ENRESSH Work Group 1

Overview of Peer Review Practices

in the SSH

ENRESSH Report

Michael Ochsner, ETH Zurich and FORS, Lausanne, Switzerland
Nina Kancewicz-Hoffman, Institute of Literary Research, Polish Academy of Sciences, Poland
Marek Hołowiecki, Adam Mickiewicz University Poznań, Poland
Jon Holm, National Research Council, Norway

With chapters by Mirjam Aeschbach, Gemma Derrick, Tim C. E. Engels, Elea Giménez-Toledo, Raf Guns, Jon Holm, Marek Hołowiecki, Sven E. Hug, Nina Kancewicz-Hoffman, Karolina Lendák-Kabók, Lai Ma, Jorge Mañana-Rodríguez, Michael Ochsner, Ginevra Peruginelli, Janne Pölönen, Elías Sanz-Casado, Tony Ross-Hellauer, Jadranka Stojanovski, Marc Vanholsbeeck

Chapters reviewed by Elea Giménez-Toledo, Aldis Gedutis, Jon Holm, Karolina Lendák-Kabók, Michael Ochsner, Ginevra Peruginelli, Janne Pölönen, Linda Sīle, Yulia Stukalina, Mimi Urbanc

WG to which the report/deliverable is related: WG 1

Grant period to which the report/deliverable is related: Grant Period 4

Version 2.1, April 2020

Executive summary

Peer review is an important method of research evaluation, and it seems that the only adequate way to evaluate SSH research involves some form of peer review. Even if bibliometrics and other quantitative ways of evaluation may provide information on some aspects of SSH research like productivity and publication strategies of research units, metrics-based indicators should be used with caution in SSH due to low coverage of SSH fields in the standard publication databases and a mismatch between dimensions of quality as defined by peers and standard bibliometric indicators. Still, peer review faces many issues and challenges. This report identifies the challenges particularly relevant for the SSH, such as different and thus often conflicting research paradigms or epistemological styles of reviewers and applicants or authors; difficulty in many SSH disciplines to define and evaluate research methodology compared to STEM disciplines; the lack of the idea of linear progress and a much longer time span necessary to evaluate academic impact of publications; the diversity of publication outputs and specific importance of books or monographs; the importance of local languages; challenges related to recent developments in research and its evaluation related to growing interdisciplinarity and the Open Science agenda. To this, the general challenges of peer review are added, such as the risk of gender bias, conservative bias, workload for all parties involved.

The report concludes that peer review fulfils different functions and that peer review practices not only need to acknowledge different disciplinary particularities but also their evaluative context. Rather than playing metrics and peer review off against each other, the focus should be on their optimal use and combination within different evaluation situations. This is especially important when it concerns the SSH because the disciplines falling under this umbrella term share the concurrency of different paradigms and a context-dependent, sometimes interpretative mode of knowledge generation and the use of a wide range of dissemination channels. This leads to a particular challenge regarding the burden of reviewers because SSH disciplines often act in a local context in national languages and include small disciplinary communities.

The SSH disciplines should develop their own ways to adequately evaluate their research, and peer review takes an important part in that. The past has shown that automatically copying evaluation procedures from STEM disciplines did not always work out well. However, the SSH community is well resourced to analyse and remediate the current tensions in research policies between funders' expectations of societal impact and the value of academic autonomy, between the ambition of mainstreaming of SSH research and the care for specific SSH methods and practices, and not least the threatened legitimacy of science in the post-factual society. The task of the SSH community should not only be to defend the integrity of scholarly disciplines, but to contribute to the development of new practices of research assessments that may build bridges between different communities of researchers and between the world of research and society at large.

Keywords

Peer Review, Evaluation, Criteria, Societal Impact, Books, Funding, Open Access

Acknowledgements

This report is a result of work from COST Action CA 15137 "European Network for Research Evaluation in the SSH (ENRESSH)", supported by COST (European Cooperation in Science and Technology).
The editors would like to thank all participants from Work Group 1 "Conceptual Frameworks of Research Evaluation in the SSH" who participated in one of the sessions regarding this report (alphabetical order):

Georgios Afxentiou, Judit Bar-Ilan, Ondřej Daniel, Katya De Giovanni, Ioana Galleron, Aldis Gedutis, Haris Gekić, Elea Giménez-Toledo, Agnė Girkontaitė, Raf Guns, Jon Holm, Marek Hołowiecki, Todor Hristov Dechev, Dragan Ivanović, Nina Kancewicz-Hoffman, Andreja Isenič Starčič, Arnis Kokorevics, Litauras Kraniauskas, Emanuel Kulczycki, Karolina Lendák-Kabók, Lai Ma, Jorge Mañana-Rodríguez, Ketrina Mijo, Michael Ochsner, Elena Papanastasiou, Sanja Peković, Ginevra Peruginelli, Janne Pölönen, Hulda Proppé, Tony Ross-Hellauer, Elías Sanz-Casado, Ana Ramos, Linda Sīle, Karel Šima, Dagmar Simon, Jolanta Šinkūnienė, Gunnar Sivertsen, Jack Spaapen, Jadranka Stojanovski, Mimi Urbanc, Marc Vanholsbeeck, Maja Vehovec, Albena Vutsova, Alesia Zuccala

Recommended citation of the whole report:
Ochsner, M., Kancewicz-Hoffman, N., Hołowiecki, M., & Holm, J. (Eds.). (2020). *Overview of peer review practices in the SSH. ENRESSH Report.* European Network of Research Evaluation in the Social Sciences and Humanities. https://dx.doi.org/10.6084/m9.figshare.12032589

Recommended citation of a chapter in the report:
Derrick, G., & Ross-Hellauer, T. (2020). Peer Review in SSH: In Need of Development? In M. Ochsner, N. Kancewicz-Hoffman, M. Hołowiecki, J. Holm (Eds.), *Overview of peer review practices in the SSH. ENRESSH Report* (pp. 10–14). European Network of Research Evaluation in the Social Sciences and Humanities. https://dx.doi.org/10.6084/m9.figshare.12032589

Table of Content

Practices of peer review in the SSH I: A systematic review of peer review criteria

By Sven E. Hug, Marek Hołowiecki, Lai Ma, Mirjam Aeschbach & Michael Ochsner

Introduction

Criteria are an essential component of any procedure for judging merit. This is widely acknowledged in the literature on peer review. Yet pertinent literature reviews and compendia do not mention or only briefly discuss peer review criteria. To address this research gap, a systematic review of studies on peer review criteria has been conducted. The review focused on the most fundamental question in any evaluation: what criteria are employed in the evaluation? The systematic review was restricted to the two most common forms of peer review: the assessment of manuscripts submitted to journals and the assessment of grant applications. The objectives of the review were (a) to identify studies that develop or derive criteria inductively,[8] (b) to determine how many of these studies focus on the social sciences and humanities, and (c) to provide a taxonomy of criteria. In the following, preliminary findings on objectives (a) and (b) will be reported. Methodological details and final results will be published in scholarly journals (Hug & Aeschbach, 2019; Hug et al., in prepration).

Preliminary findings

Twelve studies on grant review criteria and twice as many on manuscript review criteria were identified (see Table 1). While the first inductive study on manuscript criteria (i.e. Bonjean & Hullum, 1978) dates back to the time when modern peer review emerged (see Baldwin, 2017, 2018; Moxham & Fyfe, 2018), the first study on funding criteria was only carried out in the 1990s (i.e. Hartmann, 1990). Most studies have examined criteria in the medical and health sciences and the social sciences. Studies on other fields are scarce and there are no studies on manuscript criteria in the natural sciences and in engineering and technology. A possible explanation for the latter could be the fact that all studies on manuscript criteria were done by "insiders" (i.e. researchers examined the criteria employed in a journal of their own field). Since qualitative-inductive approaches are not in the (standard) repertoire of researchers in the natural sciences and in engineering and technology, it is unlikely that criteria are inductively studied in these fields. The systematic review showed that manuscript criteria are mainly examined with data from actual reviews and comments. In contrast, data collection methods such as interviews, surveys and the Delphi method are as important as actual reviews in studies on grant criteria.

[8] While an inductive approach generates criteria from empirical data, a deductive approach employs theoretically determined or otherwise predefined criteria. The very first studies on peer review criteria employed a deductive approach (e.g. Chase, 1970; Frantz, 1968). The systematic review, however, did not focus on theoretically derived or otherwise predefined criteria but on inductively and empirically established criteria, which were, for example, based on quality conceptions of scholars or on actual comments of reviewers.

Table 1. Characteristics of studies that develop or derive peer review criteria inductively (preliminary data).

	Studies on manuscript review criteria	Studies on grant review criteria
Total number of studies included in the review	24	12
Publication year of studies		
First study	1978	1990
Latest study	2018	2018
Median	2004	2007
Number of studies analysing criteria in the		
Natural sciences	–	2
Engineering and technology	–	2
Medical and health sciences	8	8
Agricultural sciences	1	–
Social sciences	14	4
Humanities	2	3
Data collection		
Interview, survey, Delphi method, etc.	5	7
Actual reviews and comments	19	5
Number of criteria per study		
Minimum	8	7
Maximum	223	66
Mean	44	26
Median	19.5	21

Studies on manuscript criteria on average report more criteria than studies on grant criteria (44 and 26, respectively). In particular, while the study that reports the most grant criteria (Pollitt et al., 1996) lists 66 criteria, there are six studies on manuscripts that list more criteria. For example, Campion (1993) lists no less than 223 criteria for reviewing research articles in applied psychology. A possible reason for this difference could be the strong improvement focus of the manuscript review process, which could promote more detailed comments of reviewers or prompt authors of studies on manuscript criteria to perform more fine-grained analyses. If, however, one ignores the studies that report a large number of criteria (i.e. those larger than the median), a similar pattern emerges: 50% of the manuscript and grant studies report 8 to 19 and 7 to 21 criteria, respectively.

Preliminary conclusions

Although there are tens of thousands of publications on peer review (see Batagelj et al., 2017) and although criteria are an essential component of any evaluation process, there are only very few studies that focus on criteria peers actually use or prefer. In particular, 24 inductive studies on manuscript review criteria and 12 inductive studies on grant review criteria were identified in the systematic review. With respect to research fields, the systematic review showed that most studies analysed criteria in the medical and health sciences and in the social sciences. These findings suggest that there is a need for more studies on peer review criteria in general and more studies on the natural sciences and humanities in particular. In addition, future studies should develop a comparative perspective to improve the understanding of the commonalities and peculiarities of the evaluation cultures of different fields and disciplines. From a practical standpoint, studies on peer review

criteria are relevant as they contribute to increasing the transparency of peer review processes and they support early career researchers in learning the basics of peer assessment.

References

References with one asterisk (*) indicate studies on manuscript peer review, references with two asterisks (**) indicate studies on grant peer review.

**Abdoul, H., Perrey, C., Amiel, P., Tubach, F., Gottot, S., Durand-Zaleski, I., & Alberti, C. (2012). Peer review of grant applications: criteria used and qualitative study of reviewer practices. *PLoS ONE, 7*(9), 1–15. https://doi.org/10.1371/journal.pone.0046054

*Albers, C. A., Floyd, R. G., Fuhrmann, M. J., & Martinez, R. S. (2011). Publication criteria and recommended areas of improvement within school psychology journals as reported by editors, journal board members, and manuscript authors. *Journal of School Psychology, 49*(6), 669–689. https://doi.org/10.1016/j.jsp.2011.10.002

*Bakanic, V., McPhail, C., & Simon, R. J. (1989). Mixed messages: referees' comments on the manuscripts they review. *The Sociological Quarterly, 30*(4), 639–654. https://doi.org/10.1111/j.1533-8525.1989.tb01540.x

Baldwin, M. (2017). In referees we trust? *Physics Today, 70*(2), 44–49. https://doi.org/10.1063/PT.3.3463

Baldwin, M. (2018). Scientific autonomy, public accountability, and the rise of "peer review" in the Cold War United States. *Isis, 109*(3), 538–558. https://doi.org/10.1086/700070

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

*Becker, B. (1991). The quality and credibility of research reviews: What the editors say. *Personality and Social Psychology Bulletin, 17*(3), 267-272. https://doi.org/10.1177/0146167291173006

*Bonjean, C. M., & Hullum, J. (1978). Reasons for journal rejection. An analysis of 600 manuscripts. *PS: Political Science & Politics, 11*(4), 480–483. https://doi.org/10.1017/S1049096500004844

*Bordage, G. (2001). Reasons reviewers reject and accept manuscripts: The strengths and weaknesses in medical education reports. *Academic medicine, 76*(9), 889–896. https://doi.org/10.1097/00001888-200109000-00010

*Campion, M. A. (1993). Article review checklist. A criterion checklist for reviewing research articles in applied-psychology. *Personnel Psychology, 46*(3), 705–706. https://doi.org/10.1111/j.1744-6570.1993.tb00896.x

Chase, J. M. (1970). Normative criteria for scientific publication. *American Sociologist, 5*(3), 262–265.

*Chong, C. (1994). Editorial: We aim to publish – not to reject. *Canadian Journal of Plant Science, 74*(4). https://doi.org/10.4141/cjps94-120

*Coniam, D. (2011). Systematising System: one reviewer's analysis of the review process. *System, 39*(4), 539–553. https://doi.org/10.1016/j.system.2011.10.018

*Daft, R. L. (1995). Why I recommended that your manuscript be rejected and what you can do about it. *Publishing in the Organizational Sciences*, 164–183. https://doi.org/10.4135/9781452240466.n14

*Day, F. C., Schriger, D. L., Todd, C., & Wears, R. L. (2002). The use of dedicated methodology and statistical reviewers for peer review: A content analysis of comments to authors made by methodology and regular reviewers. *Annals of Emergency Medicine, 40*(3), 329–333. https://doi.org/10.1067/mem.2002.127326

*Dickersin, K., Ssemanda, E., Mansell, C., & Rennie, D. (2007). What do the JAMA editors say when they discuss manuscripts that they are considering for publication? Developing a schema for classifying the content of editorial discussion. *BMC Medical Research Methodology, 7*, Article 44. https://doi.org/10.1186/1471-2288-7-44

*Ezeala, C., Nweke, I., & Ezeala, M. (2013). Common errors in manuscripts submitted to medical science journals. *Annals of medical and health sciences research, 3*(3), 376–379. https://doi.org/10.4103/2141-9248.117957

*Fiske, D. W., & Fogg, L. (1990). But the reviewers are making different criticisms of my paper!: diversity and uniqueness in reviewer comments. *American Psychologist, 45*(5), 591–598. https://doi.org/10.1037/0003-066X.45.5.591

Frantz, T. T. (1968). Criteria for publishable manuscripts. *Personnel and Guidance Journal*, *47*(4), 384–386. https://doi.org/10.1002/j.2164-4918.1968.tb02944.x

*Gogolin, I. (2011). *European Educational Research Quality Indicators (EERQI). Project final report*. http://eerqi.eu/sites/default/files/Final_Report.pdf

**Guetzkow, J., Lamont, M., & Mallard, G., (2004). What is originality in the humanities and the social sciences? *American Sociological Review, 69*(2), 190–212. https://doi.org/10.1177%2F000312240406900203

Hartmann, I. (1990). *Begutachtung in der Forschungsförderung: die Argumente der Gutachter in der Deutschen Forschungsgemeinschaft*. RG Fischer.

**Hartmann, I., & Neidhardt, F. (1990). Peer review at the Deutsche Forschungs-gemeinschaft. *Scientometrics, 19*(5–6), 419–425. https://doi.org/10.1007/BF02020704

*Hesterman, C. M., Szperka, C. L., & Turner, D. P. (2018). Reasons for manuscript rejection after peer review from the journal headache. *Headache: The Journal of Head and Face Pain*. https://doi.org/10.1111/head.13343

*Hewings, M. (2004). An 'important contribution'or 'tiresome reading'? A study of evaluation in peer reviews of journal article submissions. *Journal of Applied Linguistics, 1*(3), 247–274. https://doi.org/10.1558/jal.v1i3.247

Hug, S. E., & Aeschbach, M. (2019). Criteria for assessing grant applications. A systematic review. *Palgrave Communications, 6*, Article 37. https://doi.org/10.1057/s41599-020-0412-9

Hug, S. E., Holowiecki, M., Ma, L., Aeschbach, M., & Ochsner, M. (in prepration). *Criteria for reviewing journal manuscripts. A systematic review*.

**Lahtinen, E., Koskinen-Ollonqvist, P., Rouvinen-Wilenius, P., Tuominen, P., & Mittelmark, M. B. (2005). The development of quality criteria for research: A Finnish approach. *Health Promotion International, 20*(3), 306–315. https://doi.org/10.1093/heapro/dai008

**Lamont, M. (2009). *How professors think: inside the curious world of academic judgment.* Harvard University Press.

*Leung, D., Law, R., Kucukusta, D., & Guillet, B. D. (2014). How to review journal manuscripts: A lesson learnt from the world's excellent reviewers. *Tourism Management Perspectives, 10*, 46–56. https://doi.org/10.1016/j.tmp.2014.01.003

*McKercher, B., Law, R., Weber, K., Song, H., & Hsu, C. (2007). Why referees reject manuscripts. *Journal of Hospitality and Tourism Research, 31*(4), 455–470. https://doi.org/10.1177/1096348007302355

Moxham, N., & Fyfe, A. (2018). The Royal Society and the prehistory of peer reveiw, 1665-1965. *Historical Journal, 61*(4), 863–889. https://doi.org/10.1017/S0018246X17000334

**Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., Ford, C. E., & Carnes, M. (2018). Low agreement among reviewers evaluating the same NIH grant applications. *Proceedings of the National Academy of Sciences of the United States of America, 115*(12), 2952–2957. https://doi.org/10.1073/pnas.1714379115

**Pollitt, F. A., Notgrass, C. M., & Windle, C. (1996). Mental health services research. Peer review of rural research grant applications. *Administration and Policy in Mental Health and Mental Health Services Research*, *24*(2). 173–180. https://doi.org/10.1007/BF02042489

*Ramulu, V. G., Levine, R. B., Hebert, R. S., & Wright, S. M. (2005). Development of a case report review instrument. *International Journal of Clinical Practice, 59*(4), 457–461. https://doi.org/10.1111/j.1368-5031.2005.00319.x

**Reinhart, M. (2010). Peer review practices: A content analysis of external reviews in science funding. *Research Evaluation, 19*(5), 317–431. https://doi.org/10.3152/095820210X12809191250843

**Schmitt, J., Petzold, T., Nellessen-Martens, G., & Pfaff, H. (2015). Priorisierung und Konsentierung von Begutachtungs-, Förder- und Evaluationskriterien für Projekte aus dem Innovationsfonds: Eine multiperspektivische Delphi-Studie. [Prioritization and consentation of criteria for the appraisal, funding and evaluation of projects from the German Innovation Fund: a multi-perspective Delphi study]. *Gesundheitswesen, 77*(8–9), 570–579. https://doi.org/10.1055/s-0035-1555898

*Smith, M. U., Wandersee, J. H., & Cummins, C. L. (1993). What's wrong with this manuscript? An analysis of the reasons for rejection given by Journal of Research in Science Teaching reviewers. *Journal of Research in Science Teaching, 30*(2), 209–211. https://doi.org/10.1002/tea.3660300207

**Thomas, J. P., & Lawrence, T. S. (1991). Common deficiencies of NIDRR research applications. *American Journal of Physical Medicine & Rehabilitation, 70*(1), 161–164. https://doi.org/10.1097/00002060-199004000-00005

*Turcotte, C., Drolet, P., & Girard, M. (2004). Study design, originality and overall consistency influence acceptance or rejection of manuscripts submitted to the Journal. *Canadian Journal of Anaesthesia-Journal Canadien d'anesthésie, 51*(6), 549–556. https://doi.org/10.1007/BF03018396

**van Arensbergen, P., van der Weijden, I., & van den Besselaar, P. (2014). Different views on scholarly talent: What are the talents we are looking for in science? *Research Evaluation, 23*(4), 273–284. https://doi.org/10.1093/reseval/rvu015

*van Lent, M., IntHout, J., & Out, H. J. (2015). Peer review comments on drug trials submitted to medical journals differ depending on sponsorship, results and acceptance: A retrospective cohort study. *BMJ Open, 5*(9), e007961. https://doi.org/10.1136/bmjopen-2015-007961

**Whaley, A. L., Rodriguez, R., & Alexander, L. A. (2006). Development of a rating form to evaluate grant applications to the Hogg Foundation for Mental Health. *Evaluation Review, 30*(1), 3–26. https://doi.org/10.1177/0193841X05275586

*Yarbrough, C., & Whitaker, J. A. (2009). Analysis of reviewer comments about quantitative manuscripts accepted by the Journal of Research in Music Education. *Journal of Research in Music Education, 56*(4), 287–292. https://doi.org/10.1177/0022429408329448

*Zaruba, K. E., Toma, J. D., & Stark, J. S. (1996). Criteria used for qualitative research in the refereeing process. *Review of Higher Education, 19*(4), 435–460. https://doi.org/10.1353/rhe.1996.0017