New approaches to the research of the peer review system

Henk F. Moed

PEERE Workshop on “Peer Review: Old challenges and new advances”,
Vilnius, Lithuania, 7-9 March 2017

Short CV Henk F. Moed

1981-2009	Staff member at Centre for Science and Technology Studies (CWTS), Leiden Univ.
Feb 2010 – Aug 2012	Elsevier, SciVal Dept. Senior Scientific Advisor
Sept 2012 – Oct 2014	Elsevier, AGRM Dept; Head of Informetric Research Group
As from Sept 2014	Visiting professor at Sapienza Univ Rome (2014) and Univ Granada (2017); Private consultant;

Objects of peer review

Object of peer review	Assessment criteria
Scholarly manuscripts submitted to journal, book publishers	Significance; originality; methodological soundness; fitting to scope; quality of exposure;
Research proposals submitted to funding organizations	Contribution to scientific progress; originality; viability; methodological soundness; adequacy of requested resources
Research entities (groups, programs, organisations, etc.)	Research performance: quality; impact; productivity; relevance; viability;

Contents

Part	Subject
1	Assumptions, challenges, solutions
2	Minimal assessment criteria
3	Indicators of the manuscript peer review process
4	Experiments
5	Conclusions

Contents

Part	Subject
1	**Assumptions, challenges, solutions**
2	Minimal assessment criteria
3	Indicators of the manuscript peer review process
4	Experiments
5	Conclusions

Basic assumptions on the nature of peer review

1. Manuscript peer review aims to facilitate the scientific debate, not to settle it.

2. The combined formative and summative functions of peer review aim to set and apply minimum quality criteria

3. A journal serves a research community as it provides a pre-selection of potentially relevant papers meeting minimum quality criteria

4. Even if peer review provides invalid outcomes in individual cases, it may be beneficial to the scholarly system as a whole.

1. Manuscript peer review aims to facilitate the scientific debate, not to settle it

- Some communications between referee and author should take place after publication of the paper, not before
- Reviewers should maintain a certain distance towards their views and preferences
- A reviewer may disagree with a statement yet consider it defensible

Quote from a peer review report - 1

“Although I keep my doubts about the reliability of method M, which I find extremely high, I find the current version interesting and informative, and balanced in its discussion”.

2. The combined formative and summative functions of peer review aim to set and apply minimum quality criteria

- Formative: focus on development, improvement, evaluated entity
- Summative: focus on outcome, test against a norm, evaluating entity
- Empirical foundation of the notion of minimum criteria is needed
- Key studies on peer review process of manuscripts and grant proposals: Cole & Cole; Chichetti
- Additional evidence from a peer review process of research groups

3. A journal serves a research community as it provides a pre-selection of potentially relevant papers meeting minimal quality criteria

- Journals as research communities (e.g., Belver Griffith)
- Communities share an intellectual focus and a common set of methods and quality criteria
- A journal's core identity is its peer review process
- Pre-selection (filtering) is informetrically useful for – and appreciated by – a research community

4. Even if peer review provides invalid outcomes in individual cases, it may be beneficial to the scholarly system as a whole

- Systemic approach vs. individual case is a key issue in applied informetrics
- The grounding of a norm for assessing whether a risk of invalid outcomes is acceptable, is a social-political, not an informetric issue
- Sometimes anecdotal evidence seems to discredit the peer review system

Challenges

- Increase the visibility of the efforts of reviewers without violating anonymity in the process
- Make peer review reports more instrumental to improving the quality of a manuscript
- Improve standardization and transparency of the peer review process
- Develop and apply indicators of the quality of the manuscript review process and its effects

Towards solutions

- It is essential to differentiate in peer review processes between experienced (‘senior’) and less experienced (‘junior’) researchers
- Further develop teaching courses on research assessment techniques incl. peer review in academic doctoral programs
- Further explication of notions as genre, document type, type of study
- Set up a pilot project aimed to develop indicators of the manuscript peer review process

Contents

Part	Subject
1	Assumptions, challenges, solutions
2	**Minimal assessment criteria**
3	Indicators of the manuscript peer review process
4	Experiments
5	Conclusions

Chance and consensus in peer review (Cole, Cole & Simon, 1981)

- Comparing peer ratings of NSF grant proposals from 3 fields to ratings given in a second review by independently selected panels of reviewers:
- "… The fate of a particular grant application is roughly half determined by the characteristics of the proposal and the principal investigator, and about half by apparently random elements which might be characterised as the 'luck of the reviewer draw'" (p. 885).

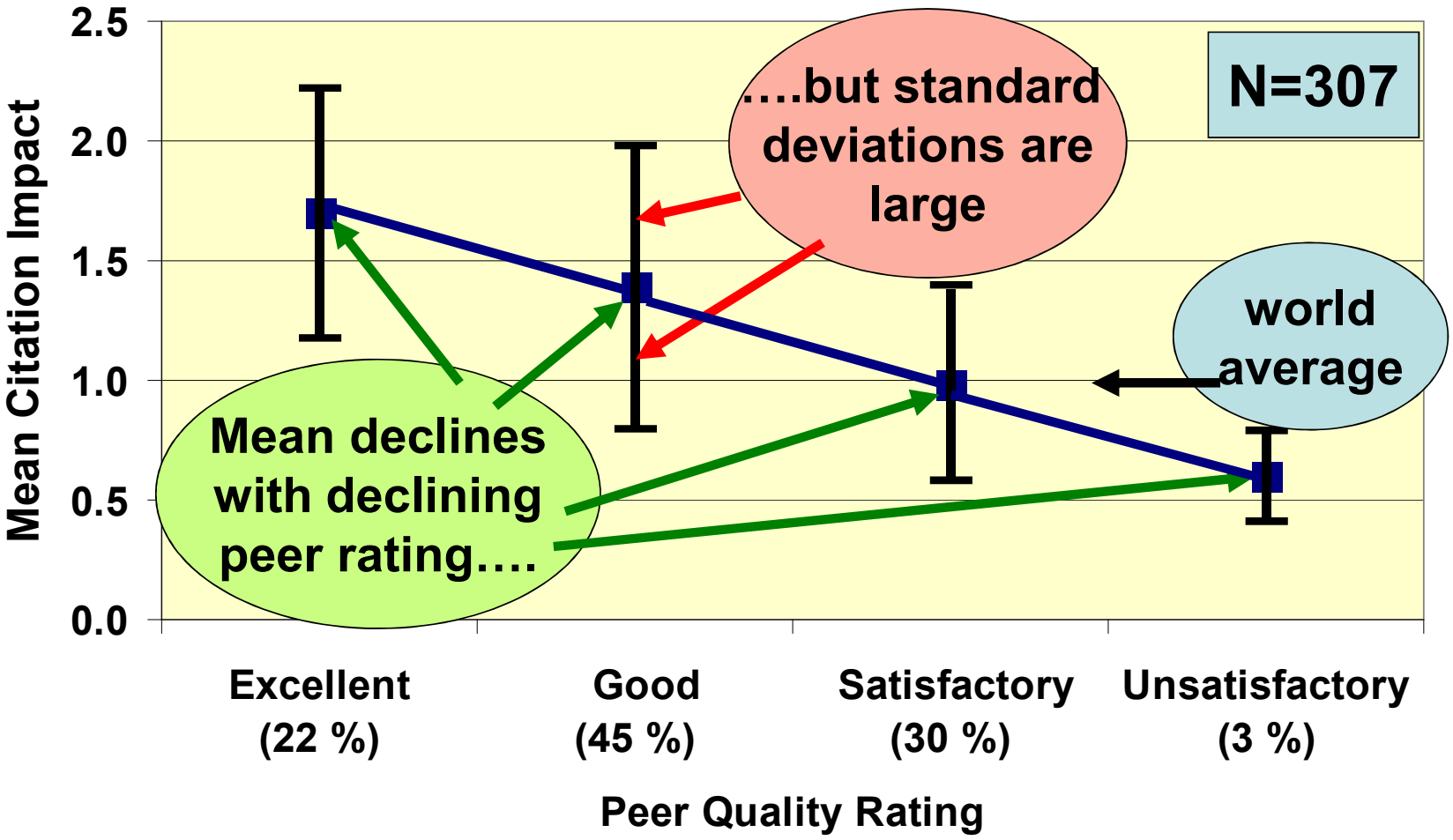
Cicchetti (1991) conclusions on agreement among referees

Evaluation Object	Subject field	Agreement among reviewers
Journal manuscripts	General and diffuse disciplines	Reviewers tend to agree more on rejection than on acceptance of manuscript
	Specialized and focused disciplines	Reviewers agree more on acceptance than on rejection
Grant proposals	Science (NSF proposals)	Reviewers agree much more about what is unworthy of support than about what does have scientific value

Case Study: Quality assessment of research in NL

- Organized by Association of Universities in the Netherlands (VSNU)
- International Committees of 7-10 independent experts in the field
- Biology (1995); Chemistry (1996);
  Physics (1996)

Correlation between peer ratings and citation impact is positive but far from perfect



VSNU: Cross table Peer Quality rating (5=excellent) vs Citation Impact class (5=highest)

Peer Quality rating	Citation Impact Class				Total
	2	3	4	5	
2	3	7	0	0	10
3	6	51	31	3	91
4	1	27	80	29	137
5	0	6	26	36	68
Total	10	91	137	68	306

VSNU: Cross table Peer Quality rating (5=excellent) vs Citation Impact class (5=highest)

Peer Quality rating	Citation Impact Class				Total
	2	3	4	5	
2	3	7	0	0	10
3	6	51	31	3	91
4	1	27	80	29	137
5	0	6	26	36	68
Total	10	91	137	68	306

Correlations between peer ratings and citation impact (NL-VSNU)

- Citation impact discriminated very well between groups rated excellent or good and those receiving lower peer ratings
- But it did not discriminate well between good and excellent groups in the perception of the peers

Assuming that the applied citation impact indicator reflects excellence adequately

- Peer review committees were able to identify ‘good’ or ‘valuable’ research meeting minimum quality standards
- But they were not very successful in identifying genuinely excellent or top research
- Possibly, peers agree more upon what is qualitatively ‘less good’ than what is ‘excellent’ or ‘genuine top’ research
- This finding is consistent with Cicchetti’s conclusions on proposal and manuscript pee reiew

Contents

Part	Subject
1	Assumptions, challenges, solutions
2	Minimal assessment criteria
3	**Indicators of the manuscript peer review process**
4	Experiments
5	Conclusions

Towards indicators of the manuscript peer review process

- Journal paper in: Scholarly Metrics and Analytics, July 2016
- Discussion paper in: Proceedings OECD Blue Sky Conference, Ghent, 19-21 Sept 2016
- News item in: Research Europe, 12 Jan 2017

Introduction-1

- Any academic journal is only as good as its peer review. And yet, the manuscript peer review process itself is still strikingly opaque
- Journal publishers and editors-in-chief rightly acknowledge that reviewers' independence must be preserved
- But reviewers tend to receive little training in what is one of the key academic activities
- There is little evidence of any standardisation in how review reports are composed.
- There is little systematic information on the quality of the process across journals and subjects, and on its effect upon the quality of submitted papers.

Introduction-2

- Journal quality is important to authors' decision on where to publish their articles—and from this, to decisions by libraries, research evaluators, appointment panels, etc.
- With peer review largely a black box, proxies for its quality have grown up, most notably the journal impact factor (JIF) based on citation counts.
- Its shortcomings are no secret, but its virtues—which include being highly visible, easily accessible, and relatively simple to understand—and the lack of alternatives, maintain its prominence

Introduction-3

- The digitisation of scientific information offers great potential for the development of tools to allow peer review to be analysed directly.
- Computational linguistic analysis and text-mining, combined with more traditional techniques from the humanities, offer the prospect for a better understanding of the process
- Such tools may offer more insight into differences among disciplines
- The ultimate aim of the project, though, should reach beyond comparing journals, to demonstrate the added value of the manuscript referee process, and to further enhance its transparency and efficiency.

Why are indicators of the manuscript peer review process useful?

- They provide more direct indicators of journal quality
- They further enhance the transparency of the review process
- They can help educating and training reviewers
- They can help assessing the effect of peer review upon manuscript quality.
- They can be used to monitor and further improve the review process

Two phases

	Phase	Brief description
1	Exploration phase "Classical-humanities approach"	❑ Development of a conceptual model ❑ Construction of referee report profiles ❑ Communication modes between actors ❑ Based on well-selected, small data samples
2	Data mining "Digital humanities approach"	❑ Use of computational linguistics tools ❑ Natural language processing ❑ Statistical analysis ❑ Data mining of large data samples

Quote from a peer review report - 2

“The paper lacks originality as this formula was developed by author A in paper P […]. Not surprisingly, the empirical findings in the paper, such as Finding F have been found before in several published papers.”

Quote from a peer review report - 3

“I have problems to identify the main purpose of your research […]. I would suggest to formulate explicit research questions and also explain why you use data from dataset D which do not reflect up-to-date field F very well.”

Quote from a peer review report - 4

“I have severe difficulties with this paper. For me an article should start with defining a hypothesis. Next empirical research is conducted examining its validity, and conclusions are drawn. This paper does not have this structure. It is more exploratory, it discusses a series of properties of database D without explaining their significance”

Tasks

- Take into account a journal's scope and instructions to reviewers
- Identify the different elements of a review report
- Categorize statements in terms of aspect, modality, etc.
- Identify standards or a-priories that reviewers apply
- Analyze how standards are expressed in the content of reviews
- Develop relevant concepts, e.g., 'formative content'
- Develop and validate hypotheses

Tentative hypothesis

Review reports that….

- fail to apply any assumed key standards
- or apply vague standards
- and that contain no reference to the manuscript's text

…. are ….

- less informative
- and of lower utility or 'quality'

…. than those adopting a series of clear assessment criteria backed up by citing the manuscript's text or tables

Objectives

- Build up an understanding, for each discipline, of what a reasonable quality threshold for publication looks like;
- how it differs among journals;
- what distinguishes an acceptable paper from one that is rejected.
- It would also be valuable to assess the degree of agreement among referees.

Contents

Part	Subject
1	Assumptions, challenges, solutions
2	Minimal assessment criteria
3	Indicators of the manuscript peer review process
4	**Experiments**
5	Conclusions

Experiments with peer review

- Differentiate between specialist journals and general, multi-disciplinary journals
- Differentiate between experienced ('senior') and less experienced ('junior') researchers
- For manuscripts in the lower part of the quality distribution focus on (double) blind, mostly formative review applying minimal criteria
- For the upper part: open, more summative review of current manuscript version

Contents

Part	Subject
1	Assumptions, challenges, solutions
2	Minimal assessment criteria
3	Indicators of the manuscript peer review process
4	Experiments
5	**Conclusions**

Potential and limits

- Identifying the point at which a paper becomes acceptable for publication, will be more tractable, and useful, than trying to analyse the top end of publishing, where fashion and politics play much larger roles.
- Ultimately, these analyses might lead to new ways to measure journal quality directly from peer review.
- Such metrics would need to be complex enough to reflect the peer-review process
- They should also take into account the qualitative level of submitted manuscripts and authors' publication strategies incl self selection practices
- They should be simple enough to be conceptually transparent and allow validation by users.

A possible practical outcome

- Publishers and/or editorial boards could set out per journal their assessment criteria by publishing a list of related statements made by the reviewers in previous reports
- Frequency of occurrence of such statements could be added
- Statements are anonymous both in terms of reviewer and reviewed author

Data collection

- Exploratory studies aimed to develop a base methodology, and to show the feasibility of the approach, could focus on peer reviewed *proceedings of international conferences*
- In a next step, journal publishers who are prepared to participate disclose under strict conditions of confidentiality parts of their electronic submission systems
- Alternatively, researchers could be invited to share the referee reports of their own manuscripts

Thank you for your attention

Short CV Henk F. Moed

1981-2009	Staff member at Centre for Science and Technology Studies (CWTS), Leiden Univ.
2009	Professor of Research Assessment Methodologies at Leiden University
Feb 2010 – Aug 2012	Elsevier, SciVal Dept. Senior Scientific Advisor
Sept 2012 – Oct 2014	Elsevier, AGRM Dept; Head of Informetric Research Group
As from Sept 2014	Visiting professor at Sapienza Univ Rome (2014) and Univ Granada (2017); Private consultant;