



Good Review Practice – how should we assess research?

Dr Matthias Kiesselbach
Division “Research Culture”

Dr Djawed Nauroozi
Division “Physics and Chemistry”



Good Review Practice – how should we assess research?

(...and how do we want
others to assess our
research?)

Dr Matthias Kiesselbach
Division “Research Culture”

Dr Djawed Nauroozi
Division “Physics and Chemistry”

Today's session

- 16:00** Overview presentation and Q&A
Research assessment – the current debate
- 16:25** Group discussions I
- Table A: Dealing with indirect markers of scientific quality**
Table B: Considering individual circumstances
- 16:45** Short break
- 16:50** Group discussions II
- Table A: Admissible (and inadmissible) uses of A.I. in research assessment**
Table B: “Out of the box” ideas to improve research assessment
- 17:10** Wrapping up: learnings, expectations, self-commitments?

Today's session

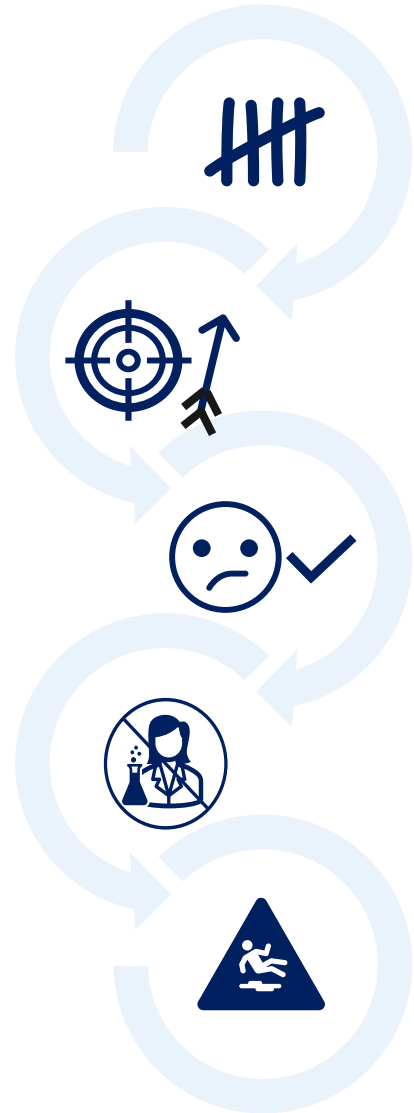
16:00 Overview presentation and Q&A
Research assessment – the current debate

- 1. Responsible research assessment**
– the current debate and movement
- 2. How does the DFG support** responsible research assessment?
- 3. Overview of the four thematic discussions:**
 - a) Dealing with indirect markers of quality
 - b) Considering individual circumstances
 - c) Admissible (and inadmissible) uses of AI in review
 - d) “Out of the box” ideas for research assessment



DFG Deutsche
Forschungsgemeinschaft

Responsible research assessment – the current debate



If research assessment is **too focused on (journal based) metrics...**

...it can **fail to identify excellent science** (false negatives)

...and come to wrongly consider **ordinary science as excellent** (false positives)...

...thus **harming excellent researchers...**

...and **harming science** and the science system, **directly and via problematic incentives**, making science less broad, less creative, less efficient, and promoting an unhealthy research culture.

Problems in research assessment

7

Matthew effect

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), S. 56-63. doi: [10.1126/science.159.3810.56](https://doi.org/10.1126/science.159.3810.56)

Matilda effect

Rossiter, M. W. (1993). The Matthew Matilda Effect in Science. *Social Studies of Science*, 23(2), 325-341. doi: [10.1177/030631293023002004](https://doi.org/10.1177/030631293023002004)

Order of authors in a publication: the order can determine the likelihood of being cited

Stevens, J. R., & Duque, J. F. (2018). Order matters: Alphabetizing in-text citations biases citation rates. *Psychonomic Bulletin & Review*, S. 1–7. doi: [10.3758/s13423-018-1532-8](https://doi.org/10.3758/s13423-018-1532-8)

The reputation of a journal is not a good predictor of the reception of a piece of research

Osterloh, M., & Frey, B. (2015 A). Ranking Games. *Evaluation Review*, 39(1), S. 102–129. doi: [10.1177/0193841X14524957](https://doi.org/10.1177/0193841X14524957)

Osterloh, M., & Frey, B. (2015 B). Rankings und der Preis der Wissenschaft. *Zeitschrift für Kulturwissenschaften*. *Der Preis der Wissenschaft*, 9(1), S. 51–63. doi: [10.25969/mediarep/13927](https://doi.org/10.25969/mediarep/13927)

Focus on speedy publications can lower the quality of research

Steen, G., Casadevall, A., & Fang, F. (2013). Why Has the Number of Scientific Retractions Increased? *PLoS ONE* (8. Juli 2013). doi: [10.1371/journal.pone.0068397](https://doi.org/10.1371/journal.pone.0068397)

Hill, R. & Stein, C. (2025) Race to the Bottom: Competition and Quality in Science, *The Quarterly Journal of Economics*, 140(2), S. 1111–1185, <https://doi.org/10.1093/qje/qjaf010>

In some fields: an inversely proportional relation between journal „rank“ and research quality

Brembs, B. (2018). Prestigious Science Journals Struggle to Reach Even Average Reliability. *Front. Hum. Neurosci.* 12:38. doi: [10.3389/fnhum.2018.00037](https://doi.org/10.3389/fnhum.2018.00037)

Brembs, B., Button, K., Munafo, M. (2013). Deep impact: unintended consequences of journal rank. *Front. Hum. Neurosci.* 7. doi: [10.3389/fnhum.2013.00291](https://doi.org/10.3389/fnhum.2013.00291)



Predatory Journals

- publish manuscripts for money - „review“ is only a show
- falsify metrics or hijack journals (e.g. using fake landing pages)
- advertise their journals / editions with fake or stolen articles of well known researchers

Grove J. (2023): Leading Scientists Worldwide Are Victims of Fake Articles, IHE:

<https://www.insidehighered.com/news/2023/02/10/leading-scientists-worldwide-are-victims-fake-articles>

Rogue reviewers blackmailing authors

Baas J. (2019): When Peer Reviewers Go Rogue, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3339568

Singh Chawla D. (2019): Elsevier investigates hundreds of peer reviewers for manipulating citations,

<https://doi.org/10.1038/d41586-019-02639-9>

Fake papers

Science (2024): Fake Papers are alarmingly common

<https://www.science.org/content/article/fake-scientific-papers-are-alarmingly-common>

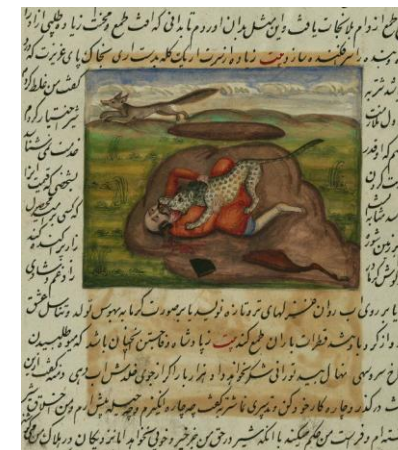
Forschung und Lehre (2023): „PAPER MILLS“ - Ein bis drei Prozent der Paper möglicherweise gefälscht

<https://www.forschung-und-lehre.de/management/ein-bis-drei-prozent-der-paper-sind-gefaelscht-6025>

Citation cartels, citations for money

Catanzaro M. (2024): Citation cartels help some mathematicians—and their universities—climb the rankings

<https://www.science.org/content/article/citation-cartels-help-some-mathematicians-and-their-universities-climb-rankings>



How easy is it to fudge your scientific rank? Meet Larry, the world's most cited cat

“Exercise in absurdity” reveals flaws in Google Scholar's productivity metrics

21 JUL 2024 - 4:30 PM ET - BY CHRISTIE WILCOX





*The Leiden Manifesto
for research metrics*

Several reform initiatives – two core demands



When assessing research, a **broad(er) diversity of types of output and activity should count.**

(Not **just journal articles** – also software code, datasets, work on research infrastructure...)

Assessing research (e.g. proposals, previous work) **should be done in a predominantly qualitative way.**

(The use of quantative proxies – such as JIF, h-index... – should be limited and extremely careful.)



How does the **DFG** support responsible research assessment?

previous works listed in
applications and CVs:

maximum of ten entries
instated



A tradition of qualitative assessment

- 1998 **Memo „Safeguarding Good Scientific Practice“**
- 2004 **Senate Commission for Clinical Research**
- 2010 **„Quality before Quantity“
in proposals and CVs**

2022 Policy package supporting qualitative assessment

- **Novel CV template** with optional narrative elements (on biography, activities in the research system e.g.)
- Previous work mentioned in application must be **explicitly summarized and linked to current proposal**
- Reviewers asked to consider **full breadth of research contributions** beyond (just) journal articles

see https://www.dfg.de/en/news/news-topics/announcements_proposals/2022/info-wissenschaft-22-61



Supplementary Career Information optional; free text

[free text, please overwrite]

Activities in the Research System optional, free text

Here you can provide information on other activities you have pursued within the research system. This includes committee involvement, activities in the field of academic self-governance, organisation of academic events, activities in teaching and mentoring.

[free text, please overwrite]

Supervision of Researchers in Early Career Phases optional, free text/

Curriculum Vitae [09/22]:
https://www.dfg.de/formulare/53_200_elan/53_200_de_elan.rtf

Two categories of scientific results
... to ensure that „non-classical“ contributions
are visible

(i.e. that Category A does not “cannibalize”
Category B)

Curriculum Vitae [09/22]:
https://www.dfg.de/formulare/53_200_elan/53_200_de_elan.rtf

Scientific Results *Part A required, Part B optional; free text*

Please indicate here your most important published scientific results (see also “Guidelines for Preparing Publication Lists”, [DFG form 1.91](#). If available, please also provide persistent identifiers (e.g. DOI/Digital Object Identifier), preferably by stating the number, otherwise by naming the URL. Open access publications should be designated accordingly.

Details of quantitative metrics such as impact factors and h-indices are not required and are not considered as part of the review.

Please also explain – where possible – how you were involved in the published findings and/or explain why you have listed the publication or the academic contribution here.

These details fall into two categories:

Category A *required, free text*

*In this category please enter articles in peer-reviewed journals, peer-reviewed contributions to conferences or anthology volumes, and book publications (see also [DFG form 1.91](#)). A **maximum of ten items** may be listed.*

[free text, please overwrite]

Category B *optional, free text*

*Here you can cite **any other form of published research results**. This might include articles on preprint servers and non-peer-reviewed contributions to conferences or anthology volumes, data sets, protocols of clinical trials, software packages, patents applied for and granted, blog contributions, infrastructures or transfer. You may also **indicate other forms of academic output** here, such as contributions to the (technical) infrastructure of an academic community (including in an international context) and contributions to science communication. **This second category is also restricted to a maximum of ten items.***

[free text, please overwrite]

Applicants are asked to discuss their previous work / publications
– not just list them

1 Starting point

State of the art and preliminary work

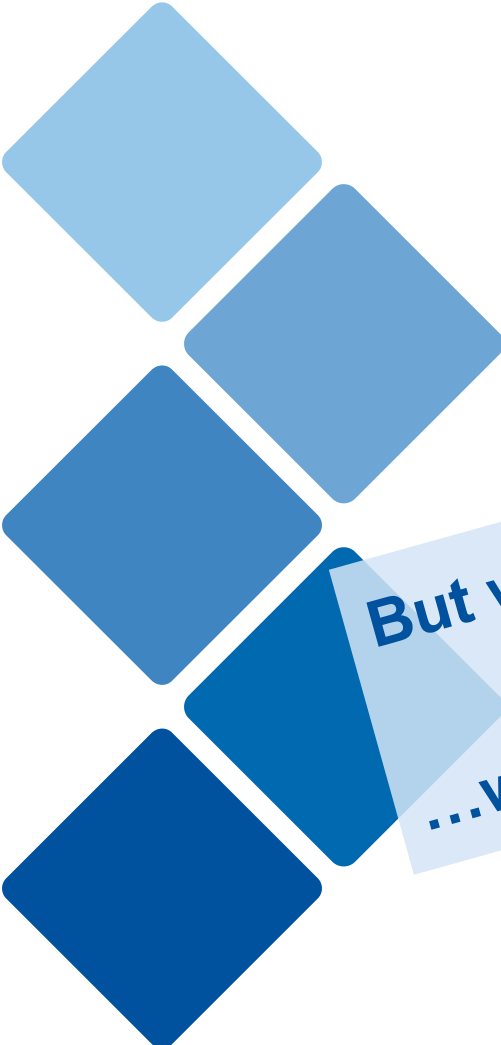
For new proposals please explain briefly and precisely the state of the art in your field in its direct relationship to your project. This description should make **clear in which** context **you situate your own research and in what areas you intend to make a unique, innovative, promising contribution**. Indicate the current state of your preliminary work. This description must be concise and understandable without referring to additional literature.

Proposal preparation instructions [03/24] („Leitfaden“)
https://www.dfg.de/formulare/54_01/54_01_en.pdf

Reviewers asked to recognize full breadth of scientific contributions, avoid metrics

Assessment of the achievement of a researcher must be carried out **in its entirety and based on substantive qualitative criteria**. In addition to the publication of **articles, books, data and software**, other dimensions can be taken into account, such as involvement in teaching, academic self-administration, public relations or knowledge and technology transfer. Details of **quantitative metrics such as impact factors and h-indices** are not required and are not to be considered as **part of the review**.

General Guidelines for Reviews [09/24]:
https://www.dfg.de/formulare/10_20/10_20_en.pdf



Comprehensive **information sessions for research boards** at the beginning of their tenure, highlighting research assessment (among other things)

DFG officers moderating review board review sessions are sensitized to issues of responsible research assessment

**But we also need a change in everyday review culture...
...within and outside the DFG's review sessions**

DFG is using short videos for reviewers, for example on implicit biases
(<https://www.youtube.com/watch?v=Kc2pPvwY&pp=ygUIZGZnIGJpYXM%3D>)

Membership in the Coalition for Advancing Research Assessment (**CoARA**), advocacy work and participation in various Working Groups (<https://www.coara.org>)

...also: **support for alternative publishing venues**

...for instance via the National Service Centre Diamond Open Access (**SeDOA**) and Open Research Europe (**ORE**)

Overview of the **four proposed group discussions**

Table 1

Dealing with indirect markers of scientific quality

„X has published four articles in top tier journals“, „Y collaborates with very renowned P, but so far, this has yielded only a pre-print“, „Z gained experience in Q lab“...

Should a funder accept these statements as valid currency of scientific expertise/experience? What biases are at play? Should funders require fuller explanations of the resp. meaning in a given context? Should funders insist on qualitative back-up (e.g. *what* did the mentioned papers establish)?

Table 2

Considering individual circumstances

How should funders deal with biographical information in external reviews? (Strike from the record? Only if negative? Only if it addresses certain illegitimate points? Allow if it's relevant to estimate the project's scientific success? Allow if it addresses certain selected dimensions? If possible, provide commentary from CV? Ask the applicant to comment? Allow such information to only have positive effect? How? How can bias be countered?) What if this happens in a panel session? How should funders introduce information from an applicant's CV? Highlight it pro-actively? Introduce only when needed?

Table 1

Admissible (and inadmissible) uses of A.I. in the review

Where should funders draw the line (if at all)? How to treat using AI to... summarize proposal, summarize supporting publications or data-sets, summarize CV, provide bullet points for review, provide first draft of review, provide answers to specific smaller questions regarding proposed project, *any AI use...*? What kind of transparency is needed, how is it usefully ensured? How does AI use relate to DFG's policy goals (equality / EDI, qualitative review)?

Table 2

„Out of the box“ ideas to improve research assessment

How to ensure qualitative review and consideration of all relevant works? How about equity and fairness?

Free discussion, possible themes:

- Organisation of the review discussion
- Moderation of the review session
- Recruitment of reviewers (who? from where?)
- ...



Thank you for your participation!

More information:

<http://www.dfg.de/en>

Matthias Kiesselbach

Matthias.Kiesselbach@dfg.de

Djawed Nauroozi

Djawed.Nauroozi@dfg.de

DFG Deutsche
Forschungsgemeinschaft



www.dfg.de



Bluesky | @dfg.de



Instagram | dfg__public



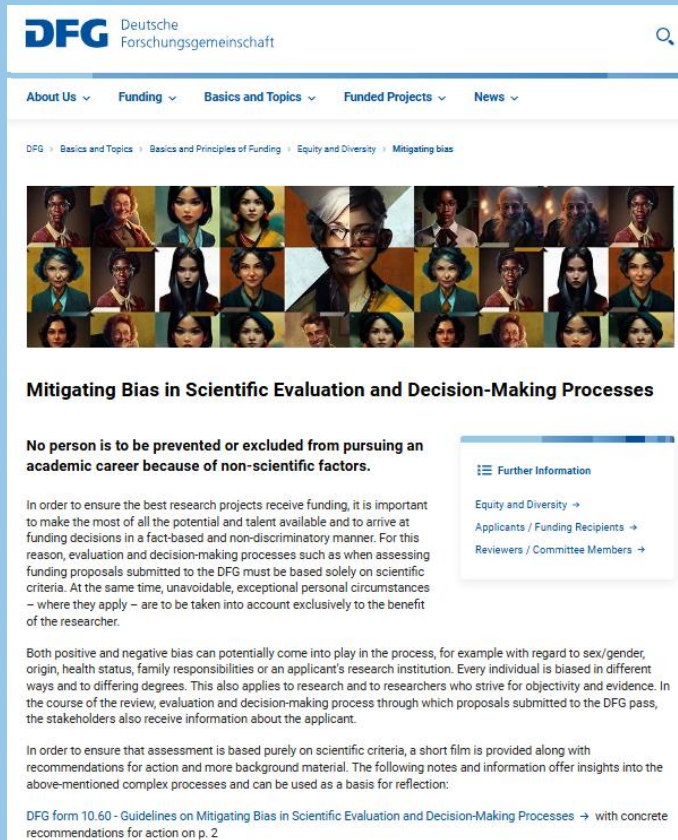
LinkedIn | Deutsche Forschungsgemeinschaft
(DFG) – German Research Foundation



Mastodon | @dfg_public



YouTube | @DFGbewegt



The screenshot shows the DFG (Deutsche Forschungsgemeinschaft) website. The header includes the DFG logo and navigation links: About Us, Funding, Basics and Topics, Funded Projects, and News. The main content area is titled 'Mitigating Bias in Scientific Evaluation and Decision-Making Processes'. It features a grid of diverse people's faces at the top. Below the title, a bold statement reads: 'No person is to be prevented or excluded from pursuing an academic career because of non-scientific factors.' The text explains the importance of ensuring the best research projects receive funding based on scientific criteria, while also acknowledging that bias can occur. It mentions that both positive and negative bias can play a role, and that the DFG provides guidelines to mitigate this. A sidebar on the right offers 'Further Information' links for Equity and Diversity, Applicants / Funding Recipients, and Reviewers / Committee Members. At the bottom, there is a link to 'DFG form 10.60 - Guidelines on Mitigating Bias in Scientific Evaluation and Decision-Making Processes'.



- **Bias** also exists in research and in science and the humanities.
- Biased funding decisions are not science-led and can be discriminatory.
- Funding decisions by the DFG must not be made on the basis of irrelevant considerations.
- At the same time, unavoidable delays in the applicant's academic career (if voluntarily stated) should be taken into account appropriately in favour of the applicant.



ased Review"

