combined, which may save time. In other cases, peer review is outcompeted by quantitative measures, e.g. when funding authorities are looking for ways to increase accountability and efficiency, and choose to replace fixed block grants to universities by funding based on quantitative performance indicators[29], rather than by research grants based on peer review.

Tension 6:
Autonomy versus responsibility

Although peer review is a key control mechanism in the academic community, it may work against responsibility in relation to society at large. Peer review serves as a mechanism of 'professional self-regulation that affords scientists a degree of autonomy from scrutiny by the public'.[30] Academic quality is one of the least politicised aspects of research, and the academic community commands full autonomy in defining and assessing its research. At times when the autonomy of academia is contested[31] and there is little overlap between public and academic research interests, society's demands to determine the research agenda will probably strengthen the researchers' efforts to fortify academic autonomy. Researchers invoke peer review in their own defence[32], one example being the evaluations of research programmes in which the reviewers recommend an increased share of resources to fundamental researcher-initiated research.[33]

The academic community's ability to protect its own independence is generally strong. The need for peer experts to assess academic quality is obvious, and researchers often succeed in defining 'quality' as 'scientific quality'. The scope for non-peer participation in the evaluation of research is clearly limited. However, there are salient differences among

different evaluator roles. While the task of assessing doctoral theses cannot be performed by non-peers, research grants may be distributed through, for example, direct awards based on previous productivity.

In addition, non-peer assessments of applications may add tensions relating to academic autonomy. Decisions on research grants are often based on additional aspects alongside academic quality, such as whether a project fits into the desired allocation of funds in terms of theme, geographical and institutional location, gender balance, potential societal impacts and ethical considerations.[34] In general, peer competence is not required for assessing such issues. Nor is it obvious who can best assess the effects of the proposed research on society. Researchers may assert that the societal merits of a project are more uncertain than the academic ones, and that effects on society depend on scientific success; accordingly, they may prioritise scientific quality over benefit to society. Moreover, if a research-funding body were to organise its own assessment process in such a way that non-experts could outvote experts, it may be perceived as illegitimate and deliberately opposed by the reviewers. Even when peer review is defined as a part of the system ensuring accountability for public expenditures, public accountability and ensuring that public funding for research serves public needs are not the priorities of the peer reviewers.[35]

The role of the research evaluator is undergoing change. In the past two decades, the work has become increasingly international, and this helps both to improve quality and to prevent conflicts of interest. Open peer review is another trend possibly contributing to more egalitarian review processes. Certain journals have introduced more open peer review processes, including open invitations to comment on manuscripts submitted.[36] Below, we discuss an agenda for future studies of the research evaluator role in the light of increased globalisation and the tensions analysed above.

Choice of evaluators, shortage of time and academic stratification

The outcome of peer review depends on who carries it out. Nonetheless, there is very little research on how evaluators are selected, or on the availability of appropriate expertise and their time constraints. Here, there are important research questions to explore. Who are the evaluators most

in demand and why? Which are the most attractive evaluation tasks and how is time spent on the various evaluation tasks today?

Selection processes and time constraints may affect work distribution and stratification in the academic community, including the global distribution of work, in various ways. There are limits to how much time researchers are willing to spend on evaluations. When the total volume of evaluations increases, the most senior and most frequently consulted evaluators will most likely have to prioritise the evaluation tasks they take on more strictly, and allow more junior scholars and/or researchers, from a broader range of countries, to be involved in evaluation. Introducing open review (self-selection of reviewers) may reinforce this tendency. Nevertheless, there is reason to assume that the top-ranking researchers can retain the assignments of greatest political importance. Relevant research questions here include how much time researchers are willing to devote to evaluation work and how they choose evaluation assignments. For example, when academic autonomy is called into question, researchers may be more willing to spend time on evaluation processes to ensure that scholarly quality is not overrun by other considerations.

Autonomy, indicators and non-peers

Another key issue is the threat to academic autonomy posed by the use of quantitative indicators and evaluations that do not require peer competence. A low approval rate in research funding agencies and concerns about bias and conflicts of interest in peer review may prompt measures to restrict the time researchers spend on carrying out peer review and writing applications, causing more emphasis to be laid on quantitative indicators. Moreover, more concern about societal relevance and the accountability of research may result in evaluations performed by non-peers.

Extending the use of quantitative measures in the allocation of research funds may reduce the importance of peer review. On the other hand, the aggregate importance of peer review is enhanced by peer review of funding based on previous performance. When funding is performance-based, important key figures are the number of peer-reviewed publications and the rate of success in obtaining research grants based on peer review. Moreover, in several countries peer review is the basis for performance-based funding; one example is the UK Research Excellence Framework.[37]

Quantitative measures give rise to two sets of research questions regarding researchers as evaluators. First, how far are evaluators aware of the aggregate importance of their assessment work, how much does this affect their willingness to take on evaluation assignments, and to what extent does it influence conflict-of-interest considerations and use of foreign reviewers? Questions in the second set relate to researchers' opinions and wishes concerning indicators and expert review. Do researchers prefer research funding to be distributed on the basis of quantitative measures and incentive systems, or of peer review of grant proposals? Both approaches are frequently criticised by researchers. Peer review is criticised for being conservative, i.e. discriminating against interdisciplinary and groundbreaking research, and time-consuming. Since performance-based funding is based on aggregate outcome of peer review, there is a risk that it may reproduce the latter's weaknesses. Quantitative measures may, moreover, produce dysfunctional incentives, such as giving higher priority to quantity than to quality. What is perceived as the fairest, least conservative and most efficient way of distributing research grants is not evident.

Quantitative measures imply a competitive regime based on (presumably) predictable criteria, while the introduction of non-peers and stronger emphasis on benefit to society in the evaluation process represent a clearer challenge to academic autonomy. Some researchers also challenge traditional peer review and propose more openness and broader selection criteria.[38]

There are many arguments against full academic self-governance in allocating research resources, and arguments for placing more emphasis on external considerations and involving non-peers in setting research priorities. It is usually held that public funds for research should serve public needs or at least some general interests. To legitimate heavy public spending on research, openness to external criticism and responsiveness to the needs and concerns of the public are also required. We have also seen rising demands for citizen influence (democratisation of research), greater emphasis on societal impacts and accountability, as well as challenges associated with controversial issues characterised by high uncertainty.[39] Including societal criteria and lay people in assessments may better ensure that research serves public needs and the needs of disadvantaged groups better.[40] Non-academic competences may reduce insider bias[41] and better ensure that a broader set of criteria — e.g.

social, economic, environmental, or health concerns — are taken into account.

Since assessment of research applications is a zero-sum game that defines the research focus in a programme or research field, this reviewer role is likely to be perceived as important in terms of protecting academic autonomy. In-depth studies are required to find out how introducing non-peers and broader relevance criteria affects review work and responsiveness to public concerns. Moreover, both evaluators' and research-funding agencies' perceptions of academic autonomy should be studied.

The text is an updated and somewhat shortened version of L. Langfeldt & S. Kyvik, 'Researchers as evaluators: Tasks, tensions and politics', *Higher Education*, 62, 2011, pp. 199–212.

★

1. J. M. Ziman, *Prometheus Bound: Science in a Dynamic Steady State*, Cambridge: Cambridge University Press, 1994, p. 103.

2. S. Johnston, 'Examining the examiners: An analysis of examiners' reports on doctoral theses', *Studies in Higher Education*, 22(3), 1997, pp. 333–347; p. 340.

3. *Tenure Denied: Cases of Sex Discrimination in Academia*, Washington, DC: American Association of University Women Education Foundation & American Association of University Women Legal Advocacy Fund, 2004.

4. J. C. Hearn & M. S. Anderson, 'Conflict in academic departments: An analysis of disputes over faculty promotion and tenure', *Research in Higher Education*, 43(5), 2002, pp. 503–529.

5. D. F. Horrobin, 'Peer review: A philosophically faulty concept which is proving disastrous for science', *Behavioral and Brain Sciences*, 5(2), 1982, pp. 217–218; R. Roy, 'Alternatives to review by peers: A contribution to the theory of scientific choice', *Minerva*, 22(3–4), 1984, pp. 316–328.

6. G. D. L. Travis & H. M. Collins, 'New light on old boys: Cognitive and institutional particularism in the peer review system', *Science, Technology & Human Values*, 16(3), 1991, pp. 322–341.

7. M. Lamont, *How Professors Think: Inside the Curious World of Academic Judgment*, Cambridge, MA: Harvard University Press, 2009.

8. A. C. Weller, *Editorial Peer Review: Its Strengths and Weaknesses*, Medford, NJ: Information Today (ASIS&T Monograph Series), 2001.

9. S. J. Ceci & D. P. Peters, 'Peer review: A study of reliability', *Change*, 14(6), 1982, pp. 44–48; J. M. Campanario, 'Peer review for journals as it stands today — part 1', *Science Communication*, 19(3), 1998, pp. 181–211; J. M. Campanario, 'Peer

review for journals as it stands today — part 2', *Science Communication*, 19(4), 1998, pp. 277–306; B. W. Speck, *Publication Peer Review: An annotated bibliography*, Westport/London: Greenwood Press, 1993; Weller 2001; C. J. Lee, C. R. Sugimoto, G. Zhang & B. Cronin, 'Bias in peer review', *Journal of the American Society for Information Science and Technology*, 64(1), 2013, pp. 2–17; K. C. Soh, 'Peer review: Has it a future?', *European Journal of Higher Education*, 3(2), 2013, pp. 129–139.

10. R. M. Friedman, *The Politics of Excellence: Behind the Nobel Prize in Science*, New York: Times Books, 2001.

11. G. Neave, 'The evaluative state reconsidered', *European Journal of Education*, 33(3), 1998, pp. 265–284.

12. H. F. Hansen, *Forskningsevaluering: Metoder, praksis og erfaringer* ('Research Evaluation: Methods, Practice and Experience'), Research: Analysis and Evaluation, 1, Copenhagen: Danish Agency for Science, Technology and Innovation, 2009.

13. S. Jasanoff, *The Fifth Branch: Science Advisers as Policymakers*, Cambridge, MA: Harvard University Press, 1990, p. 79 et seq.

14. J. Irvine & B. Martin, *Foresight in Science: Picking the Winners*, London: Pinter, 1984.

15. S. Kyvik & L. Langfeldt, 'Tidkrevende bedømmelsesarbeid' ('Time-consuming evaluation work'), *Forskningspolitikk* ('Research Policy'), 27(2), 2004.

16. Ibid., p. 15.

17. S. Kyvik & J. Wiers-Jenssen, 'Internasjonalisering av norsk høyere utdanning' ('Internationalisation of Norwegian higher education'), *Arbeidsnotat* 2/2014, Oslo: NIFU, 2014.

18. S. Cole, 'The hierarchy of the sciences?', *American Journal of Sociology*, 89(1), 1983, pp. 111–139; p. 138.

19. D. E. Chubin & E. J. Hackett, *Peerless Science: Peer Review and US Science Policy*, New York: State University of New York Press, 1990, p. 80.

20. L. Langfeldt, *Decision-making in Expert Panels Evaluating Research: Constraints, processes and bias*, Oslo: Unipub, 2002, available at www.sv.uio.no/isv/forskning/aktuelt/arrangementer/disputaser/arkiv/fulltxt/langfeldt.pdf (accessed on 16 January 2015), pp. 67–69.

21. P. Tinkler & C. Jackson, 'Examining the doctorate: Institutional policy and the PhD examination', *Studies in Higher Education*, 25(2), 2000, pp. 167–180; L. Morley, D. Leonard & M. David, 'Variations in vivas: Quality and equality in British PhD assessments', *Studies in Higher Education*, 27(3), 2002, pp. 263–273.

22. Morley et al. 2002, pp. 270–271.

23. Lamont 2009, p. 117.

24. L. Langfeldt, 'The decision-making constraints and processes of grant peer review, and their effects on the review outcome', *Social Studies of Science*, 31(6), 2001, pp. 820–841.

25. Jasanoff 1990, pp. 76–83.

26. Ziman 1994, pp. 103–105.

27. D. W. Aksnes & R. E. Taxt, 'Peer review and bibliometric indicators: A comparative study at a Norwegian university', *Research Evaluation*, 13(1), 2004, pp. 33–41; L. Bornman & H. D. Daniel, 'Reliability, fairness, and predictive validity of the peer review process for the selection of research fellowship recipients of the Boehringer Ingelheim Fonds', in B. M. Kehm (ed.), *Hochschule im Wandel: Die Universität als Forschungsgegenstand. Festschrift für Ulrich Teichler*, Frankfurt am Main: Campus, 2008, pp. 365–376; P. van den Besselaar & L. Leydesdorf, 'Past performance, peer review and project selection: A case study in the social and behavioral sciences', *Research Evaluation*, 18(4), 2009, pp. 272–288.

28. J. R. Ravetz, *Scientific Knowledge and its Social Problems*, Oxford: Clarendon Press, 1971, p. 274.

29. S. Sörlin, 'Funding diversity: Performance-based funding regimes as drivers of differentiation in higher education systems', *Higher Education Policy*, 20(4), 2007, pp. 413–440.

30. E. J. Hackett, 'Peer review in science and science policy', in M. S. Frankel & J. Cave (ed.), *Evaluating Science and Scientists*, Budapest: Central European University Press, 1997, pp. 51–60; p. 57.

31. M. Henkel, 'Can academic autonomy survive in the knowledge society? A perspective from Britain', *Higher Education Research & Development*, 26(1), 2007, pp. 87–99.

32. Chubin & Hackett 1990, p. 5.

33. Langfeldt 2002.

34. Langfeldt 2001, pp. 827–829.

35. B. Van der Meulen, 'Science policies as principal-agent games: Institutionalization and path dependency in the relation between government and science', *Research Policy*, 27(4), 1998, pp. 397–414; p. 405 ff.

36. U. Pöschl & T. Koop, 'Interactive open access publishing and collaborative peer review for improved scientific communication and quality assurance', *Information Services & Use*, 28(2), 2008, pp. 105–107.

37. Previously known as the Research Assessment Exercise; see the Research Excellence Framework, www.ref.ac.uk (accessed on 16 January 2015).

38. A. Scott, 'Peer review and the relevance of science', *Futures*, 39(7), 2007, pp. 827–845.

39. Ibid.; S. Jasanoff, 'Technologies of humility: Citizen participation in governing science', *Minerva*, 41(3), 2003, pp. 223–244.

40. E. Woodhouse & D. Sarewitz, 'Science policies for reducing societal inequities', *Science and Public Policy*, 34(3), 2007, pp. 139–150.

41. B. Martin, 'Research grants: Problems and options', *Australian Universities' Review*, 43(2), 2000, pp. 17–22, available at www.bmartin.cc/pubs/00aur.html (accessed on 16 January 2015).