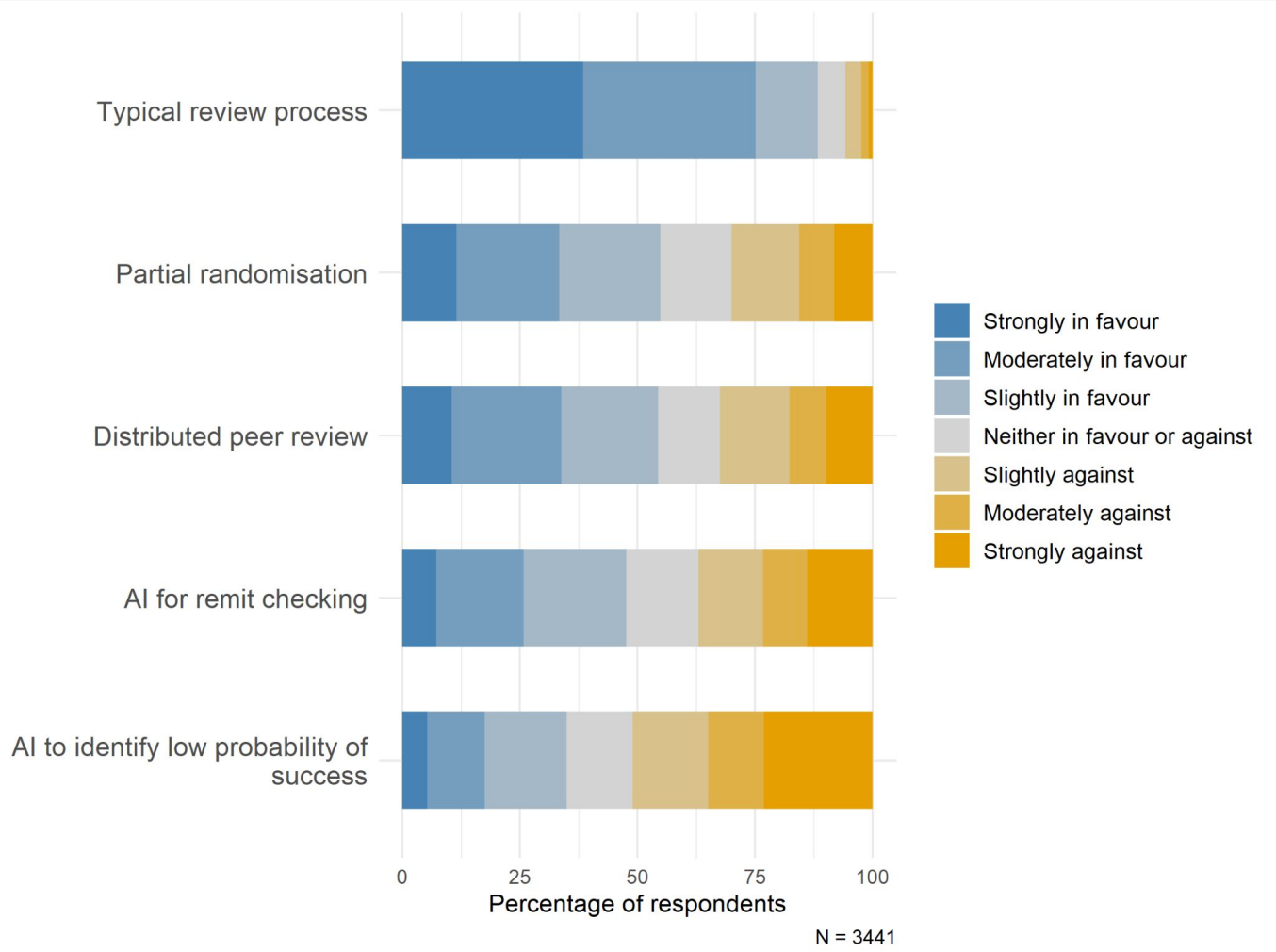
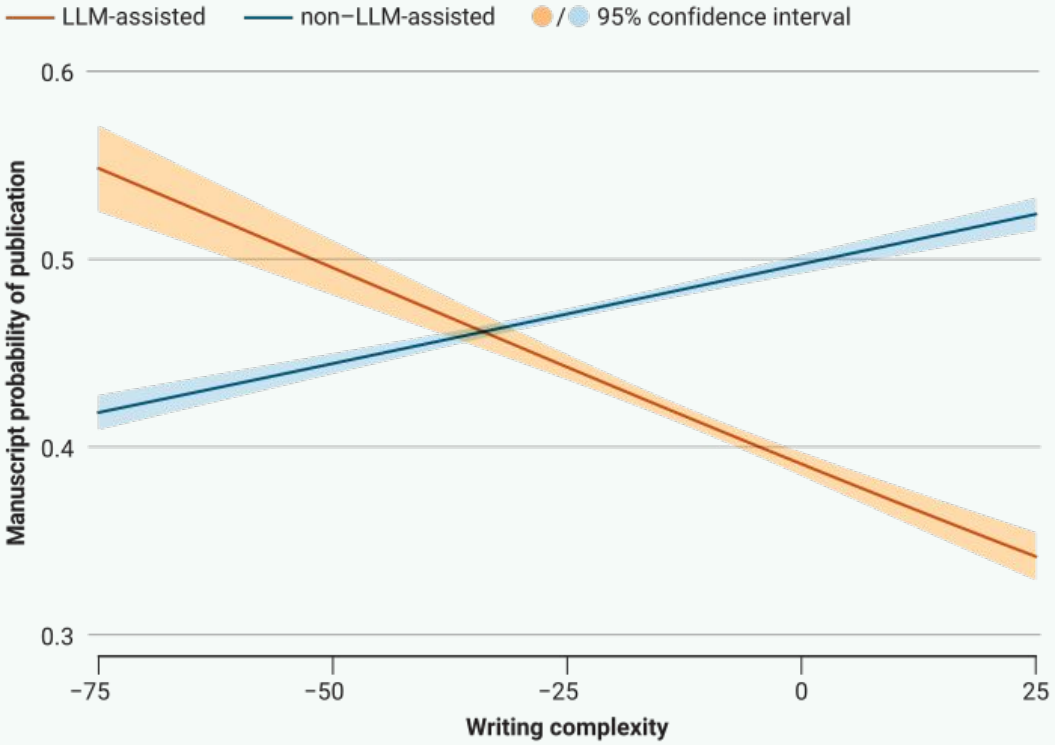

Typical review process
Partial randomisation
Distributed peer review
AI for remit checking
AI to identify low probability of success
0
25
50
75
100
Percentage of respondents
N = 3441
Strongly in favour
Moderately in favour
Slightly in favour
Neither in favour or against
Slightly against
Moderately against
Strongly against

Scientific production in the era of large language models

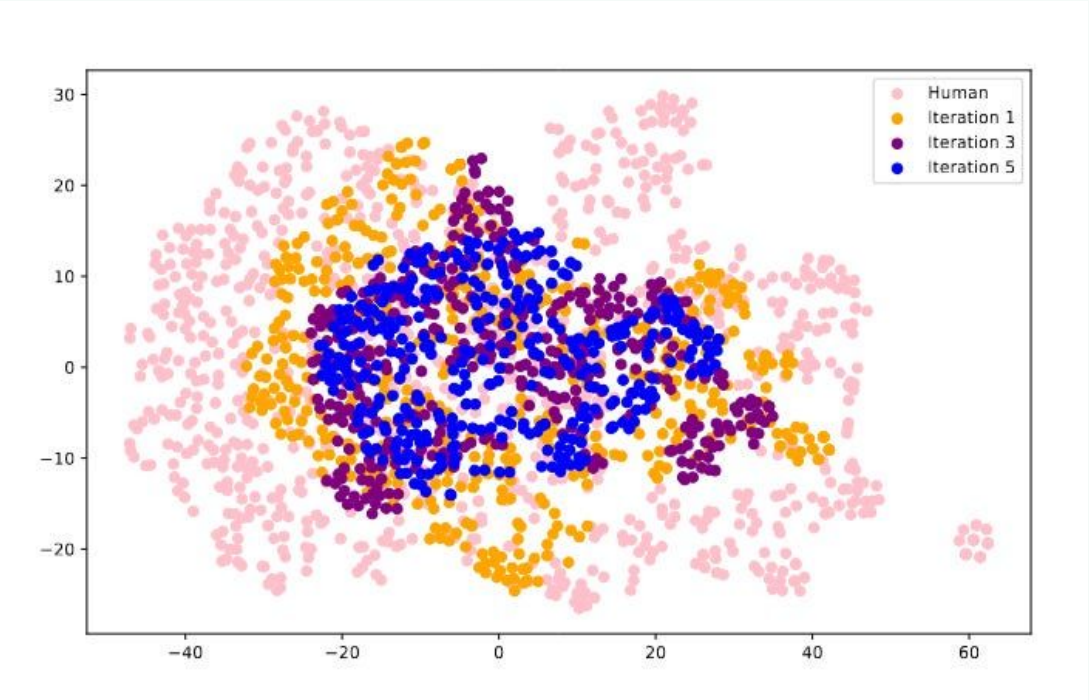
https://doi.org/10.1126/science.adw3000



Signal loss

+ Increased productivity
+ Gain greatest for non-native english speakers
+ More diverse literature cited

Model collapse



Guo et al (2024), Figure 4 https://doi.org/10.18653/v1/2024.findings-naacl.228

RESEARCH ON RESEARCH INSTITUTE

Distributed Peer Review at the Volkswagen Foundation

Tom Stafford, Anna Butters, Melanie Benson Marshall, & Stephen Pinfield
(Research on Research Institute and University of Sheffield)

Hanna Denecke, Alexander Bondarenko, Robert Nuske & Pierre Schwidlinski
(Volkswagen Foundation)

28/09/2026

These slides:
bit.ly/tomstafford

DPR experiments at The Volkswagen Foundation

- Applicants review other applications submitted for the same funding call
- “Open Up” programme – interdisciplinary projects, focus on innovation in the Humanities and Social Sciences
- Parallel implementation of DPR and established panel review process
- Additional funding provided: funding recommendations from both panel review and DPR
- Mixed methods evaluation: quantitative analysis of data from submissions and surveys of applicants, and qualitative interviews with a sample of applicants and other stakeholders
- Rich datasets to gain insight into dynamics of grant peer review e.g.:
  - ➔ Comparisons between review processes
  - ➔ Reviewer uncertainty
  - ➔ Consistency between reviewers
  - ➔ Stability of funding decision
  - ➔ Attitudes of actors

These slides:
bit.ly/tomstafford

First DPR experiment

- 2024 ‘Open Up’ call
- ‘Live’ funding experiment
- Wide interest in and dissemination of results:
  - RoRI working paper
  - RoRI insights guide
  - Journal article
  - Metascience 25, ICSSI

DPR 2.0

- Experiment repeated in 2025
- Minor changes to the call and procedure
- Evaluation extended
  - Interviews with other funders
- Comparison across years
- New outputs!

These slides:
bit.ly/tomstafford

Distributed Peer Review and Panel Review Process – Parallel Processes

Panel review process

Internal shortlisting
70 shortlisted in both years

Review by panellists
140 reviews

Panel discussion

Proposals recommended for funding

Proposals submitted
2024: 140
2025: 161

DPR

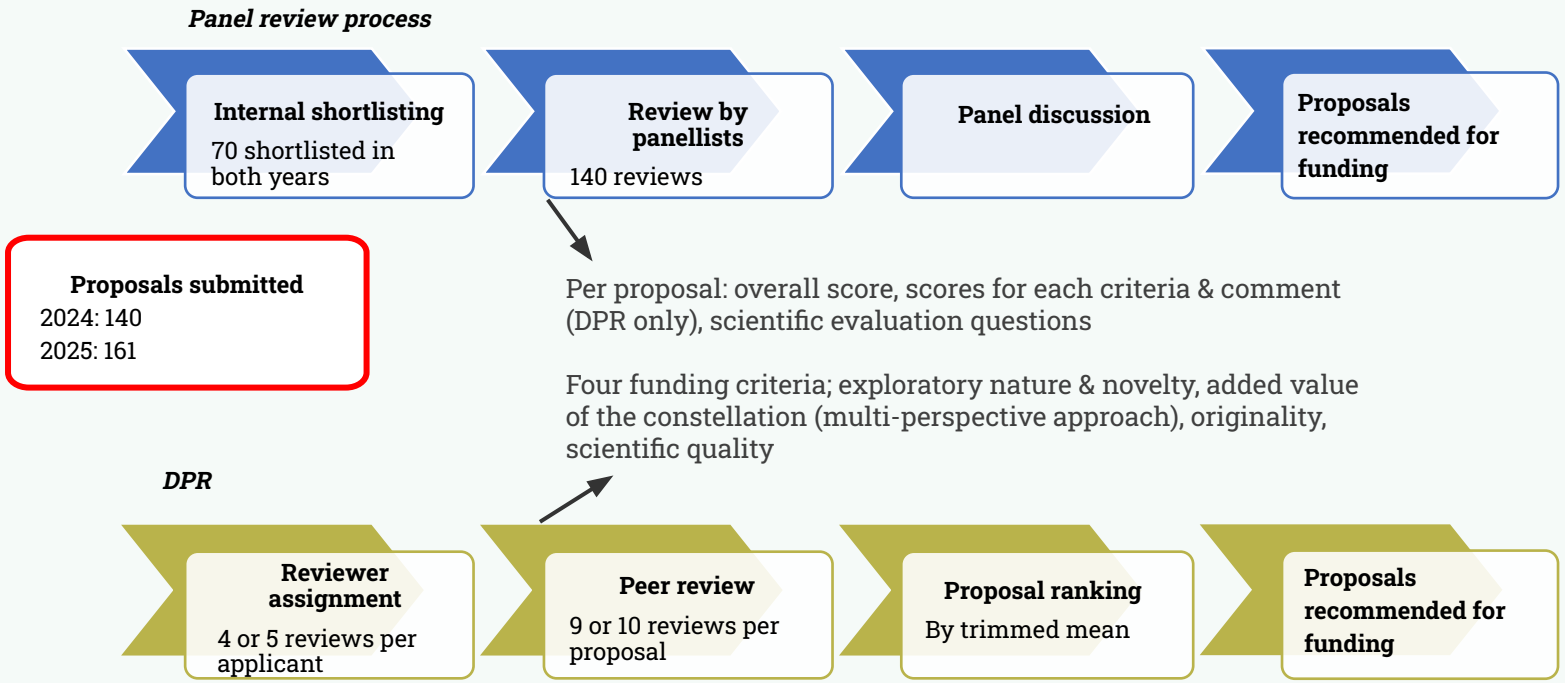
Reviewer assignment
4 or 5 reviews per applicant

Peer review
9 or 10 reviews per proposal

Proposal ranking
By trimmed mean

Proposals recommended for funding

Distributed Peer Review and Panel Review Process – Parallel Processes



18 proposals funded in each year

Panel review process

Internal shortlisting
70 shortlisted in both years

Review by panellists
140 reviews

Panel discussion

Proposals recommended for funding
2024: 11
2025: 9

Proposals submitted
2024: 140
2025: 161

15% increase in submissions

Proposals selected by both processes
2024: 3
2025: 1

DPR

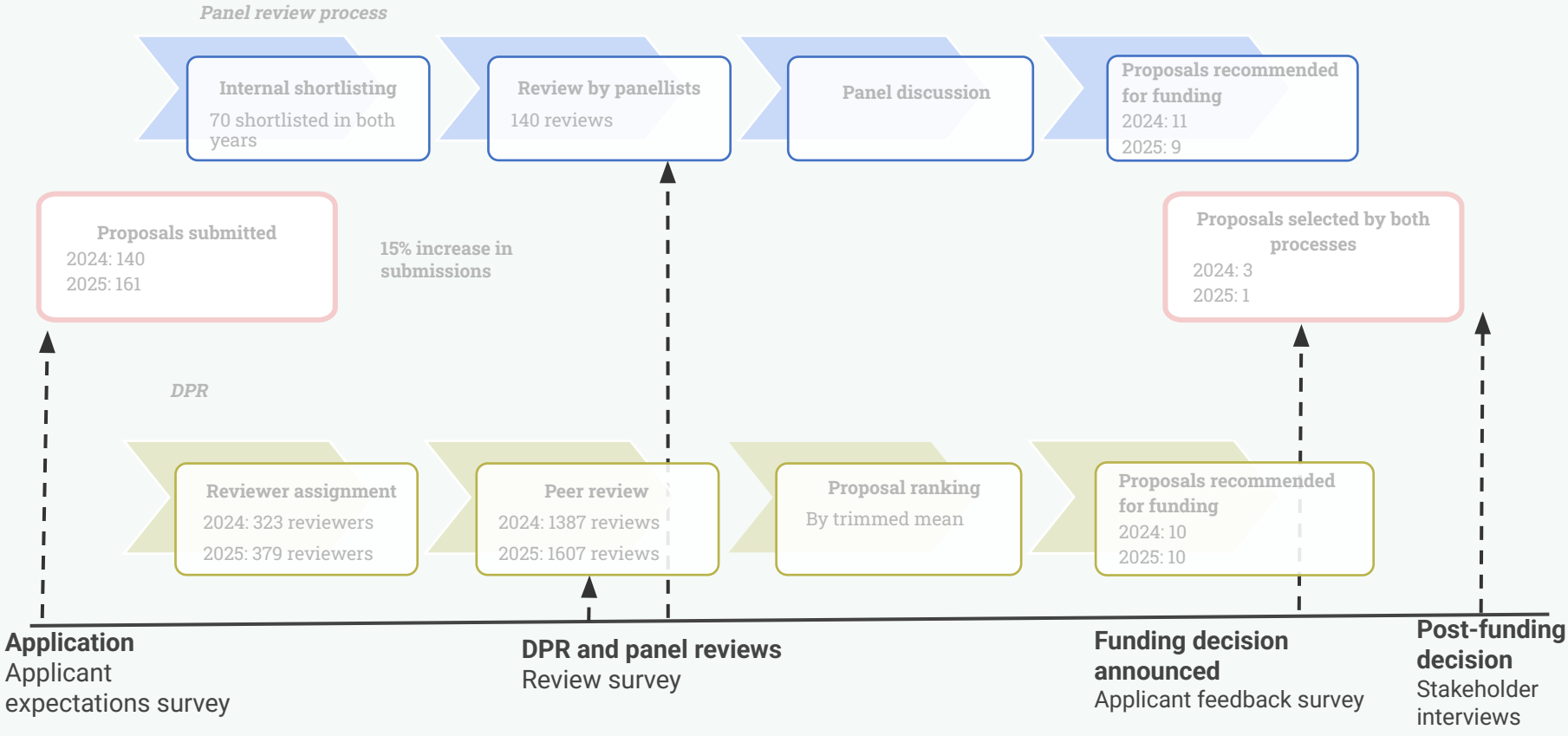
Reviewer assignment
2024: 323 reviewers
2025: 379 reviewers

Peer review
2024: 1387 reviews
2025: 1607 reviews

Proposal ranking
By trimmed mean

Proposals recommended for funding
2024: 10
2025: 10

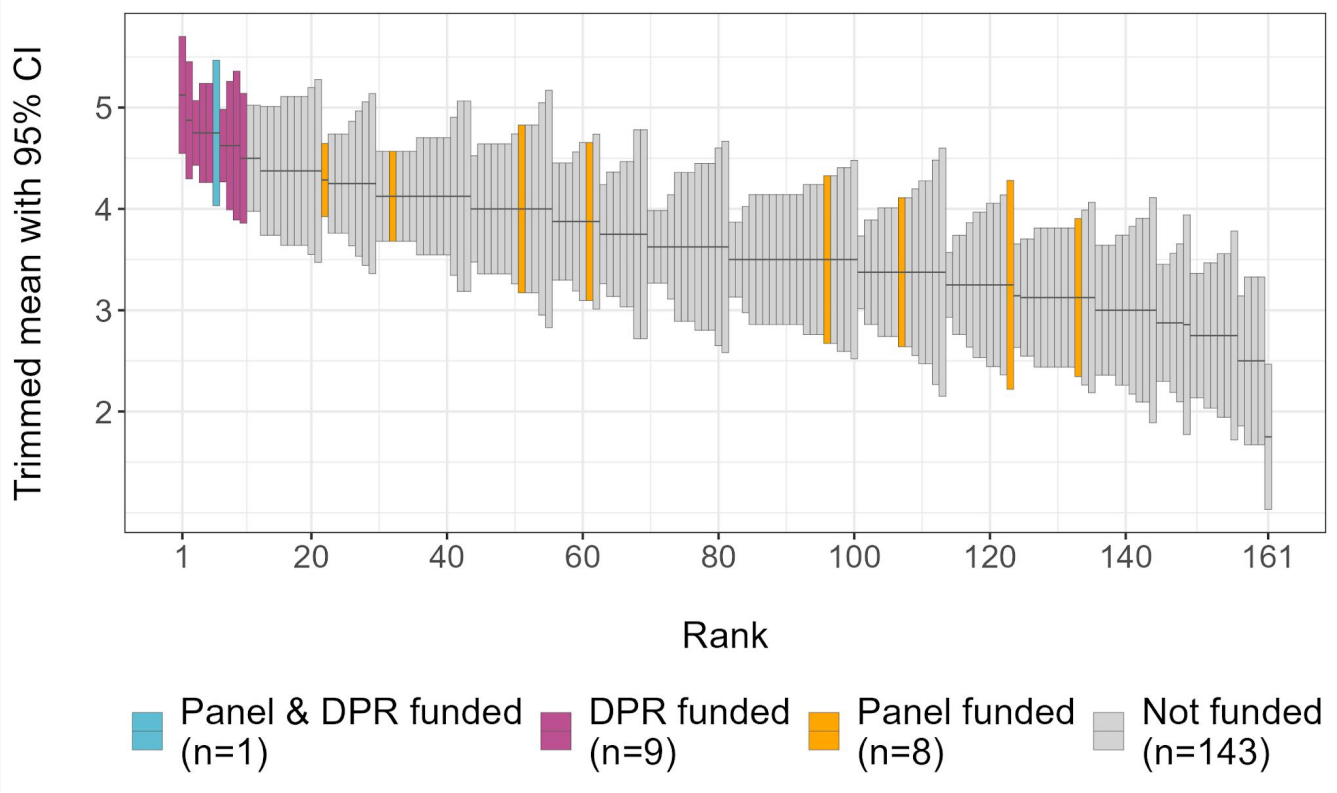
Data collected at multiple timepoints: application, on review, funding decision



Comparing DPR and panel funding decisions

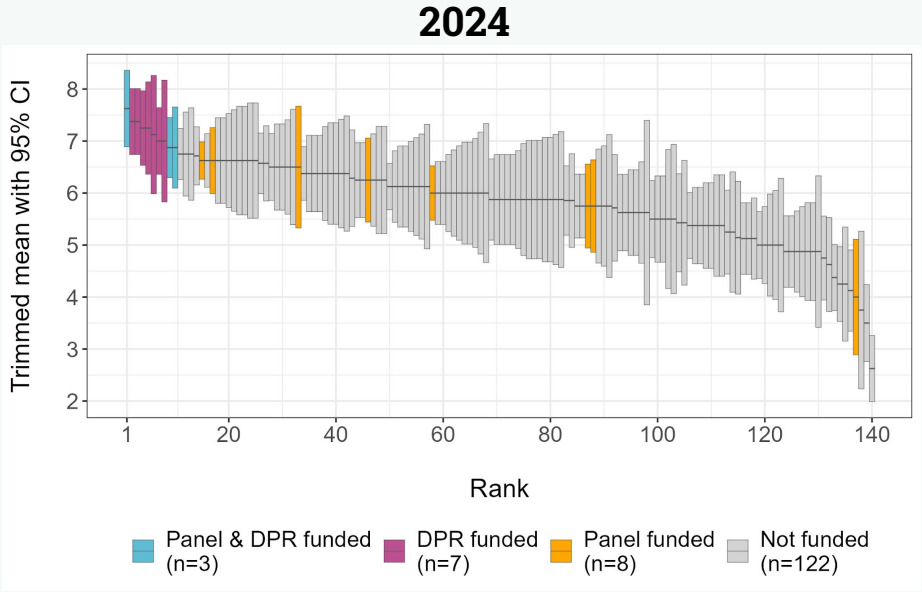


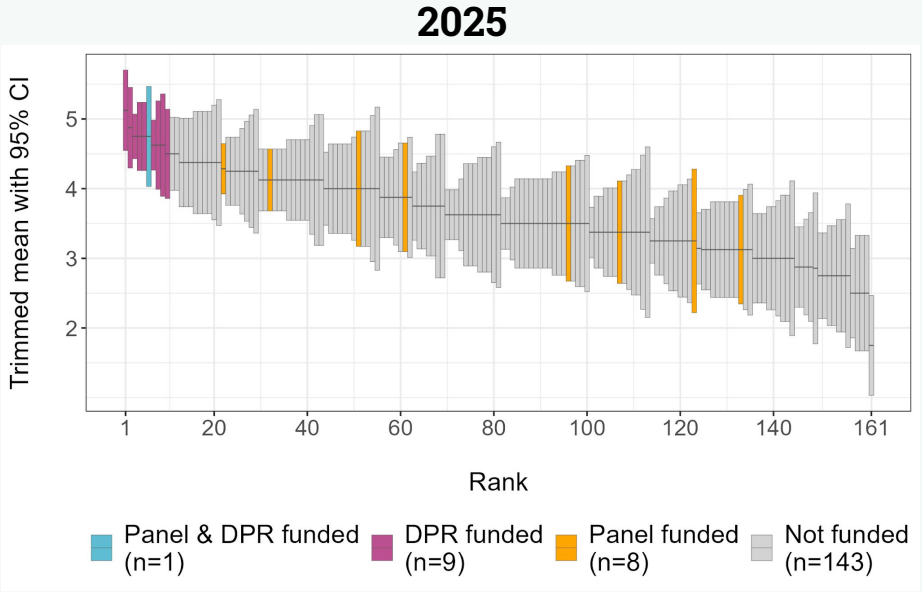
Panel selected proposals are found across the full range of DPR scores



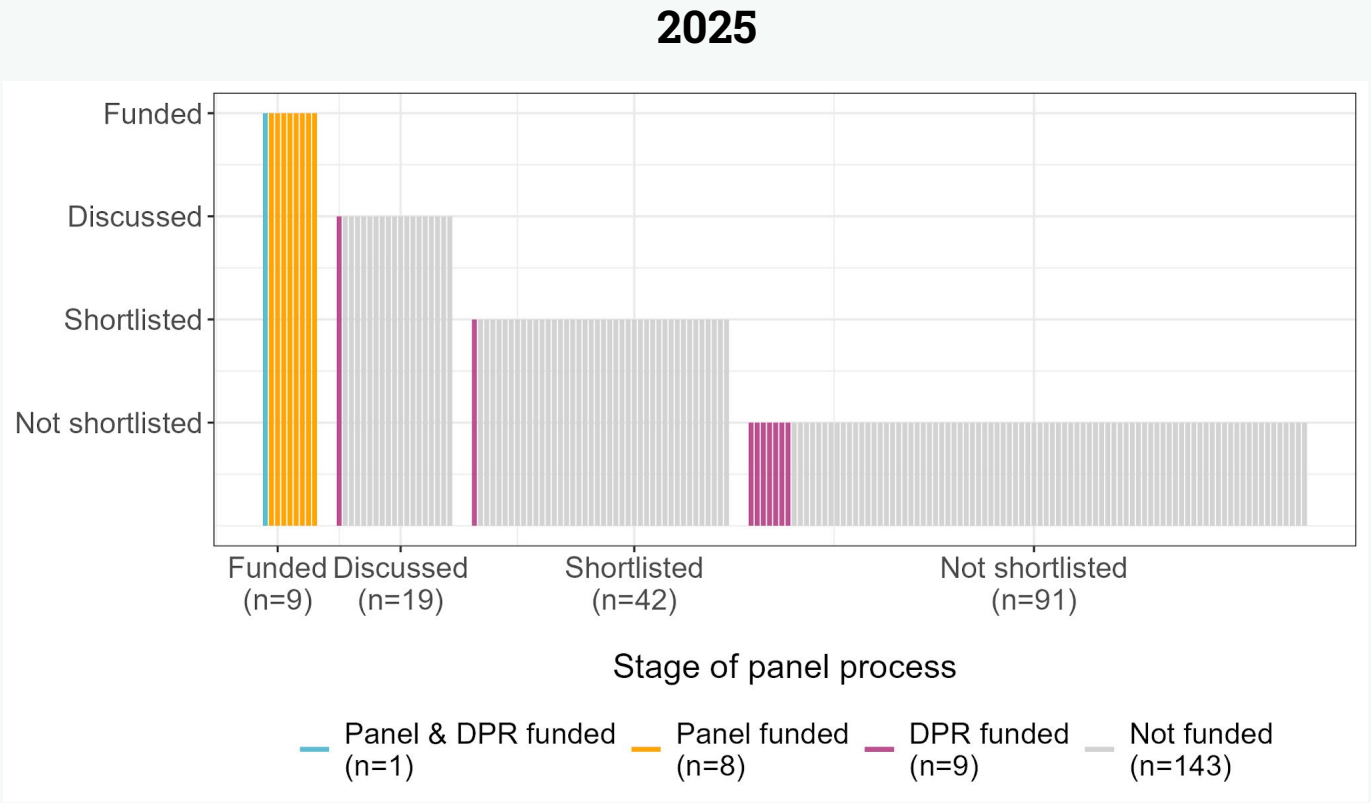
2025

Overlap was lower in 2025, but **consistent with the same level of agreement as 2024**

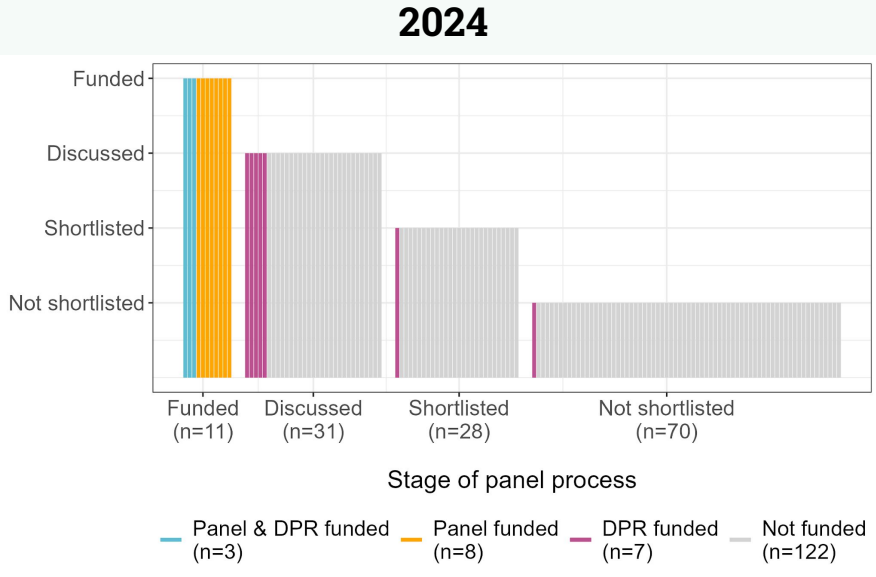


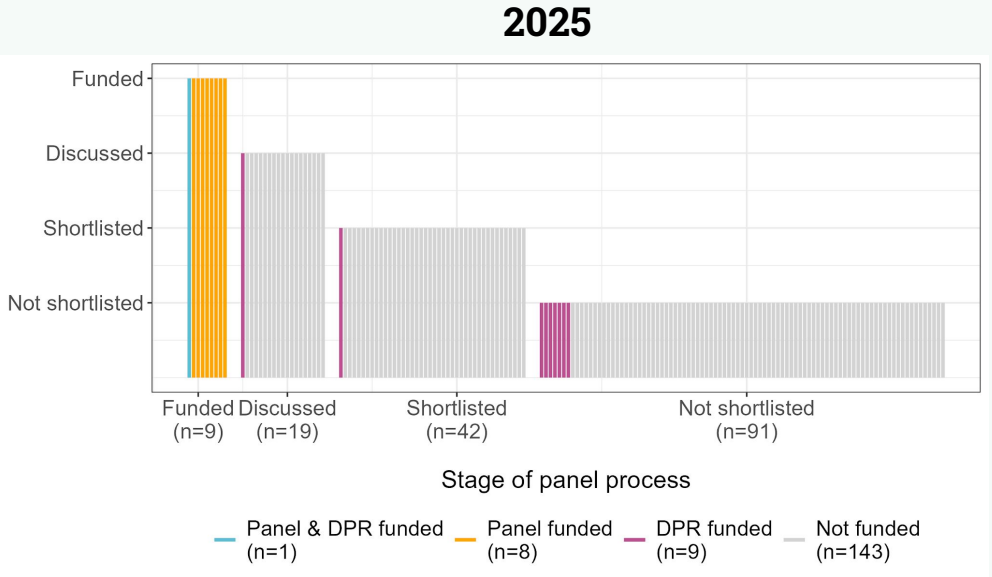


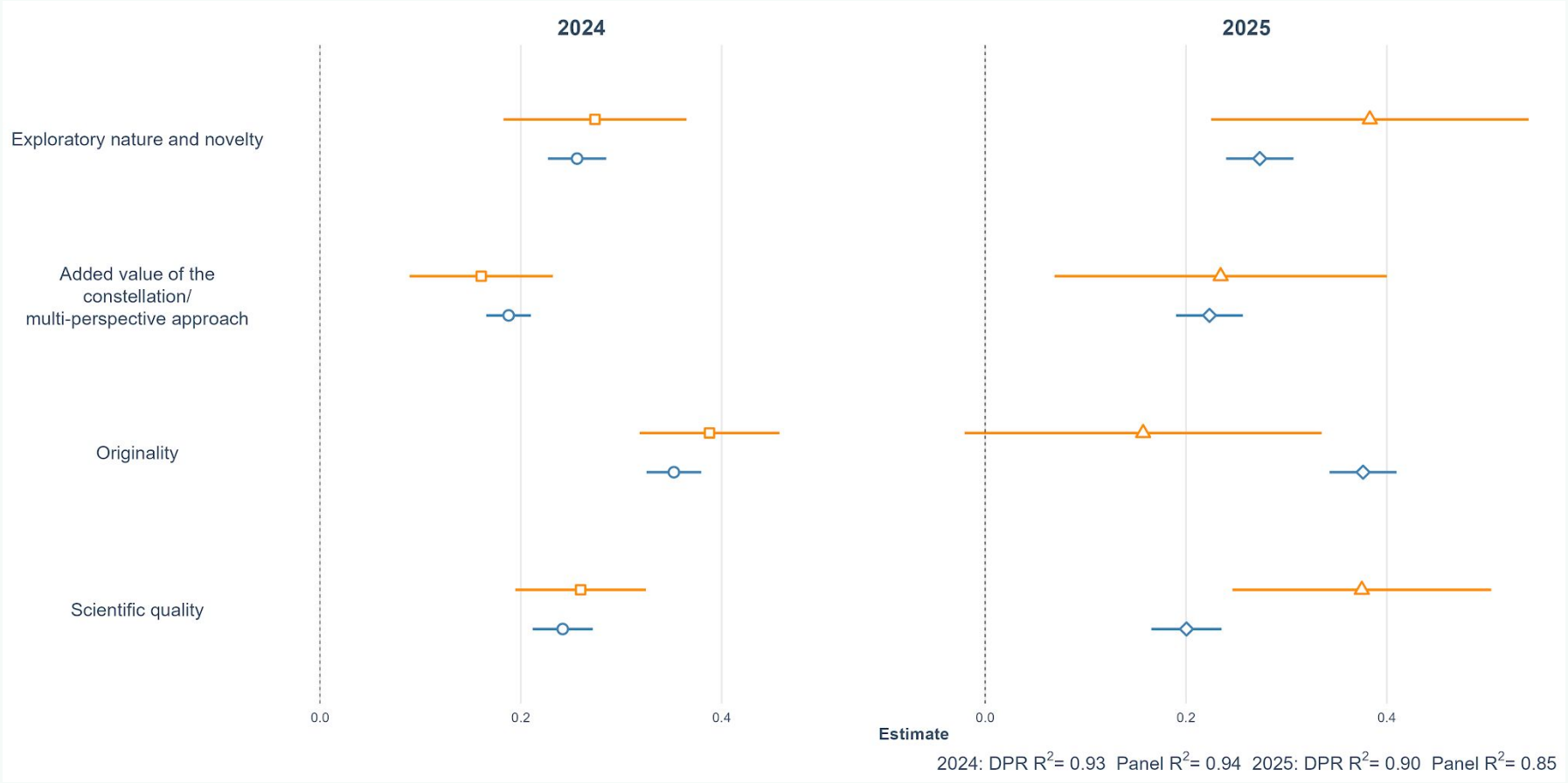
Panel selected proposals are found across the full range of DPR scores



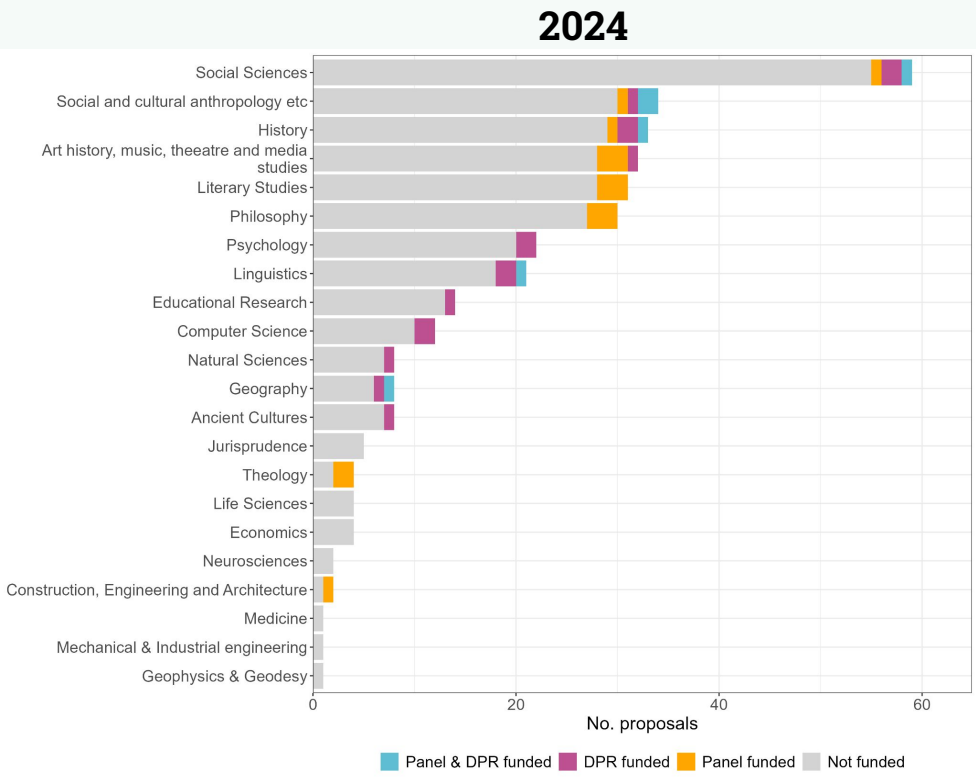
More DPR funded proposals were shortlisted in the panel process in 2024

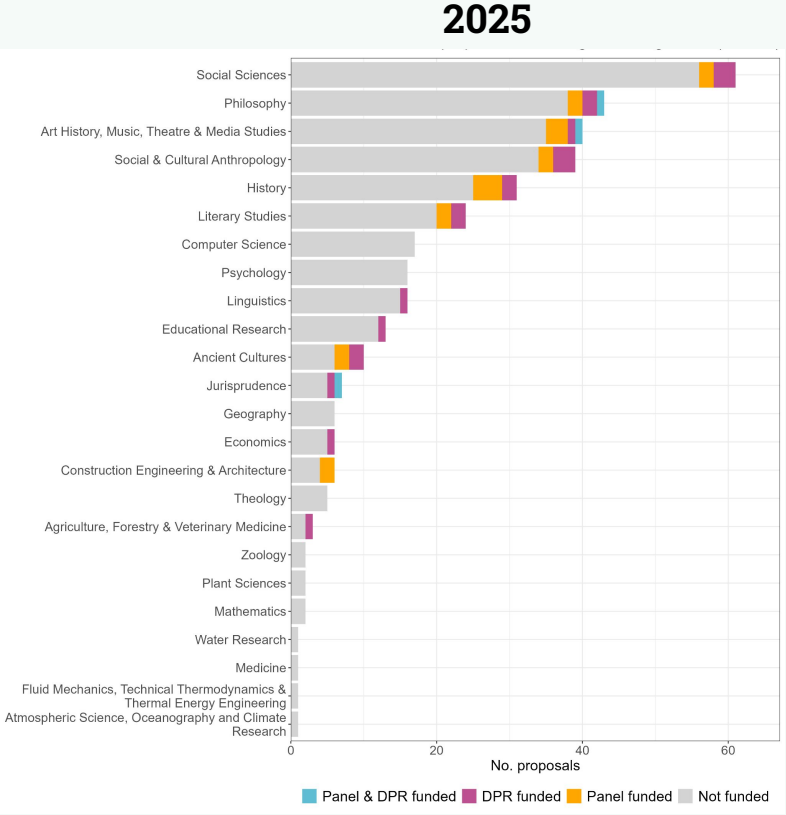





2024
2025
Exploratory nature and novelty
Added value of the constellation/ multi-perspective approach
Originality
Scientific quality
0.0
0.2
0.4
Estimate
2024: DPR R^2= 0.93 Panel R^2= 0.94 2025: DPR R^2= 0.90 Panel R^2= 0.85

DPR funded proposals spanned slightly more unique disciplines





Expectations and experience of DPR: What do stakeholders think?

Applicant expectations were consistently fairly positive

Applicant characteristics were similar in both years:

- ~ 47% identified as female
- ~ 35% first generation university student
- Working in research for, on average, 14.5 years

Before participation:

- Over 73% believed DPR was **fair**
- Over 76% thought it would **identify appropriate reviewers**
- Around 85% believed it was **suitable** to assess their proposal
- Over 72% believed it would **select the best** proposals
- Over 68% thought DPR would **select more adventurous** proposals
- **Less than 50%** expected **DPR and panel to select similar set of proposals**

But also some concerns:

- **Additional workload**
- **Insufficient expertise**
- **'Gaming'**

Feedback was also fairly positive

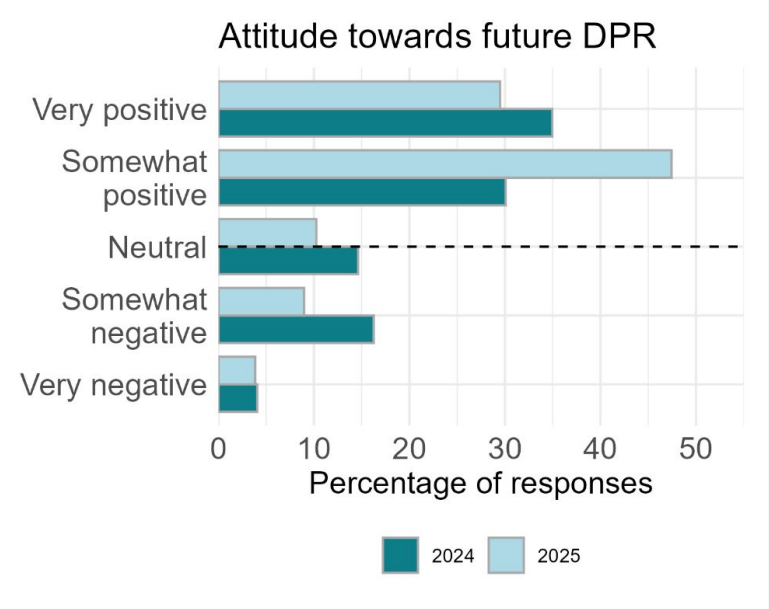
Optional feedback surveys:

- 127 applicants (39.3%) in 2024
- 91 applicants (24%) in 2025

Over 63% of review comments considered constructive

Most *funded* respondents positive about taking part in DPR in future
(2024 = 82.8%, 2025 = 94.5%)

Many *unfunded* respondents also positive
(2024 = 59.6%, 2025 = 71.7%)



Feedback comments and stakeholder interviews revealed benefits & disadvantages



Benefits

More efficient and can shorten time to decision.

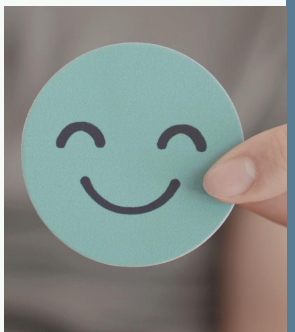
More (diverse) feedback

Addresses issues of reviewer recruitment

Educational

Enjoyable/interesting process

Fairer/more democratic



Disadvantages

Increased burden for applicants

Gaming

No opportunity for discussion/clarification

Reviewers lack experience or expertise

Reviewer experience & expertise

Minor differences in review comments according to career stage

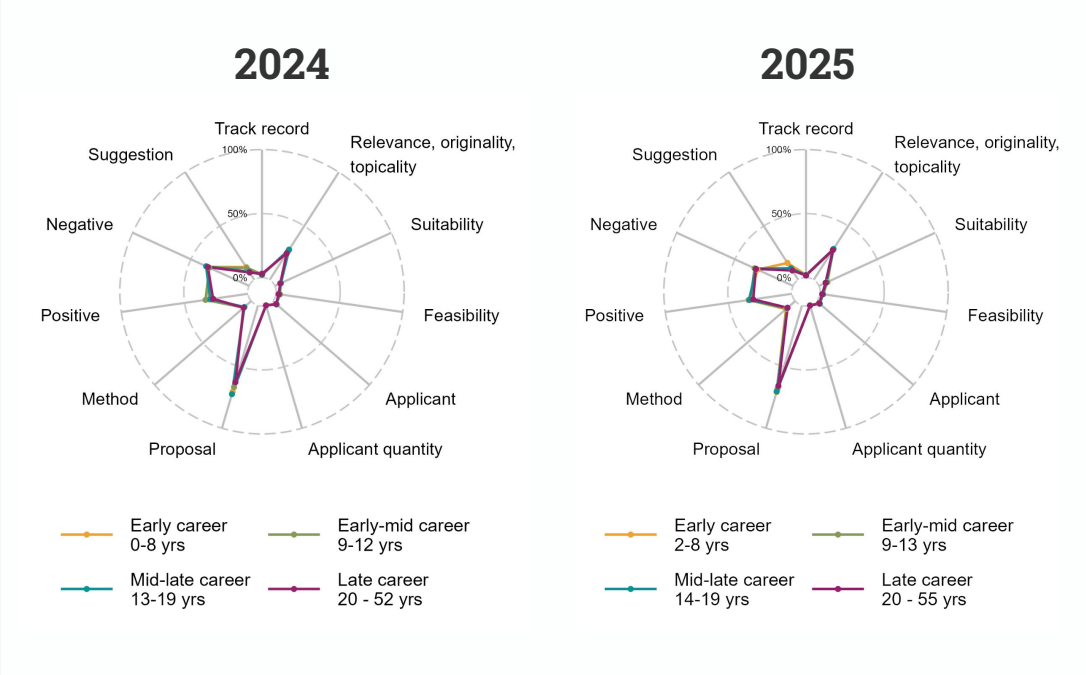
Automated analysis of review comments from 2024 and 2025

Content categories and sentiment per sentence

E.g. '*Does the sentence address the feasibility of the proposed research project?*'

Prevalence per comment and on average

Minor variation in comment content/sentiment; small effects



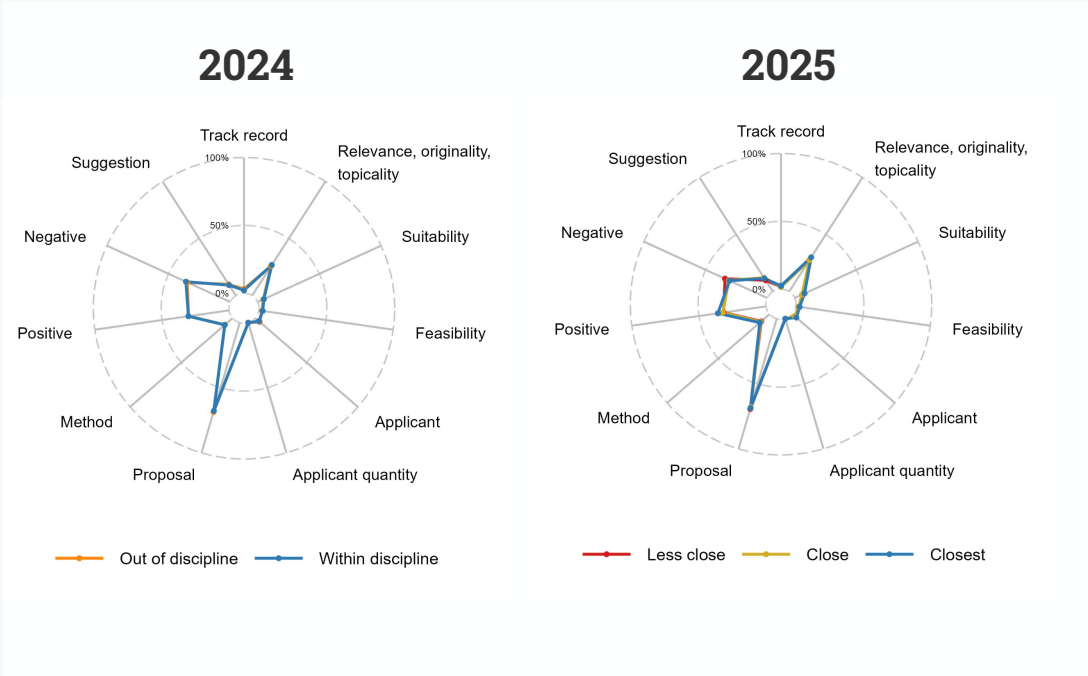
Minimal differences according to disciplinary alignment

Alignment between applicant disciplinary expertise and proposal under review:

2024: **within/ outside of discipline**

2025: **disciplinary closeness** (less close/close/closest)

Minimal differences in content/sentiment of comments or scores awarded according to disciplinary alignment



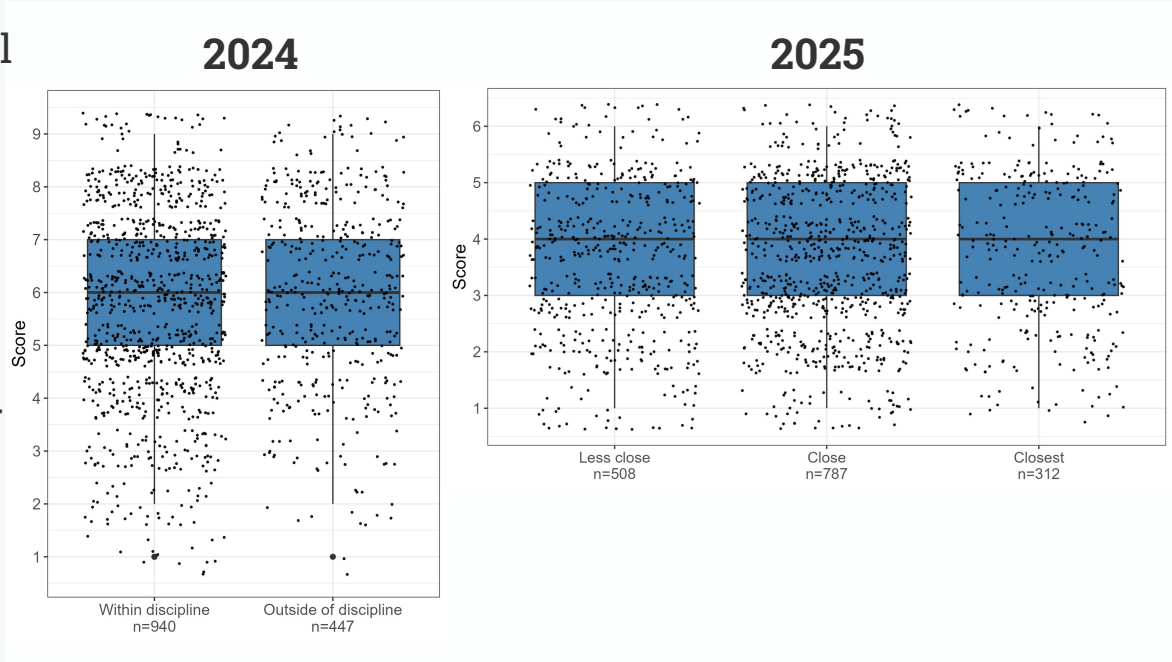
Minimal differences according to disciplinary alignment

Alignment between applicant disciplinary expertise and proposal under review

2024: **within/ outside of discipline**

2025: **disciplinary closeness** (less close/close/closest)

Minimal differences in content/sentiment of comments or scores awarded according to disciplinary alignment

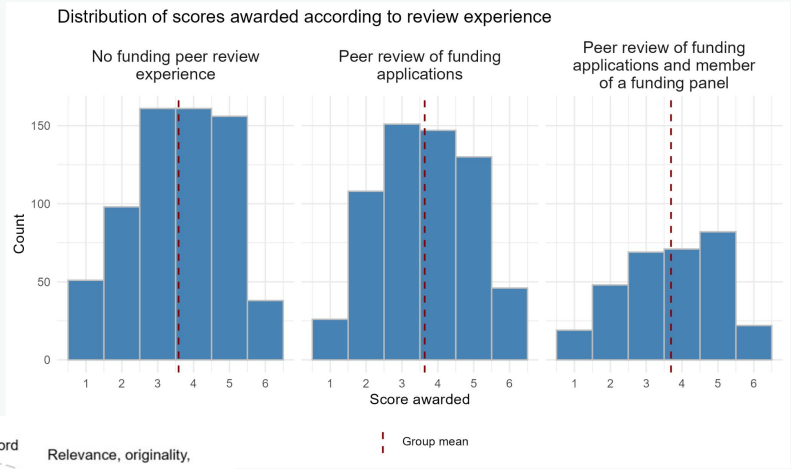


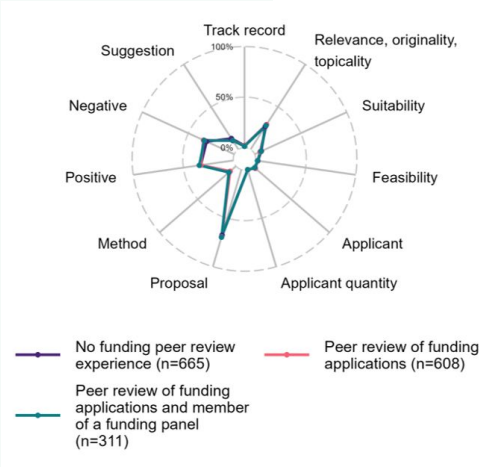
Minimal differences according to previous reviewing experience

60% of applicants had previous experience of ***funding*** peer review

No significant difference in average scores awarded; the most experienced applicant-reviewers tended to award slightly higher scores

Minimal differences in comment content and sentiment





These slides:
bit.ly/tomstafford

Findings so far…

- Many findings consistent across both years; others only partially consistent
- Some overlap between DPR and panel decisions but also divergence in processes and selected proposals
- DPR increases the total time spent reviewing but this is distributed more equally between more people
- Modest stability of funding decisions, potentially due to inter-reviewer variability; trimming helps
- No evidence that less experienced or more disciplinary distant reviewers are ‘worse’ reviewers
- Suggested benefits: faster time to decision, ‘democratisation’ of funding evaluation
- Common concerns: applicant workload, ‘gaming’, reviewer expertise

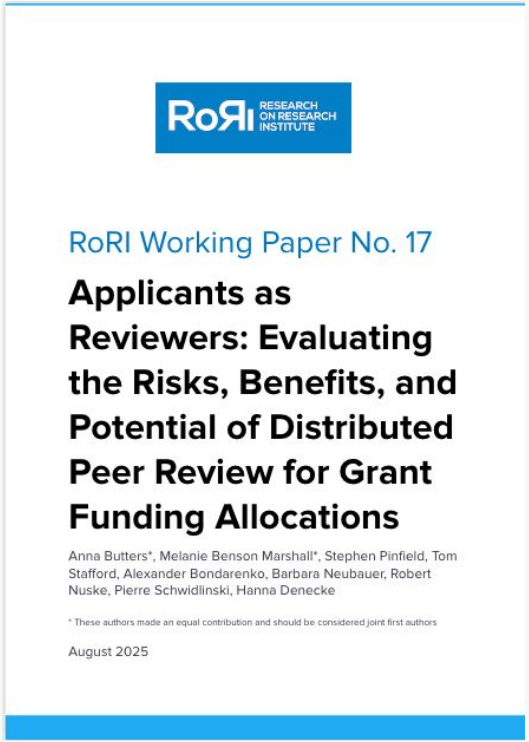
Conclusions

- Comparison between two processes is challenging; conventional system seen as ‘tried and tested’
  - Avoid considering familiar systems as normative; recognise both have trade-offs
- Implications for peer review more widely:
  - Reflects the different functions of peer review
  - From the ‘wisdom of the gatekeeper’ to the ‘wisdom of the (expert) crowd’

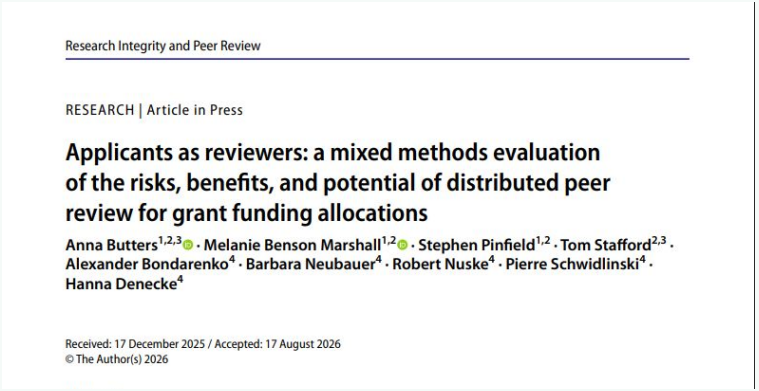
These slides:
bit.ly/tomstafford



https://doi.org/10.6084/m9.figshare.29270534.v1



https://doi.org/10.6084/m9.figshare.29994841



https://doi.org/10.1186/s41073-026-00246-7

Still to come!

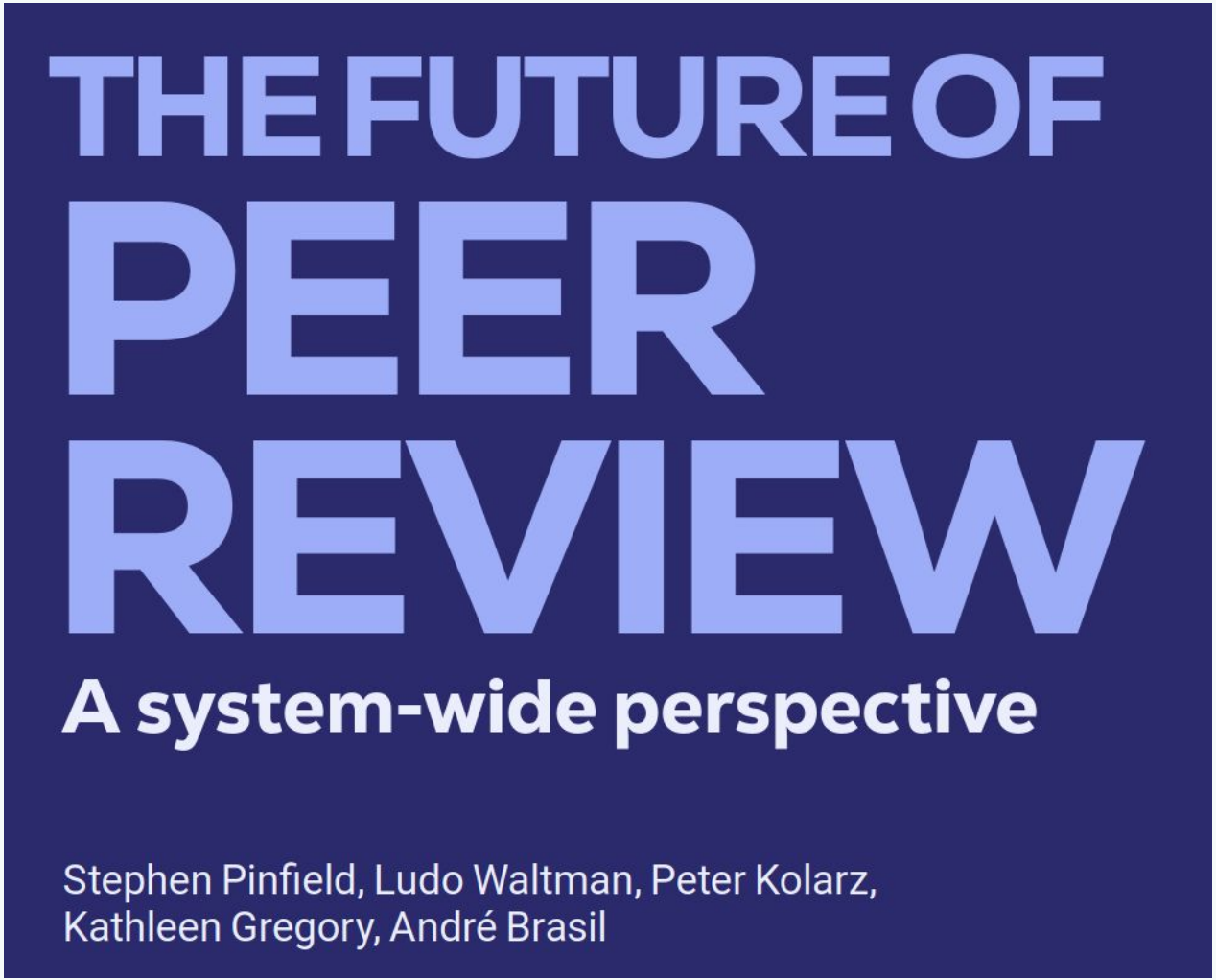
Upcoming reports/papers:

- RoRI Working Paper II: DPR at the Volkswagen Foundation
- DPR as an insight into peer review
- Automated analysis of review comments

Next:

- Comprehensive overview of DPR
  - When, why and how (not) to use
  - Informed by the Volkswagen Foundation experiments

These slides:
bit.ly/tomstafford



The Future of Peer Review:
A system-wide perspective
https://doi.org/10.6084/m9.figshare.32685351

RESEARCH ON RESEARCH INSTITUTE

Thank you

These slides:
bit.ly/tomstafford

Tom Stafford
t.stafford@researchonresearch.org

Join our Substack to get the latest news and views from our researchers