technopolis
group

June 2023

Review of Peer Review

Final report

Peter Kolarz, Anete Vingre, Aaron Vinnik, Antonio Neto, Carlos Vergara, Claudia Obando Rodriguez, Kalle Nielsen, Laura Sutinen

www.technopolis-group.com

June 2023

Review of Peer Review

Final report

Table of Contents

Tables

Figures

Executive summary

This report presents the findings of a study commissioned by UKRI to review interventions to the peer review processes used in R&I award funding. It is intended as a resource for R&I funders across the globe looking to optimise and innovate in their award-making processes.

The study assessed 38 interventions, which range from small process 'tweaks' such as increasing/decreasing the number of reviewers per application and shortening application sections, to more fundamental changes such as partial randomisation and complete bypass of peer review.

The aim of the study was to assess these 38 interventions, to establish what each of them might be useful for, and what disadvantages or hazards each might entail. We also provide an assessment of the overall strength of evidence on each intervention, i.e. which ones are well-studied and which ones are not.

Our research is underpinned by an extensive literature review (encompassing both academic and 'grey' literature), a survey of UKRI staff, and a programme of 22 interviews with representatives of UK and international research funders and a range of other stakeholders and experts in the field.

We find that all interventions we considered here are typically intended to fulfil at least one (or sometimes several) of the following seven aims:

- To save time, i.e. to speed up time-to-grant
- To optimise the relevance of applications and funded awards to the aims of the funding scheme
- To increase the ability to identify and fund high-risk / high-reward projects (sometimes known as 'frontier', 'transformative' or 'breakthrough' research)
- To reduce burden (on applicants, reviewers, panellists and/or administrators)
- To manage application volume (often a subset of reducing burden, but may also occur for other reasons)
- To reduce bias and ensure greater inclusion of disadvantaged groups, including along lines of gender, career stage, institution, or any other category
- To improve the overall quality of reviews (for instance, by ensuring optimally tailored expertise of reviewers or increased levels of transparency and feedback)

These seven aims correspond well to the known challenges of peer review documented in the 'science of science' literature. Almost all the 38 interventions considered in our review provide opportunities to fulfil the above aims.

At the same time, no intervention is a catch-all solution: none pertain to all seven aims, most are useful for certain contexts and less useful (or even problematic) in others, and almost all may entail some form of disadvantage or hazard. Few of these are insurmountable. Recent work in the UK and beyond to reduce research bureaucracy and improve research culture may help create conditions where many such hazards

can be overcome more easily. Modernised IT systems are also a prerequisite for the implementation of many interventions considered here. Often, the disadvantages of one intervention can also be offset by introducing an additional intervention. Not least for this reason, we often identify two or more interventions that typically are used together.

Our study highlights that there is a critical need to coordinate use of the interventions with the context and aims of each specific funding opportunity in question. Based on our findings, creating bespoke funding processes tailored to the needs and aims of each funding opportunity is a clear 'direction of travel' for the future of R&I funding.

We find a mixed picture when it comes to strength of evidence. For some interventions, there is plenty of evidence including experimental studies and quantified outcomes (e.g. for applicant anonymisation, 2-stage application processes, and use of non-academic reviewers), while others appear to be under-researched (e.g. group review and moderation panels). We therefore recommend that funders continue to evaluate and monitor any interventions they use and share findings with other funding organisations.

Our headline recommendation is that process design should always be a constituent part of scheme design. Every funding scheme has specific aims and characteristics, and so the design of the application, review and decision-making process should be considered for each individual funding opportunity.

We encourage funders to make extensive use of the interventions studied here and to vary their assessment processes widely. Some interventions (e.g. peer review colleges, automation-assisted reviewer allocation) even have potential to be mainstreamed across funders' entire portfolios.

We set out our full list of recommendations in the final section of this report.

1 Introduction

This report presents the findings of a study commissioned by UKRI on the use and effectiveness of interventions in peer review for grant-making processes. The study has been carried out by Technopolis from January to March 2023. The intention of this study is to act as a resource for all R&I funders across the globe.

The term 'interventions' is a catch-all word that encompasses the many different organisational and procedural refinements to the baseline application assessment process used by R&I funders across the globe involving external peer review and panel review. We provide a generalised sketch of this baseline process below. The figure below is not intended as a representation specifically of UKRI processes, but a generic heuristic of how research and innovation award funding decisions are typically made world-wide. Of the multitude of UKRI funding opportunities, those under the umbrella of 'responsive mode' funding tend to approximate most closely to the figure below.

Figure 1   The baseline application assessment process in R&I funding
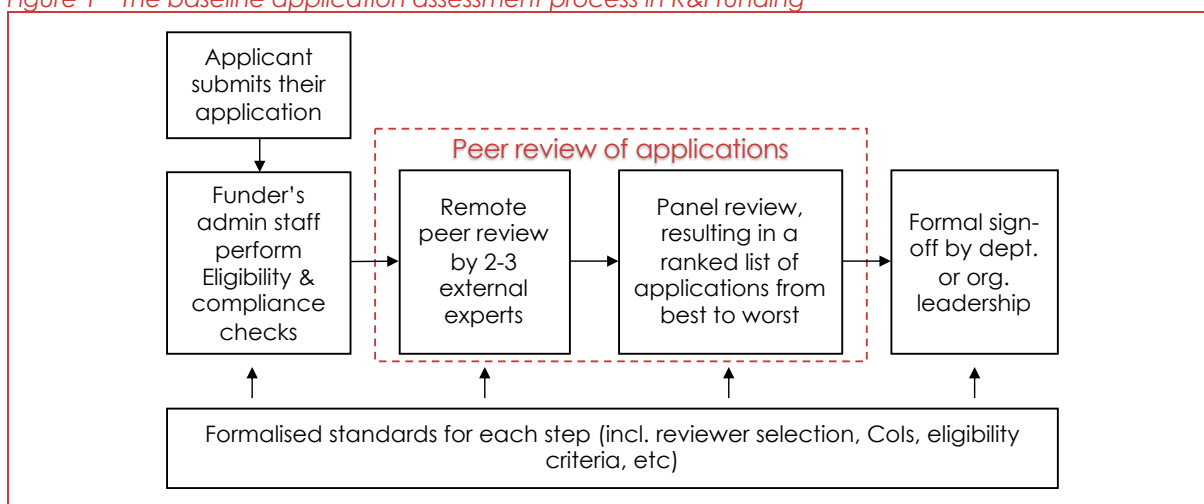
Peer review is trusted by researchers and research funders across the globe. Notwithstanding numerous advances in assessment techniques and technologies, it remains the primary means of R&I award selection. There is a large literature characterising peer review and exploring its strengths and weaknesses, which is being added to continuously for different domains and different potential solutions. Key issues with peer review include:

- It can be burdensome and time-consuming for researchers, reviewers and funders
- It tends to produce conservative decisions, avoiding risk and novelty
- It struggles to suitably assess and reward interdisciplinary research
- It can be biased in favour of established names and institutions, and there is some evidence of gender bias
- Fine-grained rankings of proposals can be influenced by reviewer choice
- It is underused as a developmental tool (e.g., investing sufficiently in feedback that has sufficient depth and quality to improve applicants' future work)

Resulting in large part from these challenges, many funders have introduced various interventions to modify and deviate from the baseline. Some change drivers are 'proactive', meaning that they signal funders' expanded remit or new strategic ambitions, e.g. to fund

research to address societal needs, to fund high-risk/high reward research, or to fund research at speed to respond to an emergency. But there is also a 'reactive' side to change drivers: there are problems with traditional R&I funding assessment processes, including the peer review burden and the risk of bias as outlined above.

1.1   Interventions covered in this review

Different interventions are intended to respond to different drivers. The key drivers for deviating from a baseline assessment process will vary depending on the aims and objectives of a given funding scheme. For example, interventions aimed at speeding up the assessment process will be important in an emergency-response funding scheme, but less so (or not at all) for long-term investments.

Further, interventions may pertain to different parts of the funding process. We distinguish between interventions at the pre-call stage, those pertaining to design of the application itself, the design of the assessment process, and the final decision-making stage. Finally, there are training or feedback interventions underpinning the entire process, so we posit this as an additional category of interventions.

In collaboration with UKRI, we compiled a list of 38 interventions to the baseline research award funding process.[1] This list forms the basis of our review.

Table 1   List of 38 interventions to the baseline peer review process

	Category	Intervention	Description
1	Pre-call	**Assessment criteria definition**	Adding assessment criteria additional to conventional ones, may involve a tiered system for assessment criteria for example, essential vs. desirable.
2	Pre-call	**Demand management: individuals (1)**	Limiting researchers to being a lead investigator only on one project or application at a time.
3	Pre-call	**Demand management: individuals (2)**	Having a 'time out' period of a year, so that after an unsuccessful application, the applicant is not allowed to apply the following year. Based on previous behaviour and includes an element of quality control
4	Pre-call	**Demand management: institutions**	Limiting the number of applications or re-submissions accepted from a single Institution.
5	Pre-call	**Working with underrepresented groups**	Providing additional support to groups that are unrepresented in the funder's portfolio to encourage them to apply and support them as they do, with the view to increasing diversity.
6	Application-design and parameters	**Applicant behaviours**	Designing application forms and processes with a view to encouraging positive behaviours among applicants (e.g. removing hierarchies of applicants to encourage consortium building and collaboration).

---

[1] The list began with a preliminary list of 29 interventions, which was included by UKRI in the terms of reference for this study. Based on our own experience of evaluating R&I funding schemes across the globe, as well as on studies we recently conducted on peer review processes in general (including for UKRI, Wellcome, Formas (Sweden) and the Global Research Council (GRC), we added to this list, and also split or combined various interventions from the preliminary list. Further consultation led to the final list of 38 interventions. We kept open the possibility to include additional interventions if we identified any interesting additional ones during our research. We summarise additional interventions in section 0.

	Category	Intervention	Description
7	Application-design and parameters	**Expression of interest/pre-proposal**	A reduced application is submitted in an expression of interest phase (may simply be a short project description and CV) and triage occurs before a subset are invited to submit a full proposal. See also 2-stage application process.
8	Application-design and parameters	**Reducing applications length/cutting sections**	Shortening application forms (page/word length) to reduce burden. Requiring only project description and not track record, or cutting other sections.
9	Process design	**'Sandpits'/Matching events**	Potential applicants are invited to an event to discuss possibilities and form teams for potential proposals. May involve some application-submission on the day.
10	Process design	**2-stage application process**	Two 'rounds' of peer or panel review are used, sifting out some after the first stage. May involve different parts of the application being reviewed at different stages, or a pre-proposal/EoI (see above).[2]
11	Process design	**Applicant anonymisation**	Reviewers or panels members or both do not see the identity of the applicant/s.
12	Process design	**Automation-assisted reviewer allocation**	Using algorithms/ AI / text recognition to aid allocation of reviewers to applications.
13	Process design	**Dragon's den-style pitch[3]**	Applicants are invited to pitch their proposal in front of a panel, and panels have an opportunity to ask questions. This differs from an interview in that no other form of evidence (e.g. written proposals or external expert review) is used in the assessment.
14	Process design	**External review only (no panel)**	Proposals are only assessed by external reviewers and review scores are simply combined to give the final score.
15	Process design	**Group review**	The same reviewer comments on multiple proposals.
16	Process design	**Changing the number of reviewers**	2-3 external reviews of applications is typical for responsive-mode grant funding, but this number may be lowered to 1 or significantly increased.
17	Process design	**Interviews**	Lead applicant (or several application team members) may do a presentation (optional) and are then asked questions on their application by panel members, reviewers or funder representatives.
18	Process design	**Moderation of reviews**	Reviews are processed internally by funding organisation staff and are only passed to the external panel if they are of sufficient quality.
19	Process design	**Moderation panel**	Assessment panels[4] use external reviews alongside their own expertise to assess the proposal. Moderation

---

[2] We note that the recent UK Research Bureaucracy Review uses the term differently. However, we opt here for a definition that most international funders would recognise in this form.

[3] Note for non-UK readers: the term 'Dragon's Den' originated from a UK TV show involving pitching of business ideas to investors.

[4] Compared to moderation panels, assessment panel members can bring in their own expertise and this approach is mostly part of the baseline process and therefore not considered as an intervention in this study.

	Category	Intervention	Description
			panels do not use their own expertise but can only use the reviews to inform their scores.[5]
20	Process design	**Panel only (no postal/external review)**	Proposals are only assessed by a panel of experts.
21	Process design	**Peer allocation**	The applicants are also the assessors and review the proposals they are competing against to decide who gets funding.
22	Process design	**Programme manager's discretion**	Applications go directly to the programme/scheme manager, who can recommend funding or even decide to fund unilaterally. Usually involves complete by-pass of peer and panel review.
23	Process design	**Standing panels vs. portfolio panels**	Standing panels are the same year on year (with some replacement due to retirement from the panel). Portfolio panels are assembled based on the proposals received and therefore will be comprised differently in each round of funding.
24	Process design	**Use of international assessors**	Having quotas for assessors based in countries other than the funder's 'home' country. May extend to mandating all-international panels and/or reviewers.
25	Process design	**Use of metrics**	Use of metrics and bibliometrics as part of the evidence-base to inform decision-making.
26	Process design	**Use of non-academic assessors (i.e. industry, policy & practice, patients, 'user' representatives)**	Having quotas for non-academic assessors. May extend to all-user panels and/or reviewers. May take the shape of consultation rather than directly making formal funding recommendations.
27	Process design	**Virtual panels**	Convening panels online rather than in person.
28	Decision-making	**Wildcard**	Sometimes also known as 'Golden ticket' or 'Joker'. Each panel member (or other decision-maker) is able to select one proposal (e.g. per call, per year, or similar) to guarantee funding (provided there is no conflict of interest), regardless of panel rankings or other decision-making processes.
29	Decision-making	**Partial randomisation**	Successful proposals are chosen at random. In most methodologies, randomisation is only partial. For example, proposals may be scored and sorted into bands, and only those on the border of being funded will be randomised.
30	Decision-making	**Scoring mechanisms**	Includes calibration of scores, consensus vs. voting, weighting.
31	Decision-making	**Sequential application of criteria (rather than simultaneous application of criteria)**	A proposal is scored for one set of criteria, ranked and a cut-off point determined. Then those above the cut-off point are assessed again for another set of criteria to determine the final funded list.

---

[5] Note that to ensure clarity for the widest possible readership, we are using terminology that might not align with UKRI terminology. In UKRI some moderation panels can bring in generic/system expertise.

	Category	Intervention	Description
32	Decision-making	**Use of quotas**	After ranking, proposals are reviewed to ensure sufficient numbers in certain categories including quotas related to protected characteristics, place, first-time applicants, etc.
33	Training and feedback	**Bringing in reviewers from earlier careers & providing mentoring**	Panels and reviewers tend to be very experienced researchers/innovators. Those early in their careers could be invited to review or be part of panels with additional training, bringing different perspectives and experiences. Previous calls' award winners may also be brought in as reviewers/panellists.
34	Training and feedback	**Embedding EDI in assessment**	Training or support provided to make assessors aware of their unconscious biases and to encourage them to call each other out during the assessment process.
35	Training and feedback	**Expanding or reducing the amount/detail of feedback to unsuccessful applicants**	Different levels of feedback may be provided on unsuccessful applications.
36	Training and feedback	**Funder representation on review panels**	The funder is represented on the panel to guide discussion or provide briefing on programme aims. Their role is beyond a purely administrative function, they may even be in a chair-role or similar.
37	Training and feedback	**Improving quality of reviews**	Through training/retaining good reviewers/recognition. Peer review colleges fit here too.
38	Training and feedback	**Open review/rebuttal**	Reviews are published and/or made available to the applicant before funding decisions are taken, so they can be viewed and responded to.

1.2   Method note

For each of the 38 interventions, we set out to compile an evidence base to establish the following points:

- **Definition(s):** what exactly does the intervention involve? Are there relevant differences in how different funders practise the intervention?

- **Why to do it:** what is the envisaged benefit of the intervention? What problems/issues is it supposed to solve? What, therefore, might be measures of its success?

- **Why not to do it:** does the intervention have any weaknesses, hazards or drawbacks? Are these especially problematic under certain circumstances (i.e. for particular scheme types)?

- **Evidence verdict** and strength of evidence: is there evidence to show that this intervention has (or has not) worked? What is the strength of the evidence (e.g. controlled experiments, light-touch evaluation, anecdotal)?

Our study had three data collection strands, which ran in parallel.

First, we conducted a review of literature on research award funding processes. This included academic literature as well as evidence from evaluations of various funding schemes and wider strategic studies. We conducted keyword searches for each intervention and also added resources known to us prior to the study.

We further conducted a consultation survey of UKRI staff. This survey was primarily intended to obtain views on any of our 38 interventions that may have been trialled in different parts of the

organisation, including comments on interventions that worked well and interventions that did not. This information adds additional stakeholder perspectives to the findings obtained from the literature. While the survey cannot be fully representative, we added some survey items that help quantify which interventions appear to be well known or less well known in different parts of UKRI, and whether there is particular appetite for certain interventions to be used more.

Finally, we ran a programme of interviews to obtain further viewpoints on the 38 interventions. We included in the programme a small number of follow-up conversations with UKRI survey respondents, as well as several representatives from UK funders other than UKRI, international funders, stakeholders from the UK HEI landscape, and a selection of academic experts on peer review, its modifications and alternatives.

Method details are presented in the appendices to this report.

1.3   Structure of this report

In the next section, we present some aggregate findings and general observations from our research. In the subsequent five main sections we present the evidence on each of the 38 interventions, split by our five intervention domains: interventions at the pre-call stage, interventions pertaining to design of the application itself, the design of the assessment process, the final decision-making stage, and training or feedback interventions underpinning the entire process. The findings in these five main sections are aggregate summaries from our literature review (see Appendix A), our survey (see Appendix B) and our interviews (see Appendix C).

For each intervention, we provide a write up explaining its aims, some data highlights (i.e. instances of use) and any known effects, as well as hazards or dangers associated with each intervention. For each intervention, we also provide a simple rating of the evidence strength. This rating relates only to evidence strength, not to intervention effectiveness: it does not reflect whether the intervention works, but the strength of evidence demonstrating its efficacy (or lack thereof, as the case may be):

- One star (*): very limited evidence, almost or entirely tentative or speculative
- Two stars (**): some evidence, e.g. several anecdotal pieces and perhaps some minor empirical observations
- Three stars (***): multiple sources of credible evidence, though not necessarily quantifiable conclusions, and not all parts of the intervention have been investigated thoroughly (e.g. in cases of multiple aims)
- Four stars (****): multiple sources of credible evidence, including experimental or other empirical measurement/evaluation

Finally, we provide a brief overview of a small number of other minor interventions not included in our initial list of 38, but which were discovered by the research team over the course of the study. The last section of this report provides a summary table of findings and our list of recommendations resulting from our research. We note that these recommendations are not specific to UKRI but may be considered by any R&I funder looking to optimise and innovate in their award-making processes.

2　General observations

The next main section presents findings on each of the 38 interventions and forms the bulk of this report. However, there are some general observations worth noting at the outset.

First, we find that the rationales for the interventions (as expressed in the literature and by consultees) correspond well to the problems of peer review and the 'baseline' funding process noted at the outset. Almost all interventions considered here draw their rationale from the following seven (partially related) aims:

- To **save time**, i.e. to speed up time-to grant, either from a simple efficiency point of view, or in order to be able to respond to emergencies

- To **optimise the relevance** of applications and funded awards to the aims of the funding scheme (e.g. thematic or sector relevance, maximum scope for application)

- To **increase the ability to identify and fund high-risk / high-reward projects** (also known as 'transformative', 'radical', 'frontier' or 'breakthrough' research). In a sense this is a subset of the above aim of optimising relevance (if a scheme specifically aims to fund such research) but it relates to the well-documented conservatism of peer review, which is an issue in its own right

- To **reduce burden** on applicants, reviewers, panellists and/or administrators. This is generally about efficiencies and minimising the effort and cost needed to carry out the review of applications

- To **manage application volume**. This may to an extent relate to reducing burden more generally, but also relates to discouraging applications that are out-of-scope or of unsuitably low quality

- To **reduce bias** and ensure greater inclusion of disadvantaged and/or under-represented groups, including along lines of gender, career stage, institution, or any other category

- To **improve the overall quality of reviews**. This may mean, for instance, to ensure optimally tailored expertise of reviewers, as well as increased levels of transparency and feedback

Almost without exception, every literature source, interviewee and survey respondent cites at least one of the above seven aims when discussing any of our 38 interventions. Some interventions relate only to one of these seven aims, though most are associated with several (often two or three). In the concluding section of this report, we provide a comprehensive overview of how our 38 interventions relate to each of the seven aims.[6]

At the same time, no intervention is a catch-all solution: none pertain to all seven aims, most are useful for certain contexts and less useful (or even problematic) in others, and almost all may entail some form of disadvantage or potential hazard.

This means that there is a critical need to coordinate use of the interventions with the context and aims of the specific funding opportunity in question. Creating bespoke funding processes tailored to the needs and aims of each funding opportunity is a 'direction of travel' for the future of research funding.

---

[6] While this list of intervention-aims is foremost intended to help funders decide when and why to introduce each intervention, we note that it may also be a useful tool to secure buy-in from the research community: the issue of buy-in is not covered in detail in any of the sources we consulted, but our research indicates that wider buy-in is a concern that funders occasionally have when contemplating introducing interventions. A clear rationale for introduction which draws on this list of possible aims may contribute towards mitigating such concerns.

The table below shows how many interventions out of the 38 pertain to each of the seven aims.

Table 2   Frequency of seven main aims among the 38 interventions

Aim	Frequency among the 38 interventions
Save time	9
Optimise the relevance of applications	11
Increase the ability to identify and fund high-risk / high-reward projects	7
Reduce burden	11
Manage application volume	3
Reduce bias	13
Improve the overall quality of reviews	17

Regarding the disadvantages and hazards of each intervention, few are insurmountable. We do not provide a full assessment of how easily each hazard or disadvantage may be overcome, in part because this is often context dependent: some funders' IT systems may for instance be readily able to address many noted challenges. Some hazards may be more severe in smaller research systems (be they delineated by country or research field) where conflicts of interest are more likely to occur. We note hazards and disadvantages where we identify them, but it will most often be dependent on each funder's context whether they constitute a 'showstopper' or whether they can be dealt with. We also note mitigations where these are evident from our research.

We also note that several interventions may complement each other and may often appear together. For instance, wildcard approaches tend to be used in combination with anonymised reviewing in order to minimise scope for cronyism. In such cases, a hazard associated with one intervention is mitigated by adding another intervention.

On the other hand, some interventions might also counteract each other: an intervention may increase the quality of applications but might entail additional burden and/or lengthen time-to-grant. Others may do the opposite. Pairing/combining different interventions does not appear to be a well-researched topic, but we have identified some common pairings and rationales for pairing certain interventions together (noted where relevant in the following main section of this report).

A few additional meta findings are as follows:

- **Positives in the peer review must be recognised**: several survey respondents and interviewees felt the need to emphasise the good things about the peer review system (including the 'baseline' approach). In light of the overall criticisms, consultees often judge it essential to praise the work that is globally invested in the peer review effort and the benefits it brings. Funder staff, having regularly observed panels in action, feel it is sometimes unfair to showcase only examples of failure and ignore the positives like the effort invested, the care that reviewers put into the activity and the benefits a good review brings to the research community

- **Political sensitivities and acceptance of the interventions in the research community**: some interventions, primarily partial randomisation, but also the use of quotas, demand management, interviews, and sandpits, have sometimes raised concerns in the research community or at the research funders' boards or oversight institutions. All consulted funders

that have introduced partial randomisation report investing extra effort to make a case for the funder's board or ministries that oversee their operations, regardless of geography. Private funders might be less concerned about external pressures but still have to make a solid internal case.

Some consulted funders and experts pointed out that the acceptance from the oversight bodies and wider society is a more significant concern than the acceptance in the research community itself. This is mainly because researchers are more familiar with the peer review system, its strengths and weaknesses and understand the rationale of the more experimental interventions. In some cases, the significant scrutiny and risks result in a reluctance to try new things. However, a degree of scepticism is certainly warranted: demand management, interviews, sandpits and some other interventions do raise concerns about, for instance, equity and the potential favouring of applicants who can access specific meetings and events and have good presentation skills

- **Shifting responsibility from funders to research performing organisations**: there is some tension between the responsibilities of funders and research performing organisations in addressing the equity and burden in research funding. Some of the interventions may mean less burden for the funders but more for the research administrators at the research performing organisations. Where demand management is transferred to the institutional level, research performing organisations may effectively carry out the assessment for the best application to put forward. There are also examples of funders removing some requirements (e.g. specific sections in applications or monitoring requirements). However, those are still implemented/asked for at the research performing organisations to maintain internal oversight. As a result, nothing changes regarding the burden for the 'regular' researcher and for the research system as a whole

- **Manuscript peer review**: in the academic literature, some interventions are discussed primarily or only in the context of manuscript peer review (i.e. for journal publication). Literature and our expert consultation show that journals frequently experiment with new interventions and assess the results of the experiments. Examples include efforts to improve the review process through open peer review (making reviews public) and improving the quality and reliability of peer review through training reviewers. Literature on manuscript peer review shows improvements in review reliability in terms of identification of errors/recommending manuscripts for rejection after the introduction of reviewer training.

Although not without controversy (e.g. concerns about less critical comments if the review is open), manuscript peer review may provide examples worth considering for research funders, even though it is beyond the scope of this review. Another example is scientific publishers reacting quickly and introducing rules and guidance to specify the use of large language models (AI algorithms like ChatGPT) in manuscript preparation and review process. Several consulted research funders were concerned about the impact of large language models on research funding processes. Examples of addressing the matter in manuscript peer review might be worth considering. It must also be noted that grant peer review processes like interviews and panels make it more complicated and difficult to compare to journal peer review. Grant peer review and journal peer review happen at different stages of the research process. Journal peer review looks at completed work, while grant peer review looks at proposed work.

3 Main findings: Interventions prior to a call

3.1 Assessment criteria definition

Adding assessment criteria additional to conventional ones, may involve a tiered system for assessment criteria for example, essential vs. desirable.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase relevance of funded projects to call aims	Reviewers may not follow guidance; too many criteria risk over-complicating discussions	***

This intervention may include: clear guidance with definitions of criteria, non-biased language (e.g. gender) and weighting of criteria, and ensuring criteria are suitably discussed and applied during panel meetings.

The aim is to make sure that proposals are assessed according to the intended criteria, and therefore to the aims of the call. Emphasis is on increasing transparency, consistency, simplification, as well as the need to ensure that the selection reflects the objectives of the specific funding scheme (especially when including new criteria that might be under-valued).

Impact evaluation of one scheme[7] shows the effectiveness of the intervention in supporting projects aimed at achieving non-academic impact. However, it cannot be attributed solely to the criteria. Funders have observed that this approach meant they funded projects that went on to have impact, and that these would not have been funded if the assessment was purely based on an assessment of research quality.

Several authors appear to agree that more explicit criteria are desirable to avoid bias and inconsistency. However, the evidence also highlights a perception that criteria that go beyond research excellence can still be challenging for reviewers and/or panellists to apply. There is also a limit to the number and complexity of criteria that panels can handle.

Further, behaviour of reviewers does not necessarily conform to guidance. Evidence suggests that external reviewers pay more attention to written guidance than panel members.

References	Interviewees & survey responses
• 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). https://doi.org/10.1371/journal.pone.0046054 • Hug, S. E., & Aeschbach, M. (2020). Criteria for assessing grant applications: a systematic review. Palgrave Communications, 6(1), 1–15. https://doi.org/10.1057/s41599-020-0412-9 • Hug, S. E., & Ochsner, M. (2022). Do peers share the same criteria for assessing grant applications? Research Evaluation, 31(1), 104–117. https://doi.org/10.1093/reseval/rvab034 • Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. Social Studies of Science, 31(6), 820–841. https://doi.org/10.1177/030631201031006002 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme, European Commission, https://doi.org/10.2777/16211 • Tamblyn, R., Girard, N., Qian, C. J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. Cmaj, 190(16), E489–E499. https://doi.org/10.1503/cmaj.170901	• 2 survey responses • 2 interviews

---

[7] In a small number of cases, consultees were hesitant about publicly sharing certain examples. This is one such case.

3.2   Demand management: Individuals (1)

Limiting researchers to being a lead investigator only on one project or application at a time.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce application numbers and concentration of awards	Shifts burden to other funders, savings are minimal	*

Many funders limit applicants to one application per call. However, this may be expanded to one application across the funders' entire portfolio. This intervention is intended to reduce the number of applications (by limiting / excluding the participation of current awardees). There may also be a motivation to limit the concentration of awards to a small number of continuously successful researchers.

This intervention is rare, not least as it requires a comprehensive research information system (preferably covering multiple funders so that applications cannot be re-submitted to other funders instead). There is ongoing use at the Swedish Research Council, though no assessments or feedback could be identified.

Our research also highlights sceptical views around this intervention; the Royal Society has stated that it does not support disincentives to apply for funding in the first place (though this statement is from 2007), and a Rand report concluded that savings gained from individual-targeting restrictions were marginal if proposals became complex as a result, and recommended institutional quotas for more substantive savings. As noted above, re-submission to other funders is a risk, so burden and application-influx is shifted rather than lessened.

Most UK-based evidence we find is from around 2006-07 after the publication of a peer review report by Research Councils UK (RCUK, UKRI's precursor). It does not appear to be a heavily studied intervention or one for which there is much 'appetite'.

References	Interviewees & survey responses
• Grove, L. (2017). The effects of funding policies on academic research. Doctoral thesis, University College London. http://eprints.lse.ac.uk/88207/ • Ismail, S., Farrands, A., & Wooding, S. (2009). Evaluating Grant Peer Review in the Health Sciences: A review of the literature. RAND Europe. • Research Councils UK. (2006). Report of the Research Councils UK Efficiency and Effectiveness of Peer Review Project. http://www.rcuk.ac.uk/documents/documents/rcukprreport.pdf • Royal Society. (2007). Response to the RCUK consultation on the Efficiency and Effectiveness of Peer Review. RS Policy Document 04/07. https://royalsociety.org/~/media/royal_society_content/policy/publications/2007/8090.pdf • Swedish Research Council. (n.d.). "Several grants simultaneously". Retrieved March 1, 2023, from https://www.vr.se/english/applying-for-funding/applying-for-a-grant/several-grants-simultaneously.html • Swedish Research Council. (n.d.). "What requirements apply if I already have a grant from the Swedish Research Council". Retrieved March 1, 2023, from What requirements apply if I already have a grant from the Swedish Research Council?	• No survey responses or interviews

3.3 Demand management: individuals (2)

Having a 'time out' period of a year, so that after an unsuccessful application, the applicant is not allowed to apply the following year. Based on previous behaviour and includes an element of quality control		
Main intended aim(s)	Main hazard(s)	Evidence strength
Limit application-volume, increase success rates, save burden	May simply shift re-submission to other funders, may not be well received by applicant community	**

This intervention aims to control application-based demand (e.g. application volume and overall success rates) and reduce the workload for funders and reviewers.

The EPSRC already operates a variant of this intervention, whereby any investigator that is repeatedly unsuccessful in the preceding two years will be written to and limited to just one application in the following 12 months.

This approach has received positive feedback from reviewers and senior university personnel. Paired with the approach to ban identical resubmissions, has been found to reduce application volumes.

Some comments note that researchers suffering from biases may be put in an increasingly disadvantaged position, and that this approach may damage individuals' confidence, experience, career, and well-being, though this has not been studied (logically though, it appears plausible that this approach would penalise at least some potential applicants).

In the absence of comprehensive international research information systems, it is also impossible to control for researchers re-submitting applications to other funders. While this approach therefore limits burden for the funder in question, it is unlikely to lead to burden-reduction in the wider research system.

Our research indicates that introducing this intervention has occasionally been controversial: generally, it appears to divide researchers (especially early career researchers) on the one hand, and funders / senior personnel on the other, who typically view the approach fairly positively. However, this assessment is based on various pieces of anecdotal evidence only.

References	Interviewees & survey responses
- Bhattacharya, A. (2012). Science funding: Duel to the death. Social Forces, 488, 20–52. https://doi.org/10.1093/sf/4.1.189 - Huntoon, L. R. (2019). Sham Peer Review : the Destruction of Medical Careers. Journal of American Physicians and Surgeons, 24(4), 99–100. https://jpands.org/vol24no4/huntoon.pdf - Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: a realist synthesis. Research Integrity and Peer Review, 7(1), 1–28. https://doi.org/10.1186/s41073-022-00120-2 - Tough love. (2010). Nature, 464(7288), 465. https://doi.org/10.1038/464465a	2 survey responses 4 interviews

3.4 Demand management: institutions

Limiting the number of applications or re-submissions accepted from a single Institution.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Limit application-volume, increase success rates, save burden in the funder and reviewer community	Largely shifts burden to institutions; potential additional bias, depending on institutional processes	****

This intervention aims to reduce workload and administrative burden for the funder and reviewer community. Key indicators would be the number of applications received (lower than without this intervention) and a higher application success rate.

This intervention is known to accomplish what it intends: variations are in continued use in multiple funding organisations in US and Europe. NERC limits the number of applications per institution where the HEI in question has failed to meet a 20% success rate over the 6 most recent grant rounds; the European Society for Paediatric Research allows an unlimited number of applications, but only awards one per institution; National Institutes of Health (NIH) allows two applications per institution in the Director's Early Independence Awards; ESRC allows a limited number of applications per institution for its Research Centres competition (alongside the use of outline proposals); the US National Science Foundation (NSF) allows three expressions of interest (EoIs) per institution, of which a maximum of one can result in an invite to submit a full proposal.

The National Natural Science Foundation of China (NSFC) has a different version of demand management. The 2011 evaluation of the NSFC noted that applicants submit their proposals via their host institution and they may not submit them directly to NSFC. An applicant may not apply more than once per year to any single NSFC programme or hold more than 3 NSFC grants at the same time (this example is also relevant to the previous section).

This approach is known to have reduced the number of applications in cases of schemes that had previous funding rounds without the intervention.

A major problem with this intervention consistently mentioned throughout the evidence is that it largely shifts selection burden from the funder to the institution. The institution may opt for a more limited reviewing procedure, thus still reducing overall burden to some extent. However, there is also some anecdotal evidence that institutions may be less experienced in some aspects of selection processes, leading to sub-optimal outcomes.

This intervention is not passionately debated in one way or the other: most sources agree on strengths and hazards on this approach, however there are naturally conflicting interests between funders and research performing organisations on this. As a practice, it is in use in multiple contexts, though it appears to be quite commonplace in the USA especially, particularly in calls aimed at early career reseachers. We also note this approach has often been found paired with use of 'expressions of interest'.

References	Interviewees & survey responses
• Arnold, E., Lu, Y., & Xue, L. (2011). International evaluation of funding and management of the National Natural Science Foundation of China. Technopolis Group. http://www.technopolis-group.com/resources/downloads/reports/nsfc_evaluation_report.pdf • ESPR. (2022). ESPR Research Grant Programme 2017-2022. European Society for Paediatric Research. https://www.espr.eu/funding/research_grant_programme.php • Mycroft, C. (n.d.). ESRC – Large Grants Competition 2016/17 (Update). University of Lincoln: Research Blog. Retrieved February 20, 2023, from https://research.blogs.lincoln.ac.uk/2016/08/03/esrc-large-grants-competition-201617-update/ • UKRI. (2022). Demand management. NERC. https://www.ukri.org/councils/nerc/guidance-for-applicants/types-of-funding-we-offer/discovery-science/demand-management/ • de Winde, C. M., Sarabipour, S., Carignano, H., Davla, S., Eccles, D., Hainer, S. J., Haidar, M., Ilangovan, V., Jadavji, N. M., Kritsiligkou, P., Lee, T.-Y., & Ólafsdóttir, H. F. (2021). Towards inclusive funding practices for early career researchers. Journal of Science Policy & Governance, 18(01). https://doi.org/10.38126/jspg180105	• 1 survey response • No interviews

3.5   Working with underrepresented groups

Providing additional support to groups that are unrepresented in the funder's portfolio to encourage them to apply and support them as they do, with the view to increasing diversity.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase diversity of applicants and award winners	May take some time to show effect; may entail administrative burden	****

This intervention intends to increase the number of applicants (and their success rate when applying for an award) of underrepresented groups, e.g. ethnic minority groups or younger / early career researchers.

The AHRC's 2020-22 Equality, Diversity and Inclusion Engagement Fellowship (EDIEF) pilot specifically targeted arts and humanities researchers whose work has a significant Equality, Diversity and Inclusion dimension. The call sought to enable researchers to engage a variety of relevant stakeholders with their research, to embed their work into policy and practice, and to work with relevant communities to realise the full potential benefits of their research. The intervention emerged as a response to previous studies identifying barriers to collaborative research partnerships with the minority ethnic communities (Common Cause research) and a commitment to improving EDI. An evaluation of the pilot showed that 28% of applicants were Asian, Black or mixed ethnicity, whilst only 9% of applicants to the standard research grant scheme were from an ethnic minority.

The UK Equality Challenge Unit (ECU) launched the Athena SWAN charter in 2005 to recognise universities' work to improve gender equality and diversity of women in science, technology, engineering, medicine, and mathematics. As a voluntary action, universities are not set with goals, but are instead encouraged to assess their current gender gaps and adopt measures to reduce disparities. The Athena SWAN Charter offers different levels of accreditation (bronze, silver and gold) to universities depending on the type of interventions and strategies adopted to alleviate gender gaps. Universities need to gain Athena SWAN Charter membership first to apply for accreditation. For the Bronze accreditation, universities undertake an assessment of gender disparities and propose a 5-year plan to address this. Silver recognition requires the implementation of specific actions, while Gold is awarded to those achieving or improving gender parity levels.

In 2011, the National Institute for Health Research (NIHR) linked its research funding for biomedical research centres to actions towards gender equality through the Athena SWAN charter. In 2016, academic institutions must hold at least silver accreditation to be shortlisted for funding. This intervention has led to an increase in women theme leads from 8% in 2006 to 24% in 2016. It may also have contributed to the increase in the number of universities in the field implementing action plans from one in 2011 to 69 in 2016. According to the literature, this intervention has been replicated by funders and science organisations in Ireland, Australia, the US and Canada.

A final example found in the literature is the National Research Mentoring Network delivered by the NIH (US) as part of the 'Diversity Program Consortium'. It is reported that intensive and sustained training of early-career researchers of underrepresented minority groups can help participants to achieve the benchmarks of proposal submission and funding. This mentoring can also have an impact on other areas, such as teaching.

Our interviewees also note that using positive language to encourage women's participation has led to an increase in the number of applications from women.

Action to improve not just the application but the success rate of underrepresented groups is a rather broader issue with many possible techniques. Most notably we come back to this issue when we address anonymised reviewing, as well as various training interventions covered in the latter parts of this report.

However, on a final point here it is worth mentioning efforts to diversify reviewers: in its 2022 Race and Ethnicity Inequity report, the EPSRC noted its action to increase the representation of ethnic minority researchers on its peer review college to 20% by actively encouraging self-nominations to the peer review college from all researchers but particularly seeking nominations from minority ethnic researchers. In the first six months of the campaign, EPSRC observed a positive response with a 2.5x increase in self-nominations compared to the previous year.

There is general agreement on the benefits and effectiveness of the intervention. However, while not noted explicitly in the literature, these actions likely take time, which can limit their applicability. There is also an associated administrative burden, which, according to some sources, may be disproportionately carried out by women.[8]

References	Interviewees & survey responses
• Blackburn, M., Coutinho, K., & Suviste, H. (2022). The Equality, Diversity and Inclusion Engagement Fellowship Pilot AHRC Funding Scheme Report 2020 - 2022. The Open University. http://oro.open.ac.uk/87420/ • EPSRC (2022). Ethnicity and Race Inequity In Our Portfolio: Findings of our community engagement and actions for change. https://www.ukri.org/wp-content/uploads/2022/11/21112022-EPSRC-Ethnicity-and-Race-Inequity-In-Our-Portfolio_Nov-2022.pdf • Gamage, D. D. K., & Sevilla, A. (2019). Gender Equality and Positive Action: Evidence from UK Universities. AEA Papers and Proceedings, 109, 105–109. https://doi.org/10.2139/ssrn.3390198 • Gregory-Smith, I. (2017). Positive action towards gender equality? Evidence from the Athena SWAN Charter in UK Medical Schools. British Journal of Industrial Relations, 56(3), 463–483. • Ovseiko, P. V, Taylor, M., Gilligan, R. E., Birks, J., Elhussein, L., Rogers, M., Tesanovic, S., Hernandez, J., Wells, G., Greenhalgh, T., & Buchan, A. M. (2020). Effect of Athena SWAN funding incentives on women's research leadership. BMJ, 371, m3975. https://www.bmj.com/content/371/bmj.m3975 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211 • Sato, S., Gygax, P.M., Randall, J. *et al.* (2021). The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. *High Educ* 82, 145–162. https://doi.org/10.1007/s10734-020-00626-y • Weber-Main AM, McGee R, Eide Boman K, Hemming J, Hall M, Unold T, et al. (2020) Grant application outcomes for biomedical researchers who participated in the National Research Mentoring Network's Grant Writing Coaching Programs. PLoS ONE 15(11): e0241851. https://doi.org/10.1371/journal.pone.0241851	• No survey responses • 2 interviews

---

[8] We note as a general comment on the evidence that historically, much of the literature around this intervention has focused mostly on gender. However, among the more recent sources we have considered here, race/ethnicity and age also feature quite strongly.

4 Main findings: Interventions in application parameters

4.1 Applicant behaviours

Designing application forms and processes with a view to encouraging positive behaviours among applicants (e.g. removing hierarchies of applicants to encourage consortium building and collaboration).		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase diversity and limit gatekeeping	None known	*

There is very limited research on this intervention and may require an additional specific review or controlled experiments. This intervention aims to ensure active management of consortia composition to refresh the membership of flagship investments. It should enable new participants to be involved without this being controlled solely by the existing incumbents.

EPSRC used this in one of its Quantum Technology Research Hubs calls. The programme had funded specific hubs for five years and launched new funding to support the continued operation of the hubs with slightly adjusted priorities. Previously funded organisations could apply again. However, EPSRC wanted to ensure existing capabilities are used and, at the same time, modify the consortia to meet the new objectives and ensure fairness in allowing new participants. In other words, some organisations might no longer be part, and new ones might join to ensure optimal composition and benefits to consortium members. Therefore, EPSRC ran a peer review process to identify participants and facilitated consortia-forming workshops. The process was resource intensive for EPSRC but resulted in new consortia that met the new funding objectives. EPSRC saved time and effort by not organising an entirely new funding call. EPSRC has organised similar consortia-forming processes also in programmes without previously funded awards.

The limited evidence we have indicates consortia with these rules in place were highly rated by peer reviewers and considered to map well to the priorities for the programme. However, this intervention potentially has many other variations (application forms may be re-designed in different ways) and we find no evidence on the effects of such modifications. It appears to be an under-researched area.

References	**Interviewees & survey responses**
• None.	• One survey response and follow-up interview with the respondent

4.2 Expression of interest/pre-proposal

A reduced application is submitted in an expression of interest phase (may simply be a short project description and CV) and triage occurs before a subset are invited to submit a full proposal. See also 2-stage application process.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce review burden, reduce application burden, increase quality and relevance of full proposals	Longer time-to-grant, influx of out-of-scope EoIs, limits information to inform decision-making	***

Expressions of interest or pre-proposals are used widely to reduce time spent on each individual review and to ensure there are fewer full applications to be reviewed, meaning less burden on the reviewer community. Potentially there is also money saved (in cases where reviewers are paid for their efforts).

Where applicants only submit a full application after success at EoI-stage, there is also a reduced burden for applicants, though this is not necessarily the case (pre-proposals may also simply be part of a full proposal, where only the pre-proposal gets reviewed at first).

To this end, the Dutch Research Council (NWO) announced the introduction of compulsory pre-proposals for its Veni Vidi Vici grants in November 2022. The method was stated to save time for applicants and reviewers alike and to limit application pressure. This update was introduced as a part of larger appetite to shorten assessment procedures at the NWO.

Additional important aims of this intervention are, firstly, to achieve increased relevance and/or quality in the pool of full applications resulting from sifting of pre-proposals/EoIs (see also the intervention on sequential application of criteria), and secondly to broaden the applicant base by attracting new constituencies of applicants through low opportunity cost. Sometimes it has also been used to assess demand/interest in a funding opportunity.

This intervention is in use in many funding schemes across the globe. Several studies and respondents note that the resulting pool of applications is of higher quality than in schemes without a pre-stage. Some evidence suggests reduced burden on reviewers, despite potential for high influx of EoIs/pre-proposals.

For instance, in the UK's Global Challenges Research Fund (GCRF), programmes were found to have a notably higher success (27%), compared to the typical corresponding success rates that can be as low as 10%. Part of this outcome was attributed to expressions of interest / pre-proposal stages in use in some of its calls.

There is also evidence that the intervention is well-received among industry applicants. For example, in the EPSRC Prosperity Partnerships programme that funds an industry-academia co-created research with 50% industry cash contribution, the assessment process includes an outline interview stage (accompanied by a short outline application) which allows for rapid go/no-go decisions. EPSRC introduced the process to encourage business involvement in the research partnership. A process evaluation of the programme found that the two-stage process allowed for interrogation of the partnership before a detailed analysis of the proposed research. The approach helped to sift traditional research projects that did not meet the objectives of the programme.

If a full application only follows after success at the pre-proposal/EoI stage, it lowers the barriers to entry, often resulting in a high influx of pre-proposals/EoIs, sometimes including speculative/out of scope/poor quality submissions. The Human Frontier Science program (HFSP) for instance saw a substantially heightened number of applicants after it introduced an initial pre-proposal phase. As a demand-management tool, effectiveness is therefore only partial when pre-proposals/EoIs are not combined with full applications in a single submission.

Two-stage processes (EoI/pre-proposal plus full application) also tend to take longer than conventional funding processes. Additionally, such short outlines may not provide enough information for reviewers to make a robust judgement.

In short, this is a widely used intervention with many possible advantages and a range of potential hazards, depending on how exactly the intervention is operationalised. While many of these are somewhat well understood, we do not find comprehensive evidence on all of them.

References	Interviewees & survey responses
• Barr, J. Et al. (2018), GCRF Evaluation Foundation Stage: Final Report, URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/810137/GCRF_Evaluation_Foundation_Stage_Final_Report.pdf • Kolarz, P. et al (2023) Organisational and Procress Review of the Human Frontier Science Program. Technopolis Group. Published by HFSP: https://www.hfsp.org/node/74873#book/ • Morgan, B., Yu, L.-M., Solomon, T., & Ziebland, S. (2020). Assessing health research grant applications: A retrospective comparative review of a one-stage versus a two-stage application assessment process. PLoS One, 15(3). • Nous Group. (2018). Consultation on the development of peer review for NHMRC's new grant program - Final Report. https://doi.org/10.1111/j.1834-7819.1982.tb04109.x • NWO (2022), Talent Scheme newsflash: simplified application process and planning 2023, URL: https://www.nwo.nl/en/news/talent-scheme-newsflash-simplified-application-process-and-planning-2023 • "Proceedings of the REvaluation Conference 2021/22". (2022). Fteval, 53. • Royal Society. (2021). Submission to the independent review of research bureaucracy. 15 October 2021. https://royalsociety.org/-/media/policy/Publications/2021/21-10-15-royal-society-response-to-the-independent-review-of-research-bureaucracy-consultation.pdf • Taylor-Robinson, A. W. (2013). Grant preparation time: the law of diminishing returns prompts a rethink on the application and review process. BMJ Open. https://acquire.cqu.edu.au/articles/journal_contribution/Grant_preparation_time_the_law_of_diminishing_returns_prompts_a_rethink_on_the_application_and_review_process/13461869 • UKRI. (2020). Diversity results for UKRI funding data: 2014-15 to 2018-19. UK Research and Innovation. https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-DiversityResultsForUKRIFundingData2014-19.pdf • UKRI. (n.d.). Comparing assessment methods: outline-to-full versus full proposal only processes. • UKRI. (n.d.). What happens after you submit your proposal. UKRI. Retrieved February 20, 2023, from https://www.ukri.org/councils/epsrc/guidance-for-applicants/what-happens-after-you-submit-your-proposal/ • UKRI. (2022). NERC funding investments between 2014 to 2015 and 2019 to 2020. UKRI. https://www.ukri.org/publications/nerc-funding-investments-2014-2015-and-2019-2020/nerc-funding-investments-between-2014-2015-and-2019-2020/ • Vingre, A., Kolarz, P., & Bryan, B. (2022). On Your Marks, Get Set, Fund! Rapid Responses to the Covid-19 Pandemic. Fteval, 53, 8–17. https://doi.org/10.22163/fteval.2022.538	• 6 survey responses • 2 interviews

4.3   Reducing applications length/cutting sections

Shortening application forms (page/word length) to reduce burden. Requiring only project description and not track record, or cutting other sections.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce burden on applicants and reviewers, improve efficiency	Limits information to inform decision-making, may not always save burden for applicants	***

Reducing the length of application forms can serve to streamline application processes, reduce burden on applicants and reviewers, improve efficiency, and accelerate the review process. It has also been done to facilitate more diverse panel discussions as more panel members can read the applications and contribute actively to panel deliberations.

This type of intervention has been implemented by several funders, and even become the "new normal" for some.

There is some resistance to the approach, as some argue that reducing the amount of information available to the reviewers makes it harder to assess the proposals. Especially for large grants, more information might be needed to ensure accountability and guard against misallocation of funds. To counter this, a study on a UKRI-funded pilot with a simpler application process and shorter applications found both applicants and reviewers to be highly satisfied with the simplified process. In addition, more than half of surveyed reviewers deemed the amount of information sufficient for a good assessment and the overall review time shortened.

There is also some evidence that some applicants spend more time on shorter applications, so it is not clear whether time is saved for applicants. This is noted in a study of the EU's Horizon 2020 proposal evaluation system, which concluded that strict page limitations may increase the applicant burden. Evidence on this is mixed, however; the Australian Centre for Health Services Innovation (AusHSI) has taken on a streamlined application process featuring a short proposal of less than 1,200 words. An observational study from 2015 identified increased success rates (from 6% to 16%) as a result, and an overall decrease in time to just 8 weeks from application deadline to outcome announcements.

As a sidenote, it is worth mentioning that this intervention and the previous one (expression of interest/pre-proposal) are sometimes linked, as there is always a question not only of what kinds of information are needed, but also of what information is needed at what point. The recent review of research bureaucracy in the UK noted the importance of this. However, our review finds no sources that look into effects of this. Funders may consider monitoring and assessing further what kinds of information are best requested at early stages of application and which ones might be better left for later stages.

References	Interviewees & survey responses
• Barnett, A. G., Herbert, D. L., Campbell, M., Daly, N., Roberts, J. A., Mudge, A., & Graves, N. (2015). Streamlined research funding using short proposals and accelerated peer review: An observational study. BMC Health Services Research, 15(1), 2–7. https://doi.org/10.1186/s12913-015-0721-7 • Bendiscioli, S. et al., (2022), The experimental research funder's handbook, *Research on Research Institute,* URL: https://rori.figshare.com/articles/report/The_experimental_research_funder_s_handbook_final_version_/19459328 • GAO. (2016). Federal Research Grants: Opportunities Remain for Agencies to Streamline Administrative Requirements (Issue June). United States Government Accountability Office. https://www.gao.gov/products/gao-16-573 • Guthrie, S. et al. (2018), What do we know about grant peer review in the health sciences?, *F1000Research*, 6(1335), https://doi.org/10.12688/f1000research.11917.2 • Herbert, D. L., Graves, N., Clarke, P., & Barnett, A. G. (2015). Using simplified peer review processes to fund research: A prospective study. BMJ Open, 5(7), 1–7. https://doi.org/10.1136/bmjopen-2015-008380 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211 • Shepherd, J. et al. (2018), Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. PLoS ONE 13(5): e0196914. https://doi.org/10.1371/journal.pone.0196914 • Plus one confidential UKRI document shared for information with the study • Tickell, A. (2022) Independent Review of Research Bureaucracy: Final Report. Published by the department of Business, Energy and Industrial Strategy (BEIS), UK: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1094648/independent-review-research-bureaucracy-final-report.pdf	• 5 survey responses • 2 interviews

5 Main findings: Interventions in assessment process design

5.1 'Sandpits'/Matching events

Potential applicants are invited to an event to discuss possibilities and form teams for potential proposals. May involve some application-submission on the day.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Foster inter/multidisciplinary research, new collaborations and transformative research	Problems for access, EDI issues; can be partially resolved through remote events	****

This intervention intends to foster interdisciplinary research and more innovative proposals and solutions to research challenges, particularly when seeking to promote transformative research.

EPSRC has used the Ideas Factory Sandpit for over 10 years with positive outcomes in terms of the establishment of research communities. There is an observable culture change amongst participants embracing creativity and originality and an increase in the capacity of multidisciplinary researchers and their interaction in the UK. EPSRC has also run sandpits at distance (remote) with positive results. For respondents, this intervention creates opportunities for building new multidisciplinary partnerships and foster blue-skies ideas.[9]

There are however some negative effects from an EDI perspective reported in the literature and from consultees due sandpits' setup: intensive face-to-face interaction, mostly away from home, with durations of 1-5 days reduces the opportunities of participation of those with caring responsibilities and potentially for those with disabilities/sensory needs.

Remote sandpits offer more flexibility but do not overcome all the limitations identified. The EPSRC implemented a number of further mitigation measures: inviting and paying for carers to sandpits to enable the applicant to attend; adapting the facilitation style of the sandpit to make it more accessible; embedding more breaks into the sandpit and changing the model of the sandpit to be accessible virtually over a different timescale to ensure a reduction in screen time.

There is clear evidence and strong agreement on the positive impact of sandpits/matching events on fostering multidisciplinary research and innovative solutions to research challenges. Sources also converge on limitations and negative effects for EDI, though the mitigation efforts noted above may provide important ways forward.

References	Interviewees & survey responses
• Bendiscioli, S. et al. (2022), The experimental research funder's handbook, Research on Research Institute, https://rori.figshare.com/articles/report/The_experimental_research_funder_s_handbook_final_version_/19459328 • Guthrie, S. et al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf • Lodge, H. (2020). Sandpit Methodology: Results of a rapid literature search to inform a sandpit exercise for PETRA. https://petranetwork.org/wp-content/uploads/2020/05/Sandpit-methodology-overview-March-2020.pdf	• 3 survey responses • 2 interviews

---

[9] The EPSRC sandpits also include elements of group review, so this example also pertains to the group review section (5.13) of this report. However, we focus here on the collaboration-building aspect of the sandpits.

- Maxwell, K., Benneworth, P., & Siefkes, M. (2018). Sandpits can develop cross-disciplinary projects, but funders need to be as open-minded as researchers. LSE Impact Blog. https://doi.org/10.1057/s41599-018-0105-9
- Meadmore, K., Fackrell, K., Recio-Saucedo, A., Bull, A., Fraser, S. D. S., & Blatch-Jones, A. (2020). Decision-making approaches used by UK and international health funding organisations for allocating research funds: A survey of current practice. PLoS ONE, 15(11), 1–17. https://doi.org/10.1371/journal.pone.0239757
- Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: a realist synthesis. Research Integrity and Peer Review, 7(1), 1–28. https://doi.org/10.1186/s41073-022-00120-2
- Schnädelbach, H., Sun, X., Kefalidou, G., Coughlan, T., Meese, R., Norris, J., & McAuley, D. (2016). Creativity Greenhouse: At-a-distance collaboration and competition over research funding. International Journal of Human Computer Studies, 87, 1–19. https://doi.org/10.1016/j.ijhcs.2015.10.006
- Treem, J. W., Schneider, M., Zender, R. L., & Sorkin, D. H. (2018). Exploring the potential role of community engagement in evaluating clinical and translational science grant proposals. Journal of Clinical and Translational Science, 2(3), 139–146. https://doi.org/10.1017/cts.2018.311

5.2   2-stage application process

Two 'rounds' of peer or panel review are used, sifting out some after the first stage. May involve different parts of the application being reviewed at different stages, or a pre-proposal/EoI (see above).

Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce burden for reviewers, applicants and programme officers, increase relevance of stage-2 proposals	Slight danger of reduced levels of feedback	****

This intervention is strongly linked to the 'Pre-proposal/EoI' intervention. Often, they may be interchangeable. Stage 1 may involve a pre-proposal, though it is also possible that the same proposal document will be reviewed at stage 1 and stage 2. In such cases, this intervention is distinct. Review at the first stage is typically conducted by a review panel (often specifically put together for the call to reflect the thematic nature of the applications), though in some cases remote reviewers are also used.

The purpose of this intervention is to reduce overall burden of the evaluation process (on applicants, administrators and reviewers). It is also used to sift out applications that do not meet particular requirements (e.g. out of scope).

There are verified positive outcomes from this intervention for both funders and applicants. Wellcome has adopted it and become a regular practice in their evaluation process, reducing burden on written review by 50%. NIHR adopted it with successful results: increased number of applications, reduced number of applications per reviewer, lower cost per evaluation round (40% reduction) and shorter notification periods to applicants.

There is wide agreement among our sources and consultees on the positive effects of this intervention, sifting out applications that do not meet programme requirements. The only noted concerns are about limited feedback to first stage applications, meaning overall less feedback in the research system and consequent less scope for learning. Some scepticism was voiced during the study of the Horizon 2020 proposal assessment process, where the study team found that the length of the stage 2 applications did not significantly differ from single-stage applications and as such, claimed that at least 65% of stage 1 applications must be rejected in order for the overall process to reduce the burden rather than increase it.

References	Interviewees & survey responses
• Barnett, A. G., Herbert, D. L., Campbell, M., Daly, N., Roberts, J. A., Mudge, A., & Graves, N. (2015). Streamlined research funding using short proposals and accelerated peer review: An observational study. BMC Health Services Research, 15(1), 2–7. https://doi.org/10.1186/s12913-015-0721-7 • Mervis, J. (2016). NSF tries two-step review, drawing praise - And darts. Science, 353(6299), 528–529. https://doi.org/10.1126/science.353.6299.528 • Morgan, B., Yu, L. M., Solomon, T., & Ziebland, S. (2020). Assessing health research grant applications: A retrospective comparative review of a one-stage versus a two-stage application assessment process. PLoS ONE, 15(3), 1–18. https://doi.org/10.1371/journal.pone.0230118 • Treem, J. W., Schneider, M., Zender, R. L., & Sorkin, D. H. (2018). Exploring the potential role of community engagement in evaluating clinical and translational science grant proposals. Journal of Clinical and Translational Science, 2(3), 139–146. https://doi.org/10.1017/cts.2018.311 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211	• 3 survey responses • 3 interviews

5.3   Applicant anonymisation

Reviewers or panels members or both do not see the identity of the applicant/s.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce bias, foster innovative/transformative ideas	Limited ability to judge feasibility of projects	****

This intervention aims to reduce bias (e.g. in relation to institution, gender, career stage, etc), and to focus reviewers' attention on project idea rather than person to identify and fund more unconventional research.

This intervention is widely used. Among the examples we find are the NIH Director's Transformative Research Award, EPSRC's New Horizons scheme, ESRC's Transformative Research Scheme (currently paused) and the New Zealand HRC Explorer Grant. It is also in use at the FWF, VW Foundation (where it is variously paired with other interventions), and it has been piloted at the Swiss SNSF.

SNSF's Spark Fund evaluation found that anonymising applications attracts more unconventional research ideas. Evidence from VW foundation also shows increased success rates for women applicants and early career researchers. Similarly, using anonymisation in the FWF's 1000 Ideas programme attracted a more diverse pool of applicants than other programmes. Anonymisation is generally considered a 'gold standard' for reducing bias.

There is however some evidence to suggest that in the absence of personal information to judge the suitability of the applicant(s), reviewers/panellists sometimes report that they struggle to assess feasibility of projects from this point of view (though 'feasibility' as an assessment criterion usually covers aspects besides applicants' abilities). FWF also reported instances of some jury members saying they knew who the applicant was and why they should be funded. FWF then reminds the jury that this information is irrelevant to this assessment process.

There is also evidence to suggest that not all bias is eliminated through applicant anonymisation: one reviewed study finds that men and women tend to use language differently and reviewers reward some language uses more associated with men. FWF reported

that some applications included information about affiliation by accident, for example, referring to the ethics policy of a particular university.

Our research also finds that anonymisation often is coupled with other interventions, and that funders suspect it may be the combination rather than necessarily the intervention by itself that leads to positive outcomes (includes review without panel and partial randomisation). A separate application stage that is not anonymised can help mitigate the issue around judging feasibility (as practiced in ESRC's Transformative Research scheme).

References	Interviewees & survey responses
• Bendiscioli, S. et al. (2022), The experimental research funder's handbook, Research on Research Institute, https://rori.figshare.com/articles/report/The_experimental_research_funder_s_handbook_final_version_/19459328 • Boudreau, K., Guinan, E., Lakhani, K., & Riedl, C. (2016). Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science. Manage Science, 62(10). https://pubsonline.informs.org/doi/10.1287/mnsc.2015.2285 • Guthrie, S. et al. (2018), What do we know about grant peer review in the health sciences?, F1000Research, 6(1335), https://doi.org/10.12688/f1000research.11917.2 • Kolev, J., Murray, Y., & Fuentes-Medel, F. (2019). Is Blinded Review Enough? How Gendered Outcomes Arise Even Under Anonymous Evaluation. NBER Working Paper Series. http://www.nber.org/papers/w25759 • Lanei, J. N., Teplitskiy, M., Gray, G., Ranu, H., Menietti, M., Guinan, E. C., & Lakhani, K. R. (2022). Conservatism Gets Funded? A Field Experiment on the Role of Negative Information in Novel Project Evaluation. Management Science, 68(6), 4478–4495. https://doi.org/10.1287/mnsc.2021.4107 • Langfeldt, L., Ingeborgrud, L., Reymert, I., Svartefoss, S. M., & Borlaug, S. B. (2022). Evaluation of the Spark pilot. • Lauer, M. (2020). Anonymizing Peer Review for the NIH Director's Transformative Research Award Applications. NIH Extramural Nexus. https://nexus.od.nih.gov/all/2020/05/27/anonymizing-peer-review-for-the-nih-directors-transformative-research-award-applications/ • Matthews, D. (2020), German funder sees early success in grant-by-lottery trial, Times Higher Education, URL: https://www.timeshighereducation.com/news/german-funder-sees-early-success-grant-lottery-trial • Nakamura, R., Mann, L. S., Lindner, M. D., Braithwaite, J., Chen, M. C., Vancea, A., Byrnes, N., Durrant, V., & Reed, B. (2021). An experimental test of the effects of redacting grant applicant identifiers on peer review outcomes. ELife, 10, 1–15. https://doi.org/10.7554/eLife.71368 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211 • Solans-Domenech, M. et al., (2017), Blinding applicants in a first-stage peer-review process of biomedical grants: An observational study, Research Evaluation, 26:3, 181-189, URL: https://academic.oup.com/rev/article/26/3/181/3858258 • One additional confidential UKRI document shared for information with the study	• 1 survey response • 4 interviews

5.4   Automation-assisted reviewer allocation

Using algorithms/ AI / text recognition to aid allocation of reviewers to applications.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase efficiency/ reduce burden in reviewer allocation; better matching of applications to reviewers	Technology is not widely tested; some algorithms may have problems	***

This is an intervention that has become possible with some modern application management systems. It typically involves matching applications' keywords or other machine-readable details to reviewers who are associated with those keywords (e.g. via applications they have reviewed in the past). We use the term 'automation-assisted' rather than just 'automated' to denote that a human element still remains in the process at all times. I.e. whatever the automated system recommends still needs to be checked by funder staff.

The objective of this approach is to increase efficiency in expert allocation, to reduce administrative burden, and enable a higher degree of quality in application reviews due to identifying the most knowledgeable experts on the topics. It may also lead to a decrease in declined review-invitations (reviewers declining due to subject matter being outside their expertise). This technology can also be used to identify potential conflicts of interest.

Automation-assisted reviewer allocation is in ongoing use at the Australian Research Council (ARC) with reported satisfaction. We find mentions of previous use at CIHR but with poorer reception, although a review study suggests this is due to avoidable challenges with implementation. The Research Council of Norway (RCN) uses an online tool to find experts to assess applications. RCN reports significant time savings and access to a broader pool of reviewers. In addition, we find anecdotes of numerous instances of use in journal peer review with a high level of reviewer satisfaction.[10]

In short, this is a very promising intervention. We find no difficulties at a general level. However, it may be subject to pitfalls simply because it is a relatively new technological approach. For example, if reviewers are identified based on past reviews, then there is a potential challenge around how to integrate new first-time reviewers into the system. It is not an insurmountable challenge, but one that requires consideration.

The approach has been studied (and algorithms have been developed), but implementation so far is somewhat limited (however, we note that we found no high-profile announcements of implementation even where it had occurred, so it is possible that it is used more than it appears). Potential hazards can likely be avoided through sharing of successful algorithms and technical procedures by funders who have had positive results.

References	Interviewees & survey responses
• Cechlárová, K., Fleiner, T., & Potpinková, E. (2014). Assigning evaluators to research grant applications: the case of Slovak Research and Development Agency. Scientometrics, 99, 495–506. • Guthrie, S. (2019), Innovating in the research funding process: peer review alternatives and adaptations, URL: https://academyhealth.org/publications/2019-11/innovating-research-funding-process-peer-review-alternatives-and-adaptations • Kalmukov, Y. (2020). An algorithm for automatic assignment of reviewers to papers. Scientometrics, 124(3), 1811–1850. https://dl.acm.org/doi/abs/10.1007/s11192-020-03519-0 • Kolarz, P., Arnold, E., Davé, A., Andreasson, H., & Bryan, B. (2018). How Research Funders Ensure the Scientific Legitimacy of their Decisions: Investigation in support of the design of Formas (Issue November). Technopolis Group. https://www.researchgate.net/publication/333853175_How_Research_Funders_Ensure_the_Scientific_Legitimacy_of_their_Decisions_Investigation_in_support_of_the_design_of_Formas_scientific_management	• No survey responses • 1 interview

[10] RCN's experience with the LookUp tool has been detailed in a case study by Elsevier: https://www.elsevier.com/__data/assets/pdf_file/0009/484083/Expert-LookupResearch-Council-of-NorwayCSWEB.pdf

- Kolarz, P. et al. (2018), UKRI Research and Innovation Funding Service (RIFS) visioning work, Technopolis Group, URL: https://www.technopolis-group.com/wp-content/uploads/2020/02/Support-to-the-generation-of-a-Research-and-Innovation-Funding-Service.pdf
- Pupella, V., Monteverde, M. E., Lombardo, C., Belardelli, F., & Giacomini, M. (2014). A semi-automatic web based tool for the selection of research projects reviewers. Studies in Health Technology and Informatics, 205, 950–954. https://doi.org/10.3233/978-1-61499-432-9-950
- Recio-Saucedo, A. et al. (2022), What works for peer review and decision-making in research funding: a realist synthesis, Research Integrity and Peer Review, 7(2), https://doi.org/10.1186/s41073-022-00120-2
- One additional confidential UKRI document shared for information with the study

5.5   Dragon's den-style pitch

Applicants are invited to pitch their proposal in front of a panel, and panels have an opportunity to ask questions. This differs from an interview in that no other form of evidence (e.g. written proposals or external expert review) is used in the assessment.

Main intended aim(s)	Main hazard(s)	Evidence strength
Increase stakeholder involvement; fund novel, transformative ideas	Favours applicants with sharp presenting skills; may present access-problems	*

This intervention seeks to provide an innovate way of funding allocation by facilitating stakeholder engagement with the research ideas. Fostering more diverse and transformative research projects has also been noted as an aim here, though 'transformative' may here suggest societal transformation rather than transformation of scientific practice itself.

EPSRC has used Dragon's Den style events in the Bright IDEAS Award programme. There is no evaluation of the programme but it claims to have funded highly diverse set of applicants and potentially transformative research.

It has also been used by the Hounslow and Richmond Community Healthcare NHS Trust to ensure that some of the most innovative practices are captured and supported. Two pitching panels were carried out with positive effects in terms of mentoring, fostering collaborative work and innovation in the trust. In another case at the National Cancer Research Institute, a Dragon's Den event was used to facilitate patients' involvement in epidemiological research, which resulted in positive feedback from participants in terms of their interest in continuing to engage with the research.

Authors emphasise the role of independent facilitators to run the process. In other words: there needs to be sufficient briefing and oversight of the 'dragons'. More generally, there is a perceived difficulty in that these events will only suit specific types of individuals (good presentation skills, able to access the events, native speakers) and disadvantage others. This is therefore unlikely to be a widely suitable intervention.

References	Interviewees & survey responses
- Mazhindu, D., & Gregory, S. (2015). Dragons' Den: promoting healthcare research and innovation. Nursing Management, 22(4), 28–33. - Morris, M., Alencar, Y., Rachet, B., Stephens, R., & Coleman, M. P. (2020). Fleshing out the data: when epidemiological researchers engage with patients and carers. Learning lessons from a patient involvement activity. BMJ Open, 10(9), e036311. https://doi.org/10.1136/bmjopen-2019-036311	- No survey responses - 1 interview

5.6   External review only (no panel)

Proposals are only assessed by external reviewers and review scores are simply combined to give the final score.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce risk-averseness of panels, reduce burden and costs, better match applications to expertise	Reduced layers of risk control, potential lack of transparency	**

This approach is intended to reduce risk-averseness in panel discussions and also to reduce burden (in this case for panellists rather than reviewers in general). Further, it gives more flexibility in matching reviewers with applications as the choice is not limited to a relatively small number of panellists – the aim being better matching between reviewers' expertise and applications. This intervention may also potentially cut costs of in-person panels. We find examples of its use at Australia's NHMRC, Switzerland's SNSF and at NERC in the UK.[11]

For SNSF's Sinergia, the evaluation found that original and unconventional research was given better chances by including originality and unconventionality as key review criteria and funding proposals based on aggregated reviewer grades (rather than panel discussions). Omitting panel meetings was also a way of reducing review costs for small grants.

NHMRC's data (based on reviewers' declared suitability before peer review) and responses to NHMRC's panel member survey suggested better matching of reviewers to applications in 2020 than in 2019.

For the NHMRC case, there was a perceived lack of transparency in the initial round (2020), however, this was mitigated by the addition of reviewer comments in 2021, when previously only scores had been released. This intervention appears also to be largely limited to use for small grants. For larger ones, there is a perceived danger due to fewer 'layers' of risk control.

References	Interviewees & survey responses
• Kelso, A. (2021). Peer review for Ideas Grants in 2021. CEO Communique, February 2021, National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/news-centre/peer-review-ideas-grants-2021	• No survey responses • 1 interview

5.7   Group review

The same reviewer comments on multiple proposals.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Facilitate consensus-building, increase diversity of reviews	Group-bias	*

This intervention aims to facilitate consensus and deliver more comprehensive reviews particularly when reviewing manuscripts for academic journals. We find very limited evidence on this intervention.

---

[11] For the NERC example we find no evaluation evidence, which is why this example is not discussed further. We do note that while there is no panel, reviews are moderated by an external moderator who is an expert in the field and who makes a funding recommendation to NERC.

The Association of American Medical Colleges experimented with this intervention and found that more thorough feedback was provided to researchers. Reviewers changed their initial individual assessments throughout the group review process and reduced time was required to evaluate the papers compared to what reviewers would spend individually.

The sources we find note a risk of group bias and that shared views may consolidate over time.

References	Interviewees & survey responses
• Dumenco, L., Engle, D., Goodell, K., Nagler, A., Ovitsh, R., & Whicker, S. (2017). Expanding Group Peer Review: A Proposal for Medical Education Scholarship. Academic Medicine, 92(2), 147–149. • Nagler, A., Ovitsh, R., Dumenco, L., Whicker, S., Engle, D., & Goodell, K. (2019). Communities of Practice in Peer Review: Outlining a Group Review Process. Academic Medicine, 94(19), 1437–1442.	• No survey responses and interviews

5.8   Changing the number of reviewers

2-3 external reviews of applications is typical for responsive-mode grant funding, but this number may be lowered to 1 or significantly increased.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase numbers: improve robustness/reliability Decrease numbers: save time/burden/cost	Increase numbers: a single bad review can sink an application; labour intensive Decrease numbers: reduced robustness, potential for greater bias	***

Increasing the number of reviewers is done to improve quality and reliability and mitigate against random variations in individual reviews and to improve ability to address additional assessment criteria, whereas reducing the number of reviewers can be done with the aim to reduce cost and burden of reviews.

There is a broad consensus that reliability of decisions increases with the number of reviewers. This has been demonstrated in quantitative studies and confirmed by funder experience. Several studies have found that five reviewers is an optimal upper limit for robustness, but this is based on data from specific types of programmes. For very small grants, a single reviewer is sometimes used (e.g. at Germany's DFG).

In short, setting the number of reviewers balances two objectives: adding more reviewers to optimise reliability, or reducing the number of reviewers to improve resource efficiency. The optimal number will inevitably depend on situation-specific trade-offs between cost and benefit of adding more reviewers and the tolerance for mistakes in specific situations.

There is some disagreement about the appropriateness of using a single reviewer. This is sometimes done for very small grants, but some argue that the minimum should be two reviewers.

Further, inter-rater-reliability (IRR) is the subject of a large volume of technical literature across this and related topics, and beyond the scope of this discussion.

References	Interviewees & survey responses
• Bendiscioli, S., Firpo, T., Bravo-biosca, A., Czibor, E., Garfinkel, M., Stafford, T., & Wilsdon, J. (2022). The experimental Research funder's handbook. Research on Research Institute (RoRI).	• No survey responses • 1 interview

- Langfeldt, L. (2016). The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. Social Studies of Science. https://link.springer.com/book/10.1007/978-3-030-75263-7
- Nous Group. (2018). Consultation on the development of peer review for NHMRC's new grant program - Final Report. https://doi.org/10.1111/j.1834-7819.1982.tb04109.x
- UKRI. (2021). How does the number of grant reviews affect the reliability of peer review decisions? December 2021.

5.9 Interviews

Lead applicant (or several application team members) may do a presentation (optional) and are then asked questions on their application by panel members, reviewers or funder representatives.

Main intended aim(s)	Main hazard(s)	Evidence strength
Improve quality of reviews, increase scrutiny; opportunity to respond to criticism	Resource-intensive, bias and/or disadvantage for certain groups	**

Interviews serve different purposes depending on the scheme. They can serve to demonstrate the applicant's presentation skills (which may be especially relevant to commercialisation projects), improve engagement with panellists (assuming panellists are the interviewers), allow applicants to respond to comments and defend their proposal, and/or improve the overall quality of reviews as the interview will provide reviewers with additional context.

When used, interviews often occur at the end of the process: due to their resource-intensive nature, efficiencies can be gained by having interviews as the final stage of a multi-stage assessment process (by which point most applicants will have already been rejected, meaning there are fewer interviews to do). In addition to standard interviews, they can also take the form of a scientific symposium or workshop. The practice is often used for early career fellowships (i.e. strongly person-centred awards) and schemes aiming to fund particularly transformative research.

Funders have found interviews to be a helpful way of assessing proposals or candidates against specific objectives, whereas others use it more widely to improve the quality of the review. It can be difficult to evidence the exact effect of using interviews but one study found that that the interview stage had a significant impact on the final grant selection.

Interviews are typically used in addition to other types of assessment.

They are particularly resource-intensive, requiring time and space set aside for each individual applicant. For this reason, Wellcome has for instance decreased its use of interviews in recent years. Assessment through interviews can also be biased against certain personality types (e.g. introverted, nervous, non-native speakers). In-person interviews may also pose difficulties for applicants with caring responsibilities or disabilities. However, we note on a final point that there is limited research on the effectiveness of interviews in terms of achieving certain types of funding outcomes, despite their relatively frequent use.

References	Interviewees & survey responses
- van Arensbergen, P., van der Weijden, I., & van der Besselaar, P. (2014). Academic talent selection in grant review panels. (Re)Searching Scientific Careers, 25–54. - Lal, B., Wilson, A., Jonas, S., Lee, E., Richards, A., & Peña, V. (2012). An outcome evaluation of the National Institutes of Health (NIH) Director's pioneer award (NDPA) program, FY 2004-2006. 204. https://www.ida.org/~/media/Corporate/Files/Publications/STPIPubs/ida-p-4899.ashx	- No survey responses - 3 interviews

- Luukkonen, T., Stampfer, M., & Strassnig, M. (2015). Evaluation practices in the selection of ground-breaking research proposals. fte Workiing Paper. https://www.etla.fi/en/publications/evaluation-practices-in-the-selection-of-ground-breaking-research-proposals/

5.10 Moderation of reviews

Reviews are processed internally by funding organisation staff and are only passed to the external panel if they are of sufficient quality.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Ensure consistency/quality of reviews	Time-consuming for administrators; administrators may not have sufficient thematic expertise	*

Moderation of reviews is intended to ensure quality of the reviews received in order not to waste time or have an inconsistent evidence base at later stages of the evaluation process, particularly during panel reviews and for feedback from assessor/reviewers. Moderation might only involve a basic 'usability-check' (i.e. ensuring that reviews are not just one line of text or similar) or more involved engagement to check if the reviews meet a broader set of criteria.

All UKRI councils use some degree of review moderation. For example, Innovate UK introduced a moderation phase to review outlier scores from assessors to ensure consistency, since they were receiving complaints from applicants about conflicting feedback from assessors.

Our research received some anecdotal comments noting on one hand that moderation of reviews does bring benefits in terms of consistent review quality, but that it places a burden on administrators' time. Additionally, administrators may not always have all the necessary thematic expertise if moderation extends to thematic aspects.

Beyond such anecdotal points, our research found no further evidence on the efficacy or hazards of this intervention. Literature appears to be insufficient and does not distinguish from moderation panel intervention.

References	Interviewees & survey responses
- Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211 - Wieczorkowska, G., & Kowalczyk, K. (2021). Ensuring sustainable evaluation: How to improve quality of evaluating grant proposals? Sustainability (Switzerland), 13(5), 1–11. https://doi.org/10.3390/su13052842	- No survey responses or interviews

5.11 Moderation panel

Assessment panels use external reviews alongside their own expertise to assess the proposal. Moderation panels do not use their own expertise but can only use the reviews to inform their scores.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Ensure consistency, increase expertise and robustness of reviews	Not known	*

Assessment panels[12] where members can bring in their own expertise are the baseline approach funders use. Our research found no evidence of the effectiveness of using moderation panels. UKRI uses moderation panels in some programmes where assessment panels cannot cover the breadth of expertise required to assess applications from diverse disciplines. However, the effectiveness of the moderation panels is not systematically studied; therefore, evidence strength for this intervention is weak.

References	Interviewees & survey responses
None.	• No survey responses or interviews

5.12 Panel only (no postal/external review)

Proposals are only assessed by a panel of experts.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase speed of decisions, efficiency, ensure consistency of reviews, include strategic perspectives in reviewing	Difficulty to cover the required expertise in a panel, may still need additional reviews, potential bias	***

This intervention is similar to the 'group review' intervention, though it involves reviewers actually meeting as a group (review panel), which the 'group review' intervention does not. It is used for a variety of reasons:

- To speed up funding decisions

- To reduce written feedback (and its associated costs and burden)

- To improve quality and consistency of feedback to applicants

- To assess riskier research proposals and where strategic considerations play a central role in the judgement process (e.g., ensuring EDI is properly assessed)

AHRC adopted panel-only assessment for the Equality, Diversity and Inclusion Engagement Fellowship (EDIEF) pilot programme, forming a bespoke panel that embedded EDI in the evaluation process and sped up funding decision-making. The Royal Society has also used it with positive results, funding more high-risk high-reward research proposals, and an increase in the number of individuals willing to participate as panellists. The Royal Society has also found that rigour has remained high, which is also reflected in our survey responses. Cancer Research UK has implemented panel-only assessment, resulting in a significant reduction of written peer review requests and highlighted the benefits of having in-person discussions as a more valuable way of evaluating research applications.

Panel-only review was also a technique used by a range of funders in their R&I funding responses to Covid-19, as a mechanism for ensuring that awards were made quickly and could thus respond to the societal emergency at hand. Several examples are detailed in the process review of UKRI's response to Covid-19 (which contains a review of six international comparators), though long-term evaluations of effectiveness are not yet available.

---

[12] Compared to moderation panels, assessment panel members can bring in their own expertise and this approach is mostly part of the baseline process and therefore not considered as an intervention in this study.

It is a challenge to represent enough expertise on a panel to cover the potentially broad thematic and/or subject range of a large number of in-coming applications. There is therefore typically a need to have large panels and funders may still have to rely on external reviewers when applications fall outside the panel expertise, or in the absence of agreement.

Configuration of panels may be difficult as panellists may need to be recruited from distant subject domains, potentially creating some administrative burden in panel set-up.

There is broad agreement on effectiveness of this intervention (reducing burden and speeding up decision-making) but controversy around its effects regarding bias. Some evidence showed panels purposefully used to embed and ensure EDI throughout the process, while in other programmes it was found that it failed to sufficiently factor this in. Cross-disciplinary panel composition may also result in 'communication problems'.

We note that this is the most frequently discussed intervention in our UKRI staff survey.

References	Interviewees & survey responses
• Banal-Estañol, A., Liu, Q., Macho-Stadler, I., & Pérez-Castrillo, D. (2021). Similar-to-me effects in the grant application process: Applicants, panelists, and the likelihood of obtaining funds. Economics Working Paper Series, 1801. • Blackburn, M., Coutinho, K., & Suviste, H. (2022). The Equality, Diversity and Inclusion Engagement Fellowship Pilot AHRC Funding Scheme Report. The Open University. • Feliciani, T., Morreau, M., Luo, J., Lucas, P., & Shankar, K. (2022). Designing grant-review panels for better funding decisions: Lessons from an empirically calibrated simulation model. Research Policy, 51(4), 104467. https://doi.org/10.1016/j.respol.2021.104467 • Kolarz, P., Arnold, E., Bryan, B., D'hont, J., Horvath, A., Simmonds, P., Varnai, P. & Vingre, A. (2022) Process review of UKRI's research and innovation response to Covid-19. UKRI. https://www.ukri.org/wp-content/uploads/2022/01/UKRI-180122-ProcessReviewUKRIResponseCOVID19-FinalReport.pdf • Samuel, G. N., & Derrick, G. (2021). Exploring the degree of delegated authority for the peer review of societal impact. Science And Public Policy.	• 6 survey responses • 2 interviews

5.13 Peer allocation

The applicants are also the assessors and review the proposals they are competing against to decide who gets funding.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Lesson administrative burden, reduce pressure to identify reviewers	Possibly open to abuse/gaming, adds to applicant burden	***

This intervention evens out the number of applicants to reviewers and is intended to lessen the administrative burden on reviewers and shorten the overall time taken to recruit reviewers.

There are a small number of known instances where the intervention is in use, but the results are cautiously positive across the board. In one scheme, seven rounds of review were organised within a year, with a total of 614 reviews carried out by 201 reviewers (some being applicants and some not). When compared, the two groups of scorings correlated. Where successfully in use, it relieves the pressure to identify expert reviewers. It appears to be a successful way to expedite the review process without impacting the integrity of the selection.

The NSF ran an experiment on peer allocation in 2013. As a condition of application, applicants had to commit to assessing seven other proposals submitted to the scheme and then rank the proposals from best to worst. The NSF also employed a mechanism to dissuade reviewers from

downgrading a competitor's proposal in order to boost their own: reviewers earned bonus points on their own applications if their assessments of other proposals closely matched what their colleagues thought. An article in *Science* reports that the system saved time and money, but that the need for 'group consensus' may disadvantage novel, unconventional ideas.

Peer allocation may risk being abused if the consistency of scoring with non-applicant reviewers is not monitored and/or the approach is mainstreamed. It is possible that this is mainly viable in smaller, perhaps early-career settings. Moreover, peer allocation has the reverse effect on the administrative burden on applicants, particularly if additional training is required.

As a side note, the ESRC Transformative Research scheme had an element of this but in the context of a dragons' den-style event rather than application review proper (termed 'pitch-to-peers'). The evaluation found that, contrary to simple self-interest arguments, reviewers were generally supportive of their fellow applicants, resulting in collegial discussions.

References	Interviewees & survey responses
• Graham, C. L. B., Landrain, T. E., Vjestica, A., Masselot, C., Lawton, E., Blondel, L., Haenal, L., Greshake Tzovaras, B., & Santolini, M. (2022). Community review: a robust and scalable selection system for resource allocation within open science and innovation communities. F1000Research, 11, 1440. https://doi.org/10.12688/f1000research.125886.1 • Guthrie, S. (2019). Innovating in the Research Funding Process : Peer Review Alternatives and Adaptations - Horizon Scan. Paradigm Project, AcademyHealth. https://www.academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf • Kolarz, P., Arnold, E., Farla, K., Gardham, S., Nielsen, K., Rosemberg, C., & Wain, M. (2016). Evaluation of the ESRC Transformative Research Scheme. Technopolis Group, July 2016. https://esrc.ukri.org/files/research/research-and-impact-evaluation/esrc-transformative-research-scheme-evaluation/ • Mervis., J (2014) A radical change in peer review: A pilot project to ease pressure on NSF's vaunted peer-review system required grant applicants to review seven competing proposals. Science, 18 Jul 2014, Vol 345, Issue 6194, pp. 248-249. https://www.science.org/doi/full/10.1126/science.345.6194.248 • Powering Discovery: The Expert Panel on International Practices for Funding Natural Sciences and Engineering Research. (2021). The Council of Canadian Academies (CAA/CAC). https://cca-reports.ca/wp-content/uploads/2021/05/Powering-Discovery-Full-Report-EN_DIGITAL_FINAL.pdf • Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: a realist synthesis. Research Integrity and Peer Review, 7(1), 1–28. https://doi.org/10.1186/s41073-022-00120-2	• No survey responses or interviews

5.14 Programme manager's discretion

Applications go directly to the programme/scheme manager, who can recommend funding or even decide to fund unilaterally. Usually involves complete by-pass of peer and panel review.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Shorten time-to-grant, reduce overall burden, respond to emergencies, fund high-risk-high-reward projects likely to fail in peer review	Evidence that it may be under-used as programme managers themselves can be risk averse; lacks transparency, potentially a 'winners' game'	***

This approach is used to support exploratory or high-risk / high-reward projects that might not be selected through potentially more conservative peer review.

This approach has also been used to respond quickly to address urgent issues or grasp immediate opportunities for innovative developments. For example, several funders including NSF, NWO, NRC, and ANR have on occasion partly or fully bypassed peer review. Several of them relied on programme managers in parts of their Covid-19 response funding and found that this accelerated the funding decisions at a time when research projects had to start as soon as possible.

The approach can also be applied by leaving the final decision to funding staff (i.e. programme managers) after an initial shortlisting/sifting through a more traditional external review process. Furthermore, even when programme managers are tasked with assessing the applications, there is still usually an option to recruit external expertise if necessary.

This approach has been found to be successful in supporting exploratory research that often led to follow-on funding and significant results further down the line. The approach also encourages dialogue between applicants and staff at the funding organisation.

This approach is particularly common in the USA and is associated with the 'DARPA' approach, considered highly successful and attracts a lot of attention from businesses. NSF used the approach in the Small Grants for Exploratory Research programme introduced in 1990. Now it is applied in the successor programme – the RAPID instrument, which is used to fund research in response to emergencies.

In one case, actual use of discretionary allocation was found to be much lower than the allowed limit (up to 5% of grant budget).

In opposition to wide-spread use of this mechanism, it is argued that the selection process lacks transparency, effectively basing decisions on one person's opinion. It has also been argued that the successful application of the 'DARPA' model is a 'winners game' potentially benefitting the most well-established and well-connected researchers.

References	Interviewees & survey responses
• Ginther, D. K., & Heggeness, M. L. (2020). Administrative Discretion in Scientific Funding: Evidence from a Prestigious Postdoctoral Training Program. Research Policy, 49(4). • Guthrie, S. et al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf • Kolarz, P., Arnold, E., Bryan, B., D'hont, J., Horvath, A., Simmonds, P., Varnai, P., Vingre, A. (2021). Process review of UKRI's research and innovation response to COVID—19 - Final report. Technopolis Group, 15 September 2021. https://www.technopolis-group.com/wp-content/uploads/2022/01/Adriana-Badau-BE.pdf • Wagner, C. S., & Alexander, J. (2013). Evaluating transformative research programmes: A case study of the NSF Small Grants for Exploratory Research programme. Research Evaluation, 22(3), 187–197. https://doi.org/10.1093/reseval/rvt006	• No survey responses • 2 interviews

5.15 Standing panels vs. portfolio panels

Standing panels are the same year on year (with some replacement due to retirement from the panel). Portfolio panels are assembled based on the proposals received and therefore will be comprised differently in each round of funding.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Standing panels ensure consistency, and may be the site of long-term learning and interdisciplinary conversation	Standing panels may potentially lead to institutionalised bias	***

Broadly speaking, standing panels ensure greater consistency over time and creation of certain 'cultures' and understandings of specific scheme aims, while portfolio (or 'ad-hoc') panels can be assembled to better reflect the thematic and disciplinary spread of a specific pool of applications.

The literature highlights that standing panels ensure consistent evaluation. There appears also to be a link with more consistent and comprehensive feedback on applications (particularly important for resubmitted proposals), and creating opportunities for interdisciplinary conversations between panellists, reviewers and applicants, including over time as a standing panel 'matures'.

In some cases, standing panels present an opportunity to develop capacity of inquiry of reviewers/staff and for professional development of applicants. They also reduce recruitment burden on programme officers as members of standing panels are normally appointed for several-year periods. For example, the review of the National Institutes of Health National Institute on Disability and Rehabilitation Research funding processes concluded that programme staff managing programmes with standing panels face less burden in peer recruitment.

While various forms of training (e.g. EDI training) has a longer 'effect' on standing panels, there may in the absence of such training also be more institutionalised bias and narrow perspectives, so this needs to be considered when configuring them to offset these potential drawbacks.

The main advantage of the portfolio panels is a fresh view and better ability of peers to assess the specifics of the concerned funding programme as the peers are selected specifically for the call or programme. However, we find no empirical evidence assessing the functioning of portfolio panels and, in the literature, the associated benefits are reported as assumptions about how the portfolio panels would work.

It is worth noting that hybrid-versions are possible, and practiced to some degree by many funders. For example, the Human Frontier Science Program (HFSP) uses standing panels where each panellist has a finite tenure. Once a panellist's tenure expires, secretariat staff may consider any changes over time to the portfolio of applications (evolving themes and new emerging methods or interdisciplinary perspectives) when identifying new panellists.

References	Interviewees & survey responses
• Bonetta, L. (2008). Enhancing NIH Grant Peer Review: A Broader Perspective. Cell, 135(2), 201-204. https://doi.org/10.1016/j.cell.2008.09.051 • CGIAR IAES. (2023). Standing Panel on Impact Assessment (SPIA). CGIAR Independent Advisory and Evaluation Services (IAES). https://iaes.cgiar.org/spia • Gallo, S., et al. (2018). Participation and Motivations of Grant Peer Reviewers: A Comprehensive Survey of the Biomedical Research Community. bioRxiv., URL: https://www.biorxiv.org/content/10.1101/479816v2 • IES. (2019). IES Guide for Grant Peer Review Panel Members: Information, tips and instructions for members of IES scientific grant peer review panels. Institute for Education Sciences, US Department of Education. https://ies.ed.gov/director/sro/pdf/IESGuideforPeerReviewersFY2019.pdf • IES. (2019). IES Guide for Grant Peer Review Panel Members: Information, tips and instructions for members of IES scientific grant peer review panels. Institute for Education Sciences, US Department of Education. https://ies.ed.gov/director/sro/pdf/IESGuideforPeerReviewersFY2019.pdf • Kolarz, P. et al (2023) Organisational and Procress Review of the Human Frontier Science Program. Technopolis Group. Published by HFSP: https://www.hfsp.org/node/74873#book/	• 1 survey response

- Kolarz, P. et al. (2021). Process review of UKRI's research and innovation response to COVID-19. UKRI. URL: https://www.ukri.org/wp-content/uploads/2022/01/UKRI-180122-ProcessReviewUKRIResponseCOVID19-FinalReport.pdf
- Becoming a Peer Reviewer for NIJ (2023). National Institute of Justice. Available at: https://nij.ojp.gov/funding/becoming-peer-reviewer-nij
- Newton, P. & Feucht, T. (2012). Improving NIJ's Peer Review Process: The Scientific Review Panel Pilot Project. National Institute of Justice Journal. Available at: https://nij.ojp.gov/topics/articles/improving-nijs-peer-review-process-scientific-review-panel-pilot-project
- NIH. (2023). CSR Data & Evaluations. National Institutes of Health, Center for Scientific Review. https://public.csr.nih.gov/AboutCSR/Evaluations
- Rivard, J., O'Connell, M. & Wegman, D. (2012). Review of Disability and Rehabilitation Research: NIDRR Grantmaking Processes and Products. National Academy of Sciences. Available: https://www.ncbi.nlm.nih.gov/books/NBK189591/pdf/Bookshelf_NBK189591.pdf
- UKRI (2022). Responsive mode grant assessment process. Available: https://www.ukri.org/about-us/esrc/grant-assessment-panels/
- University of Florida College of Education (2015). An Overview: IES Procedures for Peers Review of Grant Applications. Available: https://education.ufl.edu/educational-research/5499
- Wigginton, N., et al. (2021). Managing internal nomination and peer review processes to reduce bias. University of Michigan. Available at: https://research.umich.edu/wp-content/uploads/2021/11/guidance_for_equitably_managing_nomination_and_peer_review_process_to_reduce_bias_on_web_2-9-21_9_pm.pdf

5.16 Use of international assessors

Having quotas for assessors based in countries other than the funder's 'home' country. May extend to mandating all-international panels and/or reviewers.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Avoid conflicts of interest, ensure required expertise and fill gaps; bring in specific country-expertise	May require more guidance/training for panellists	**

International assessors (reviewers and/or panellists) are used to ensure required expertise (in-country knowledge, international development knowledge, to fill competence gaps in the funder-country). Particularly in smaller countries, there may be minimum quotas for international reviewers, or even mandates only to use international reviewers to avoid conflicts of interest among reviewers. For example, the Austrian Science Fund FWF uses only international reviewers for this reason.

UKRI, Wellcome, the Royal Society and CRUK are among many funders who have used international reviewers extensively and found it to be effective in diversifying and expanding the pool of reviewers and ensuring review quality, particularly from developing countries.

Funders have also used international assessors to fill gaps and in cases of knowledge/context-specific needs. Several funders are also keen to use this intervention more often because of its effectiveness and benefits. However, some also expressed concerns regarding country differences in the assessment process that may require extra guidance for some panellists. Disparities across different countries' typical assessment processes may require additional training/guidance for international reviewers.

While the likelihood for conflicts of interest among national reviewers is far more significant in small countries, this issue does also apply to larger countries to some extent: some consultees for our study also expressed significant support for international reviewers to mitigate for the conflict of interest in areas where there are small numbers of potential reviewers in the UK.

References	Interviewees & survey responses
• Tamblyn, R., Girard, N., Qian, C. J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. Cmaj, 190(16), E489–E499. https://doi.org/10.1503/cmaj.170901	• 3 survey responses • 4 interviews

5.17 Use of metrics

Use of metrics and bibliometrics as part of the evidence-base to inform decision-making.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Provide additional information about applicants, increase robustness of review	Highly controversial: poor measure of excellence, open to bias and abuse, may contravene the DORA[13]	***

Metrics can be used to support the assessment of funding applications, providing additional information about the applicants' track record. If used, it is typically early in the assessment process. The most commonly used metrics are reportedly field-normalised citation measures and proportion of publications among the most cited in the field. Some UKRI schemes have also used grant income or Research Excellence Framework (REF) outcome metrics in the past, prior to UKRI becoming signatory to the DORA. While use of metrics overall is rare, when it does appear it tends to be in programmes funding research in bio-medical fields.

Recent survey evidence shows that bibliometric indicators are viewed as important by some reviewers, particularly in the early stages of the review to assess the candidate, and less important at the panel stage.

The use of metrics is controversial, and many limitations have been identified: first, various objections hold that metrics are a poor way of assessing research excellence and potential. Second, their use may lead to biases (e.g. around gender, career stage or research field) and that bibliometric indicators are often used unethically. It cannot be ruled out either that the focus on track record (as demonstrated by bibliometric analysis) can contribute to a vicious circle where those with a shorter track will be rendered with fewer research activities overall, while funding concentrates on established individuals. Survey evidence suggests that reviewers who themselves have good personal bibliometric impact scores are more likely to regard metrics as important.

In short, evidence suggests that some reviewers find bibliometric indicators useful as supporting information, but there are widespread concerns about their use in the research communities.

We note that despite the apparently widespread use and despite general controversy, none of our survey respondents and interviewees commented on this intervention.

References	Interviewees & survey responses
• Curry, S., Gadd, E., & Wilsdon, J. (2022). Harnessing the Metric Tide: indicators, infrastructures & priorities for UK responsible research assessment. Report of The Metric Tide Revisited panel (Issue December). December 2022.	• No survey responses and interviews

---

[13] https://sfdora.org/read/

- Gunashekar, S., Wooding, S., & Guthrie, S. (2017). How do NIHR peer review panels use bibliometric information to support their decisions? Scientometrics, 112(3), 1813–1835. https://doi.org/10.1007/s11192-017-2417-8
- Langfeldt, L., Reymert, I., & Aksnes, D. W. (2021). The role of metrics in peer assessments. Research Evaluation, 30(1), 112–126. https://doi.org/10.1093/reseval/rvaa032
- Viner, N., Green, R., & Powel, P. (2006). Segmenting academics: Resource targeting of research grants. Science and Public Policy, 33(3), 166–178. https://doi.org/10.3152/147154306781779037
- Bendiscioli, S., & Garfinkel, M. (2021). EMBO Science Policy Programme Dealing with the limits of peer review with innovative approaches to allocating research funding. EMBO Science Policy Programme.
- Sato, S., Gygax, P., Randall, J. & Schmid Mast, M. (2021), The leaky pipeline in research grant peer review and funding decision: challenges and future directions, Higher Education, 82, 145-162, https://doi.org/10.1007/s10734-020-00626-y

5.18 Use of non-academic assessors (i.e. industry, policy & practice, patients, etc)

Having quotas for non-academic assessors. May extend to all-user panels and/or reviewers. May take the shape of consultation rather than directly making formal funding recommendations.

Main intended aim(s)	Main hazard(s)	Evidence strength
Increase societal relevance and impact	May dilute notions of basic research, not recommended for such contexts	****

The inclusion of non-academics is closely related to the increasing priority given to societal use and impact of research. Depending on the context, including non-academic reviewers may aim to represent stakeholder concerns (e.g. patients), improve the assessment of relevance and potential impact (e.g. using industry reviewers), potential interest among users, as well as feasibility of real-world applications (e.g. using technicians to support the assessment of applications for research infrastructure).

Our consultation reveals that most health research funders (e.g., Wellcome, NIH, NIHR, UKRI MRC) involve patient representatives in at least some of their funding. Funders observe that this helps panel members assess whether the applications consider patient needs. STFC sometimes involves technical professionals and project management experts in the assessment process of applications for new major projects and project technology development funding and finds that this adds valuable information to assess the feasibility of proposed operational costs and other technical details. Relatedly, the involvement of non-academic assessors is seen as a way to overcome a perceived bias against applied research in traditional peer review.

Non-academic users are now widely used in programmes where societal or economic impact are important objectives and recommended in grey literature texts. Practically, this intervention can be implemented in a staggered review process, for example with a traditional academic peer review followed by a more diverse panel with a greater focus on relevance and impact.

Funders report that this helps improve panel discussions, understanding of the context of use (e.g. industry), and the quality of the assessment. For example, EPSRC and Norway's RCN use industry reviewers in programmes that support collaborative research and aim to deliver academic outcomes and also benefit industry partners. Using industry reviewers helps assess applications that cover industry motivation and potential commercial outcomes of the proposed projects. Feedback from applicants shows that they feel better understood when industry reviewers are involved. Although it is hard to attribute programme success to this process element, impact and process evaluations show that the programmes have succeeded in selecting the right applications that align with programme objectives.

Despite the overall positive verdict, non-academic considerations are not appropriate in all contexts, e.g. in the context of pure basic research funding schemes.

Our research also highlights a common view that there is a risk of bias in the selection of industry representatives from large enterprises in specialised roles, rather than from SMEs. As such, pools of non-academic assessors may not be representative of the wider business population (this may not necessarily be a problem depending on scheme aims, e.g. if it only targets certain types of businesses).

The value of non-academic reviewers may also be limited if the objects and types of impact sought by the funding schemes are unspecific and too open-ended.

There is evidence that some academic reviewers believe they are sufficiently aware of the wider context in which research is used to assess proposals. Some also perceive the role of industry assessors negatively, potentially blocking worthy applications due to a lack of understanding of the academic context.

Interviewed funders report some difficulties in finding/recruiting non-academic reviewers. There are limited incentives to complete reviews in academia, and there are almost none in the industry or other sectors, which can be challenging. The funders use systems to find academic reviewers that are not always appropriate for finding non-academic reviewers. Therefore recruiting non-academic reviewers may require substantial additional effort for research funders.

There is, in short, some difficulty around this intervention, meaning it is important to consider carefully when to use it. However, as a means of increasing relevance and strengthening the science/society interface it has significant importance.

References	Interviewees & survey responses
• Luo, J., Ma, L., & Shankar, K. (2021). Does the inclusion of non-academic reviewers make any difference for grant impact panels? Science and Public Policy, 48(6), 763–775. • Gallo, S. A., Sullivan, J. H., & Glisson, S. R. (2016). The Influence of Peer Reviewer Expertise on the Evaluation of Research Funding Applications. PLoS ONE, 11(10), e0165147 • Guthrie, S. et al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, URL: https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf • Guthrie, S. (2019), Innovating in the Research Funding Process: Peer review alternatives and adaptations, AcademyHealth, https://academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf • Bendiscioli, S., Firpo, T., Bravo-biosca, A., Czibor, E., Garfinkel, M., Stafford, T., & Wilsdon, J. (2022). The experimental Research funder's handbook. Research on Research Institute (RoRI). • Kolarz, P., Bryan, B., Farla, K., Krčál, A., Potau, X., & Simmonds, P. (2018). UKRI Research and Innovation Funding Service (RIFS) visioning work - Final report. Technopolis Group, 18 May 2018 • Kolarz, P. et al. (2018), How Research Funders Ensure the Scientific Legitimacy of their Decisions: Investigation in support of the design of Formas scientific management, Digitala Vetenskapliga Arkivet, URL: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1307597&dswid=-4867 • Schouten, R. et al. (2023), Evaluation of research proposals by peer review panels: broader panels for broader assessments? Science and Public Policy, 00, 1-14. • Shepherd, J., Frampton, G. K., Pickett, K., & Wyatt, J. C. (2018). Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. PLoS ONE, 13(5), 1–26.	• 3 survey responses • 5 interviews

5.19 Virtual panels

Convening panels online rather than in person.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Save costs and carbon-footprint; ensure more international panellists, generally remove barriers to participation	Potentially less robust or detailed discussion, though this is unclear	***

Online panels saw drastically increased use during the Covid-19 pandemic to overcome travel restrictions and lockdowns. More broadly, online panels can help secure participation of international panel members. At a general level, online panels aim to reduce costs and environmental impact of international (and even national) panellists travelling. Panellists with caring responsibilities or any other travel-limitations are also usually more able to participate in virtual panels.

Online panels were widely adopted by CRUK since the pandemic, and this has resulted in increased participation of international assessors. Other examples report cost reductions and greater diversity of panels. The NSF for example experimented with virtual panels in 2010, and an article in Science reports cost savings of $10,000 per panel.

There is a perceived need with virtual panel meetings to provide especially clear briefing beforehand. Some consultees see a risk of lower engagement and therefore shorter discussions compared to face-to-face panels, which can also be seen as positive in some cases with a very focused discussion.

Although use of this intervention has increased very recently, there is a lot of positive feedback and agreement around its effectiveness. However, virtual panels have seen a substantial increase in use during the Covid-19 pandemic, bringing them to the attention of many more stakeholders.

Much of our literature does not cover these recent experiences and it is possible that this intervention has more detractors than the pre-pandemic literature suggests. We note as one example the Irish Research Council's Laureate Award scheme, which shifted its panel meetings online during the pandemic and was reviewed shortly thereafter. The report surveyed panellists on the experience and whether online panels should be mainstreamed in future. While the mean opinion is reported to be in the range of 'neutral' to 'somewhat in favour', the review notes a broad range of different sentiments, indicating the need for more research and consultation on the matter.

References	Interviewees & survey responses
• Bieri, M., Roser, K., Heyard, R., & Egger, M. (2021). Face-to-face panel meetings versus remote evaluation of fellowship applications: Simulation study at the Swiss National Science Foundation. BMJ Open, 11(5), 1–8. https://doi.org/10.1136/bmjopen-2020-047386 • Bohannon, J. (2011) Meeting for Peer Review at a Resort That's Virtually Free. Science. 7 Jan 2011, Vol 331, Issue 6013, p. 27. https://www.science.org/doi/10.1126/science.331.6013.27 • Gallo, S. A., Carpenter, A. S., & Glisson, S. R. (2013). Teleconference versus Face-to-Face Scientific Peer Review of Grant Application: Effects on Review Outcomes. PLoS ONE, 8(8). https://doi.org/10.1371/journal.pone.0071693 • Guthrie, S. et al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, URL: https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf • Guthrie, S. et al. (2018), What do we know about grant peer review in the health sciences?, F1000Research, 6(1335), https://doi.org/10.12688/f1000research.11917.2	• 1 survey response • 3 interviews

- Guthrie, S. (2019), Innovating in the research funding process: peer review alternatives and adaptations, AcademyHealth, URL: https://academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf

- Kolarz, P., Arnold, E., Cimatti, R., Dobson, C. & Seth, V. (2020). Process review of the IRC Laureate Award. Irish Research Council. https://research.ie/assets/uploads/2021/02/IRC-Laureate-review-FINAL.pdf

- Recio-Saucedo, A. et al. (2022). What works for peer review and decision: making in research funding: a realist synthesis, Research Integrity and Peer Review, 7(2), https://doi.org/10.1186/s41073-022-00120-2

- Shepherd, J. et al. (2018), Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. PLoS ONE 13(5): e0196914. https://doi.org/10.1371/journal.pone.0196914

6 Main findings: Interventions to the shape of decision-making

6.1 Wildcard

Sometimes also known as 'Golden ticket' or 'Joker'. Each panel member (or other decision-maker) is able to select one proposal (e.g. per call, per year, or similar) to guarantee funding (provided there is no conflict of interest), regardless of panel rankings or other decision-making processes.

Main intended aim(s)	Main hazard(s)	Evidence strength
Fund riskier, transformative ideas; save debating time in panels	Open to abuse if conflicts of interest are not monitored very well. Requires anonymised reviewing	***

This is an intervention aimed mostly at increased funding for new and riskier ideas. The underlying assumption is that panels tend towards conservatism supported by the finding that often a single poor review may mean an application is rejected. This intervention provides a way of circumventing this type of 'group-think'.

A secondary aim of this intervention is to reach funding decisions more rapidly. Especially controversial applications (i.e. with some very positive and some very negative reviews) tend to take up a considerable time in panel meetings. A 'wildcard' option means occasionally ending long discussions where agreements seemingly cannot be reached.

There are three known instances of implementation (at Volkswagen Foundation, FWF and Villum Foundation). At Volkswagen Foundation and Villum Foundation outcomes were generally as hoped. Awarded applicants included greater numbers of young and early career researchers, and selected proposals included ones which would not have been awarded based on ranked scores. At FWF, no reviewers chose to apply their 'wildcard' in any of the three funding calls where it was used.

At the VW foundation, only 11 out of 183 possible grants (6 %) have been awarded on the basis of a wildcard. One important effect of the wildcard option was to save time in the meetings, when two opposing opinions could not be resolved by further deliberation.

There is, however, a strong risk of cronyism, which means that conflicts of interest need to be monitored extremely carefully, and anonymised reviewing needs to accompany schemes where a 'wildcard' system is used. Both are likely necessary as even in anonymised reviewing, peer reviewers / panellists may still be able to infer the identity of the applicant based on the topic or approach. In addition, giving a panellist the power to outright select an application conflicts with interventions targeting subjectivity in the selection process via training.

The literature also tends to pair 'wildcard' systems with anonymised reviewing, so positive findings (e.g., increased confidence to submit 'braver' ideas than usual) is likely contributed to by this double approach; the intervention has also been used only in experimental schemes thus far, setting a contextual predisposition to riskier research.

This appears to be a somewhat controversial approach: among the sources available to us, strengths and risks are variously emphasised, with some positive and some negative verdicts.

References	Interviewees & survey responses
- Avin, S. (2019). Mavericks and lotteries. Studies in History and Philosophy of Science Part A, 76(January 2018), 13–23. https://doi.org/10.1016/j.shpsa.2018.11.006 - Bendiscioli, S., & Garfinkel, M. (2021). Dealing with the limits of peer review with innovative approaches to allocating research funding. EMBO Science Policy Programme.	- No survey responses 3 interviews

- García-Ruiz, J. M. (2019). I won a project! Substantia, 3(1), 5–7. https://doi.org/10.13128/Substantia-182
- Guthrie, S. at al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf
- Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: a realist synthesis. Research Integrity and Peer Review, 7(1), 1–28. https://doi.org/10.1186/s41073-022-00120-2
- Sinkjær, T. (2012). Fund ideas, not pedigree, to find fresh insight. Nature, SEPTEMBER, 1. https://doi.org/10.4135/9781412971959.n413

6.2   Partial randomisation

Successful proposals are chosen at random. In most methodologies, randomisation is only partial. For example, proposals may be scored and sorted into bands, and only those on the border of being funded will be randomised.

Main intended aim(s)	Main hazard(s)	Evidence strength
Remove bias, reduce panel burden	Reputational impact on applicants	***

Literature, survey and interviews all suggest that partial randomisation aims to remove bias (both against demographic factors and riskier ideas), and to reduce administrative burden in the selection process. Mostly the burden is mentioned in connection to ranking, but the literature suggests that it has also been used (in connection with other interventions) to enable shorter applications. Use of partial randomisation is justified by increasingly overwhelming evidence that while peer/panel review reliably identifies the very highest quality applications, as well as the 'tail' of unsuitable low-quality ones, it tends towards arbitrary decision-making in the 'upper-midfield' of the quality spectrum.

Evidence on this intervention is mainly from observations from real life applications, some of which have been assessed for diversity and applicant satisfaction. However, in most cases it is too early to say anything about effect on the actual nature of the funded research. The approach is further supported with statistical analysis suggesting arbitrariness in the traditional peer review process.

The data collection identified at least six research funding bodies where partial randomisation has been used. Some assessments have been carried out on the impacts of the intervention, and at least two funders (Volkswagen Foundation and SNSF) were found to have diversified their awardee pool. In addition, applications at BA, FWF and VWF were found to increase in response to the partial randomisation introduction. In the case of VWF, this was reportedly due to a perceived higher chance of success among applicants.

We identify two main concerns; first, (from the funders' point of view) there is the risk of awarding lower quality or less relevant awards. Second (from the applicants' point of view), there is a concern of reputational impact from both rejections or successes. The first concern is inevitable but can be mitigated by narrowing down the pool of applications to those where finding consensus among reviewers and panellists is challenging (genuinely poor-quality applications will at this stage already have been sifted out). This is typically the approach taken, and use of the term 'partial randomisation' rather than simply 'lottery' is generally preferred in order to emphasise this point.

The second concern has been approached differently: at VWF, applicants were concerned about crediting their awards (if successful) to randomisation, which was mitigated by the added use of wildcards and not disclosing which applications were awarded via what method. FWF also used wildcards in combination with partial randomisation; however, in the three calls of its 1000 Ideas programme, no reviewers used the wildcard option. FWF thinks this is because panel members were concerned about their reputation in case other jury members disagreed about the value of the application supported by the wildcard. Conversely, at SNSF, all applicants were informed if partial randomisation was used in both rejection letters and award letters to ensure transparency.

While there is controversy around this intervention in general terms, all evidence on implementation is fairly positive. From the UK, we also have anecdotal evidence that the academic community's response to NERC's partial randomisation trial has been overwhelmingly positive.

We note also that there is considerable versatility in application, for instance in terms of how conservatively the process is used: RCN has thus far only randomised the selection of applications identical either in idea or scoring. It can (and often is) paired with applicant anonymisation to fully avoid bias (as some degree of filtering applications takes place in all instances of implementation). VWF relies on a quadruple process of anonymisation > outlines to full applications > wildcards > partial randomisation of those not outright selected but of good quality.

References	Interviewees & survey responses
• Avin, S. (2019). Mavericks and lotteries. Studies in History and Philosophy of Science Part A, 76(January 2018), 13–23. https://doi.org/10.1016/j.shpsa.2018.11.006 • Bendiscioli, S. (2019). The troubles with peer review for allocating research funding. EMBO Reports, 20(12), 1–5. https://doi.org/10.15252/embr.201949472 • Bendiscioli, S. et al. (2022), The experimental research funder's handbook, Research on Research Institute, URL: https://rori.figshare.com/articles/report/The_experimental_research_funder_s_handbook_final_version_/19459328 • ESRC. (n.d.). Might randomised funding increase ESRC investigator diversity? An analysis of the 'ESRC Research Grants' call. • Fang, F. C., Bowen, A., & Casadevall, A. (2016). NIH peer review percentile scores are poorly predictive of grant productivity. ELife, 5, 1–6. https://doi.org/10.7554/eLife.13323 • Fang, F. C., & Casadevall, A. (2016). Research funding: The case for a modified lottery. MBio, 7(2). https://doi.org/10.1128/mBio.00422-16 • Guthrie, S. et al. (2013), Alternatives to Peer Review in Research Project Funding, RAND, https://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf • Guthries. S. (2019), Innovating in the research funding process: peer review alternatives and adaptations, AcademyHealth, https://academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf • Höylä, T., Bartneck, C., & Tiihonen, T. (2016). The consequences of competition: simulating the effects of research grant allocation strategies. Scientometrics, 108(263–288). • Jerrim, J. & de Vries, R. (2020), Are peer-reviews of grant proposals reliable? An analysis of Economic and Social Research Council (ESRC) funding applications, The Social Science Journal, 60(1), 91-109, https://doi.org/10.1080/03623319.2020.1728506 • Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). Study on the proposal evaluation system for the EU R&I framework programme. European Commision. https://doi.org/10.2777/16211 • Roumbanis, L. (2019). Blind Luck – Could lotteries be a more efficient mechanism for allocating research funds than peer review? LSE Impact Blog.	• 1 survey response • 6 interviews

https://blogs.lse.ac.uk/impactofsocialsciences/2019/12/11/blind-luck-could-lotteries-be-a-more-efficient-mechanism-for-allocating-research-funds-than-peer-review/

- Roumbanis, L. (2019). Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-making Mechanisms for Allocation of Research Grants. Science Technology and Human Values, 44(6), 994–1019. https://doi.org/10.1177/0162243918822744
- VolkswagenStiftung. (n.d.). Partially Randomized Procedure - Lottery and Peer Review. https://www.volkswagenstiftung.de/en/partially-randomized-procedure-lottery-and-peer-review
- Walsweer, T. (2018). Give Chance a Chance. VolkswagenStiftung, May 22, 2018. https://www.volkswagenstiftung.de/en/news/story/give-chance-chance

6.3   Scoring mechanisms

Includes calibration of scores, consensus vs. voting, weighting.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase relevance of funded projects to the aims; improve review quality/reliability	None confirmed but may disadvantage high-risk/ high-reward applications	****

Consulted funders and literature point to two main variants of this intervention. The first involves applying equal weighting of some criteria to meet the specific needs of the funding scheme (e.g., wider/non-academic impact, novelty). Reviewed literature and survey respondents provided examples of the use of equal weighting of scientific merit and impact, making the scoring matrix and calibration of scores more quantitative /absolute. Consulted funders claim that the intervention was successful in making sure the right applications were funded considering the importance of impact. The intervention also appears to have been successful in that panels followed the criteria and funder instructions instead of any other considerations (e.g. applying/balancing criteria as they would in 'ordinary' schemes).

In the second variant of this intervention, reviewers may apply their own interpretation of criteria and weightings when scoring proposals. According to the literature, calibration of scores (disclosure of scores and discussion to calibrate scores between reviewers) has been found to have the effect of converging scores within a panel, but not an increase in relatability overall (as tested in experiments with multiple panels scoring the same proposals).

Our research has not identified any known hazards of this intervention.

The literature points out that, depending on the situation, the two objectives for the use of this intervention (increase reliability and meet specific funding needs) might lead to opposite recommendations. For instance, to increase reliability, one might recommend calibration and elimination of outliers, whereas to identify and fund novel research, one might want to prioritise proposals with highly variable scores.

References	Interviewees & survey responses
- Abdoul et al. (2012), Peer Review of Grant Applications: Criteria Used and Qualitative Study of Reviewer Practices, Plos one - Guthrie et al. (2013), Alternatives to Peer Review in Research Project Funding - 2013 Update, RAND Europe, Prepared for the UK Department of Health-funded Centre for Policy Research in Science and Medicine (PRiSM) - John Jerrim; Robert de Vries (2019), Are peer-reviews of grant proposals reliable? An analysis of Economic and Social Research Council (ESRC) funding applications, - Lee (2015), Commensuration Bias in Peer Review, Philosophy of Science, Vol. 82, Issue 5, pp. 1272-1283	- 2 survey respondents - No interviews

• Pier et al. (2017), Your comments are meaner than your score': score calibration talk influences intra- and inter-panel variability during scientific grant peer review, Research Evaluation, vol. 26, Issue 1, pp. 1-14
• Viner (2004), Panel Reproducibility: Results of a trial in October 2004, Presentation

6.4 Sequential application of criteria (rather than simultaneous application)

A proposal is scored for one set of criteria, ranked and a cut-off point determined. Then those above the cut-off point are assessed again for another set of criteria to determine the final funded list.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Ensure application of all criteria, increase relevance to programme aims	None known	**

This intervention is typically related to two-stage approaches and/or pre-proposals. The literature shows that funders can use this for programmes with a complex set of aims (typically research excellence as well as non-academic relevance), where assessing based on one set of criteria first can help reduce the burden of subsequent rounds.

The Dutch NWO's 'Veni' programme required submission of a CV and track record along with an initial short proposal. This first round thus included an assessment of the set of criteria related to the researchers' qualities, allowing a reduction of the number of applications progressing to the subsequent round focussed on an assessment of the proposal and potential impact. This is reported to have worked well and reduced the assessment time of reviewers by 25%.

We note at this point the evidence on this form of criteria application is too limited to arrive at strong verdict. However, it needs to be considered alongside the aforementioned two-stage approaches and/or pre-proposals, which often implicitly take this approach (however they do not always do so, hence we mention this intervention here separately.

References	Interviewees & survey responses
• Calatrava Moreno et al. (2019), Science Europe Study on Research Assessment Practices - Final Report, Technopolis Group • Kolarz et al. (2018), UKRI Research and Innovation Funding Service (RIFS) visioning work - Final report, Technopolis Group • Kolarz, P. et al. (2018), How Research Funders Ensure the Scientific Legitimacy of their Decisions: Investigation in support of the design of Formas scientific management, Digitala Vetenskapliga Arkivet, URL: • http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1307597&dswid=-3985	• No survey responses or interviews

6.5 Use of quotas

After ranking, proposals are reviewed to ensure sufficient numbers in certain categories including quotas related to protected characteristics, place, first-time applicants, etc.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Avoid/counteract bias and 'clustering'	Very drastic approach	*

Quotas are a means of avoiding clustering of investments in places or themes and ensuring equitable success rates among disadvantaged and minority researcher populations.

Our research indicates that funders use quotas to ensure diversity among reviewers and panel members (see also intervention on 'embedding EDI in assessment' below). Some literature items discuss applying quotas as an option at the decision point, but we found no evidence of implementation. Some literature items point to this measure being too drastic and that funders can achieve diversity through other means (e.g. working with underrepresented groups, partial randomisation, anonymisation).

References	Interviewees & survey responses
• Comfort, N. (2021). Addressing Racial Disparities in NIH Funding. Journal of Science Policy & Governance, 18(04). https://doi.org/10.38126/jspg180408 • Cruz-Castro, L., & Sanz-Menéndez, L. (2019). Grant Allocation Disparities from a Gender Perspective: Literature Review. Synthesis Report. 1(November), 84. http://hdl.handle.net/10261/200024 • De Peuter, Steven & S. Conix (2022) The modified lottery: Formalizing the intrinsic randomness of research funding, Accountability in Research, 29:5, 324-345	• 1 survey response • No interviews

7 Main findings: Interventions in training and feedback

7.1 Bringing in reviewers from earlier careers & providing mentoring

Panels and reviewers tend to be very experienced researchers/innovators. Those early in their careers could be invited to review or be part of panels with additional training, bringing different perspectives and experiences. Previous calls' award winners may also be brought in as reviewers/panellists.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Improve review quality, diversify reviewers	None known	**

Funders use this intervention to improve review quality, reduce burden, diversify the pool of reviewers and provide career support to early career researchers (ECRs). Most consulted funders involve ECRs in the peer review and one (CRUK) invites ECRs to observe panels and committees. One funder (Wellcome) has specific targets for the number of ECRs on its panels.

All consulted funders report positive feedback from involved ECRs. The experience helped them to learn about the process, improve their grant writing skills and made the assessment process more transparent.

All funders report significant interest and demand from ECRs to be involved in the peer review. This results in improved ability of funders to secure reviewers (because of the larger pool available). Similarly, all sources available to us report that ECRs provide very good quality reviews and are very enthusiastic. Though most evidence is based on funder observation rather than controlled experiments, there is general agreement on the effectiveness of this intervention.

References	Interviewees & survey responses
• None.	• 3 survey responses • 3 interviews

7.2 Embedding EDI in assessment

Training or support provided to make assessors aware of their unconscious biases and to encourage them to call each other out during the assessment process.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Reduce bias, increase diversity among awardees	Ineffective training may install a false sense of confidence	**

Funders introduce this intervention to reduce bias, enable fair decisions and improve diversity in the funded portfolio.

Consultees pointed out that there is no way to remove bias entirely, but they feel that highlighting potential issues (via training) helps. It is difficult to demonstrate the effectiveness of this intervention because it is hard to attribute positive changes to a single intervention (hence the relatively low star-rating for this intervention despite many sources). But anecdotal feedback from panel members is that it does make them question their biases and decisions. For example, since anti-bias training for the Austrian FWF's board ('Kuratorium'), every board meeting now starts with a one-slide reminder about bias and the need to call it out.

Submissions and success rates by demographics are periodically reviewed in organisations including the UK Research Councils and in different funding institutions in Canada, finding improvements in diversity of the funded portfolio (though as noted, causality is difficult to confirm).

We find only anecdotal evidence from funders, in some cases based on monitoring data but it is a challenge to attribute change to a single intervention. There is nevertheless broad consensus about the relevance of this intervention despite difficulty attributing change directly.

Available material discussing the effectiveness of this type of training offers mixed views and cautions against blind reliance on it. Training approaches vary greatly, and ineffective training may harmfully install a false sense of confidence. The reliability of self-reported results has also been questioned. Other points of criticism assess its impact at the institutional level and note that bias training should only form part of a more holistic approach addressing the bigger picture. That said, there is little evidence addressing unconscious bias training in grant peer review specifically (the majority of the material on the effect of unconscious bias training focuses on general professional or healthcare settings).

We note that there are several aspects to EDI, and a wide range of techniques that may be implemented. Some additional forms of embedding EDI at UKRI have included reducing the number of proposals at panel and increasing the number of breaks to help with cognitive overload, silent scoring, and management of any unacceptable behaviours and/or comments at the panel. Some councils have also had diversity panel targets in place since May 2016 (e.g. panels aim to meet at least 30% representation of the underrepresented gender) and there are gender and ethnicity targets to increase diversity of perspective in assessment.[14] These targets have since been met or surpassed.

While not fully related to this intervention, we note that all interviewed funders try to ensure balanced panels, and two funders (Wellcome and CRUK) have introduced diversity targets and quotas. Wellcome achieved the targets in 2022, which has been a big success, as it brings in a much broader breadth of voices. Wellcome did not have any challenges securing diverse membership, but there is a tendency to go to the same reviewers. Regarding the impact this intervention has had on the diversity of the portfolio, it is too early to tell. Wellcome has not seen a significant shift in, for example, the proportion of ethnic minority groups. The measure was introduced because it was deemed the right thing to do and to improve the diversity of voices. This intervention has become a new normal at Wellcome and will not change.

References	Interviewees & survey responses
- Dias Lopes, A., & Wakeling, P. (2022). Inequality in Early Career Research in the UK Life Sciences. University of York. https://www.ukri.org/wp-content/uploads/2022/11/BBSRC-301122-BBSRCInequalityInECRReport.pdf - EPSRC. (2020). Understanding our portfolio: A gender perspective. https://www.ukri.org/publications/epsrc-understanding-our-portfolio-a-gender-perspective/ - Fitzgerald, C., Martin, A., Berner, D. & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review. BMC Psychology. 7:29. https://bmcpsychology.biomedcentral.com/articles/10.1186/s40359-019-0299-7	3 survey responses 4 interviews

---

[14] For current guidance, see https://www.ukri.org/what-we-offer/supporting-healthy-research-and-innovation-culture/equality-diversity-and-inclusion/epsrc/evolving-and-upholding-fairness-in-peer-review/

- Gallo, S. A., Sullivan, J. H., & Croslan, D. J. R. (2022). Scientists from Minority-Serving Institutions and Their Participation in Grant Peer Review. BioScience, 72(3), 289–299. https://doi.org/10.1093/biosci/biab130
- Green, T. & Hagiwara, N. (2020). The Problem with Implicit Bias Training. Scientific American. https://www.scientificamerican.com/article/the-problem-with-implicit-bias-training/
- Hillier, E., Razack, S., Clary, V., & Münter, L. (2018). The State of Women in Life Sciences at McGill: A Summary and Report of the Win4Science Forum. https://www.mcgill.ca/pharma/files/pharma/final_report.win4science_report_-_nov_30_2018.pdf
- Kolehmainen, C., & Carnes, M. (2018). Who Resembles a Scientific Leader – Jack or Jill? How Implicit Bias Could Influence Research Grant Funding. Circulation. 137(8). 769.
- Pritlove, C., Juando-Prats, C., Ala-Leppilampi, K. & Parsons, J. (2019), The good, the bad, and the ugly of implicit bias. The Lancet. 393. https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(18)32267-0.pdf
- Roper, R. L. (2019). Does Gender Bias Still Affect Women in Science? Microbiology and Molecular Biology Reviews, 83(3). https://doi.org/10.1128/mmbr.00018-19
- Tamblyn, R., Girard, N., Qian, C. J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. Cmaj, 190(16), E489–E499. https://doi.org/10.1503/cmaj.170901
- The Behavioural Insights Team (2020). Unconscious bias and diversity training – what the evidence says. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/944431/20-12-14_UBT_BIT_report.pdf
- Yip, P. S. F., Xiao, Y., Wong, C. L. H., & Au, T. K. F. (2020). Is there gender bias in research grant success?: Hong Kong as a case study. Humanities and Social Sciences Communications, 7(1), 1–10. https://doi.org/10.1057/s41599-020-00656-y

7.3 Expanding or reducing feedback to unsuccessful applicants

Different levels of feedback may be provided on unsuccessful applications.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Improve transparency, improve applicants' learning from unsuccessful applications	Added burden; feedback may be of inconsistent quality	**

Consulted funders use feedback mainly to explain decisions and thus increase transparency of the assessment process. Some share feedback only if an applicant requests it. A secondary aim is to ensure there can be a better learning process for unsuccessful applicants.

Literature on the subject is limited but one study shows that well-developed, good quality feedback helps applicants to improve the quality of future applications. Consulted institutions reported that in rare instances when unsuccessful applicants receive feedback, it is very helpful and encouraged the funders do so more.

One funder changed the presentation of feedback by sending panel members' written comments verbatim instead of a summary of the panel discussion. This was not effective because applicants received several sets of comments which can conflict with each other, making it difficult for applicants to understand the rationale for the decision on their application.

A notable hazard is that it is hard to be consistent and equitable with the type of feedback given, as quality of feedback may differ at least slightly. Additionally, consolidating, checking and distributing feedback creates additional burden for the funder. There is therefore a trade-off here between transparency and learning on one hand and reduced burden on the other.

Evidence is mostly anecdotal, based on funder observations and one survey of applicants. However, there generally seems to be appetite for more feedback on unsuccessful application. Given the added burden, there is a case to consider carefully whether feedback is more useful in some funding schemes or for some applicant types than for others.

References	Interviewees & survey responses
• Derrick, G. E., Zimmermann, A., Greaves, H., Best, J., & Klavans, R. (2022). Targeted, actionable and fair: reviewer reports as feedback and its effect on ECR career choices. https://osf.io/preprints/socarxiv/a8psh/ • Rodriguez-Rincon, D. Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S. & Guthrie, S. (2021), Study on the proposal evaluation system for the EU R&I framework programme, European Commission, https://doi.org/10.2777/16211	• 1 survey response • 5 interviews

7.4　Funder representation on review panels

The funder is represented on the panel to guide discussion or provide briefing on programme aims. Their role is beyond a purely administrative function, they may even be in a chair-role or similar.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Ensure guidance is followed; may help ensure relevance of decisions to scheme aims	None known	**

Funders are usually represented on review panels to ensure the panels follow the guidance and to document the process but not in an advisory or chair role. In this intervention, representatives of the funder take a more active role in communicating scheme aims and ensuring review and discussion stays focused on the scheme's main criteria. This may happen at the start of panel meetings, but may also involve prior briefings, as well as reminders and steering while discussion of applications is taking place. However, even within this intervention, funder representatives generally do not have a role in making the funding recommendations as such (i.e. they may steer but they do not have a 'vote').

We note that 'funder representation' is a term that is somewhat open to interpretation. For example, in the Austrian FWF's Emerging Fields scheme, the FWF board is involved in the decision-making at several points in the multi-stage decision-making process. Its members are based at various research performing organisations. However, the Executive Board president chairs these board meetings. This individual has a strong academic track and experience but is closely familiar with the funder's strategy and operations. The Emerging Fields scheme is currently subject to evaluation and so the effects of this form of funder representation are at this point unknown.

A more 'clear-cut' type of funder representation on panels occurs at the Human Frontier Science Program (HFSP), which has undergone a full organisational and process review recently. The review found that, in line with its stated objectives, the HFSP process successfully identifies the most innovative, 'frontier' research ideas and recommends them for funding. However, the review further found that it relies primarily on culture rather than process structure to achieve this, and that this 'HFSP-culture' is in part perpetuated through the presence and input of secretariat staff at the panel meetings. While secretariat staff are not involved in the decision-making itself, they ensure through briefing both the panel in general and new panellists individually about the purpose of the programme and the emphasis on 'frontier' research that panellists are expected to identify and reward.

A further example worth noting (though it does not constitute 'funder representation' in the strict sense) is the ESRC's Transformative Research scheme. In its third round, an awardee from the first round two years prior was selected as panel chair, with the aim of ensuring a cultural understanding of the scheme aims (and therefore of the criteria and how select applications) would be ingrained in the panel as much as possible. While this chair did not represent the funder (ESRC) as such, their previous involvement with the scheme meant that they could be an important voice to communicate the scheme aims to the rest of the panel. Evaluation of the scheme found that, alongside other process-innovations, this selection of panel chair played an important role in maintaining the panel's focus on the 'transformative' element of submitted applications.

While evidence on this intervention is relatively limited, there is agreement among consultees that funder presence on panels helps to ensure panels follow the funder guidance and thus improves the quality of the assessment, and some evaluative evidence points in the same direction.

References	Interviewees & survey responses
• Kolarz, P., Arnold, E., Farla, K., Gardham, S., Nielsen, K., Rosemberg, C., & Wain, M. (2016). *Evaluation of the ESRC Transformative Research Scheme*. Technopolis Group, July 2016. https://esrc.ukri.org/files/research/research-and-impact-evaluation/esrc-transformative-research-scheme-evaluation/ • Kolarz, P., Davé, A., Bryan, B., Urquhart, I., Rigby, J. & Suninen, L. (2023) *Organisational and Process Review of the Human Frontier Science Program*. Technopolis Group. Published by HFSP: https://www.hfsp.org/node/74873#book/	• No survey responses • 2 interviews

7.5 Improving quality of reviews

Through training/retaining good reviewers/recognition. Peer review colleges fit here too.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Improve quality of reviews; simplify training; increase response rate for review requests	None known	****

Funders use training and peer review colleges to improve quality of reviews. The literature on this intervention mostly uses reviewer agreement as a proxy for improved quality of the review.[15] Peer review colleges are also seen as a tool to address peer-review fatigue and to increase reviewer response rates.

One controlled trial at the US National Institutes of Health showed that a training programme to increase inter-rater reliability improved scoring accuracy and reviewer agreement. Consulted funders also report that use of a peer review college provides a large number of reviewers to approach initially that are familiar with the scheme and have a proven track record of providing good reviews. Additional training has been useful and that has been developed based on common review errors.

An ongoing example of the above can be found at EPSRC, whose peer review college consists of more than 6000 members, all of whom have undergone online training upon joining the

---

[15] However, some sources speculate about high disagreement being an indication of high-risk/ high-reward nature of an application. We note this to indicate that, while the overall evidence base on this intervention is strong, there is some disagreement about whether this common form of measuring its success might have some limitations.

college. The training, among other factors in the membership, is stated to help members to increase their knowledge of proposal writing and reviewing.

Funders note that they receive more reviews per request from college members compared to 'cold' peer review invites as well as a higher percentage of reviews of suitable quality.

Evidence on this intervention is fundamentally strong – there are controlled trials reported in literature (about training) and funder observations and monitoring data on positive responses to peer review requests and improved review quality.

References	Interviewees & survey responses
• European Science Foundation. (2011). ESF Survey Analysis Report on Peer Review Practices. 60. http://www.esf.org/fileadmin/Public_documents/Publications/pr_guide_survey.pdf • Guthrie, S. et al. (2018), What do we know about grant peer review in the health sciences?, F1000Research, 6(1335), https://doi.org/10.12688/f1000research.11917.2 • Hesselberg, J.-O., Fostervold, K. I., Ulleberg, P., & Svege, I. (2021). Individual versus general structured feedback to improve agreement in grant peer review: a randomized controlled trial. Research Integrity and Peer Review, 6(1), 1–11. https://doi.org/10.1186/s41073-021-00115-5 • Publons. (2018). Global State Of Peer Review 2018. Publons and Clarivate Analytics. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjy8o-Tmg_9AhWMtKQKHY9-ASAQFnoECAQQAQ&url=https%3A%2F%2Fpublons.com%2Fstatic%2FPublons-Global-State-Of-Peer-Review-2018.pdf&usg=AOvVaw3xjM8vzC-KkdNCHrGDT3iv • Recio-Sauceda, A. et al. (2022), What works for peer review and decision-making in research funding: a realist synthesis, Research Integrity and Peer Review, 7(2), https://doi.org/10.1186/s41073-022-00120-2 • Sattler, D. N., McKnight, P. E., Naney, L., & Mathis, R. (2015). Grant peer review: Improving inter-rater reliability with training. PLoS ONE, 10(6), 1–9. https://doi.org/10.1371/journal.pone.0130450 • Shepherd, J. et al. (2018), Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. PLoS ONE 13(5): e0196914. https://doi.org/10.1371/journal.pone.0196914	• 2 survey responses • 3 interviews

7.6 Open review/rebuttal

Reviews are published and/or made available to the applicant before funding decisions are taken, so they can be viewed and responded to.		
Main intended aim(s)	Main hazard(s)	Evidence strength
Increase accountability and review quality	Possibly increased burden for funder (and longer timelines depending how rebuttal works)	***

Open reviews and rebuttals have been particularly well-known elements in journal peer review but are also becoming recognised as potential tools for grant peer review. Open peer review is considered an umbrella term inclusive of open identities, open reviews and open interaction. They are expected to lead to increased accountability, challenging unjust reviews, giving applicants more voice in the process and increase overall review quality. The latter is particularly enabled by applicants being able to clarify in case reviewers have genuinely misunderstood some of the application's content, which may be especially important where English is not the applicant's first language. Open identities are also hoped to contribute to the credit of the reviewer.

Consulted funders were positive about the intervention as it is well received by the applicants and reviewers and helps to increase the transparency of the process.

There are, however, some opposing voices. Concerns have been raised about the potential for reduced rigour and valid criticism where the identities of the reviewers are made known. Literature also points to a potential (but not evidenced) increase in burden, though consulted funders did not raise this concern.

A prominent example of the ongoing use of rebuttals is at NWO (Dutch Research Council), where reviews are shared with applicants, who then have one week to produce a short rebuttal to reflect on issues raised by reviewers. These rebuttals will be reviewed along with the review by the panel. In other words, the rebuttals may influence the funding recommendation. One reviewed study suggests that this rebuttal stage may have a corrective effect on some degree of gender bias in the review process.

A similar process is also used at UKRI. There is a 10 working day turnaround time for lead applicants to provide a rebuttal (recently extended from five working days to take into account EDI concerns). These may then influence the final funding decision – they are provided to the panel alongside the written reviews to aid their decision making.

References	Interviewees & survey responses
• Baždarić, K., Vrkić, I., Arh, E., Mavrinac, M., Marković, M. G., Bilić-Zulle, L., Stojanovski, J., & Malički, M. (2021). Attitudes and practices of open data, preprinting, and peer-review-A cross sectional study on Croatian scientists. PLoS ONE, 16(6 June), 1–18. https://doi.org/10.1371/journal.pone.0244529 • Bol, T., de Vaan, M., & van de Rijt, A. (2022). Gender-equal funding rates conceal unequal evaluations. Research Policy, 51(1), 104399. https://doi.org/10.1016/j.respol.2021.104399 • Cavallaro, M. (2022). Horizon Europe and Plan B research funding: Turning adversity into opportunity. HEPI Policy Note, 38. • D 5.1 Report on "Strategic advice for enhancing the gender dimension of Open Science and Innovation Policy". (2018). GENDer equality in the ERA Community To Innovate policy implementatiON (GENDERACTION). https://genderaction.eu/wp-content/uploads/2018/12/741466_GENDERACTION_D11_Report-on-Strategic-advice-for-enhancing-the-gender-dimension-of-OS-and-Innovation-Policy_approved-1.pdf • Guthrie, S. (2019). Innovating in the Research Funding Process : Peer Review Alternatives and Adaptations - Horizon Scan. Paradign Project, AcademyHealth. https://www.academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf • Hill, Michael (2019). Response by the Author. EMBO Reports, 20Meadmore, K., Fackrell, K., Recio-Saucedo, A., Bull, A., Fraser, S. D. S., & Blatch-Jones, A. (2020). Decision-making approaches used by UK and international health funding organisations for allocating research funds: A survey of current practice. PLoS ONE, 15(11), 1–17. https://doi.org/10.1371/journal.pone.0239757 • Kolarz, P. et al. (2018), How Research Funders Ensure the Scientific Legitimacy of their Decisions: Investigation in support of the design of Formas scientific management, Digitala Vetenskapliga Arkivet, URL: http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1307597&dswid=-3985 • Meadmore, K., Fackrell, K., Recio-Saucedo, A., Bull, A., Fraser, S. D. S., & Blatch-Jones, A. (2020). Decision-making approaches used by UK and international health funding organisations for allocating research funds: A survey of current practice. PLoS ONE, 15(11), 1–17. https://doi.org/10.1371/journal.pone.0239757 • Recio-Saucedo, A. et al. (2022), What works for peer review and decision-making in research funding: a realist synthesis, Research Integrity and Peer Review, 7(2), https://doi.org/10.1186/s41073-022-00120-2	• 1 survey response • 2 interviews

8 Additional interventions identified by our review

Our data collection focused on the 38 interventions that served as a baseline for this review. However, while going through the literature on the set of 38 and running several consultations, we identified some other interventions not included in the initial list. Two of these are recently introduced interventions facilitated by technological advances and increased use of information technology to support the peer review process. A third summarises various actions to improve behaviours and culture/supporting EDI in interviews and panel meetings. We briefly describe these interventions here.

Roving panel members

UKRI Future Leaders Fellowships scheme uses panels of experts from across the research and innovation system to consider all assessment criteria, and the panel members are roving. Panel members change panels through the assessment process to ensure consistency and quality between panels. The intervention is easy to introduce, and according to programme staff, it makes a significant difference in ensuring consistency across the panels.

Discussion boards

BBSRC uses discussion boards (shared virtual platforms for information exchange) to reduce 'on the day' peer review pressure because the discussion has already happened three weeks period before the panel meeting online. In the actual panel meeting, reviewers only have to discuss outstanding issues and agree a ranked list. Discussion boards allow panel members to be flexible with the time they commit to the review. Discussion boards support increased transparency of pre-panel meeting work and discussion and improve benchmarking of scoring before the panel meeting. Our consultees pointed out that discussion boards help to remove the 'corner of the room' discussions that happen at in-person meetings and might not be transparent and cannot be challenged. Furthermore, discussion boards enable clear and detailed feedback to applicants because all discussions are recorded.

Use of videos

Several consulted funders reported using videos as part of the application process. For example, UKRI uses video clips along with short application forms at the EOI stage of the Healthy Ageing Catalysts programme. FWF uses videos for pre-selection before soliciting full applications in its Momentum programme. The programme funds researchers 1-2 years after tenure to provide a boost for their career. The programme received many applications, and FWF decided to introduce pre-applications and used a three-minute video application for the first assessment stage. The accompanying evaluation found no bias in the assessment based on the video format (e.g., showing diagrams, not showing the speaker, etc.), and reviewers were happy with the format. Video format helped to keep the burden for reviewers low.

Improving culture/supporting EDI in interviews and panel meetings

Our consultation also reveals several small interventions or tweaks to the assessment process, all aimed at improving assessment culture and supporting EDI. None of these modifications has the potential to improve the process alone, but in combination with other measures, these small interventions can potentially have a positive impact. For example, UKRI's Future Leaders Fellowships programme introduced silent reflection periods in interviews (adapted from prior use by EPSRC in 2016-17). After the interview, a two-minute silent period is mandated when no one can speak. A silent period helps to stop initial verbal reactions about how good or bad the interview was, which may otherwise affect the rest of the discussion. The silent period is

intended for panel members to reflect on what they have heard and develop reasoning for their grading.

In the same programme, UKRI introduced the numbers-first approach. This means that panel members first give their grades without commentary to avoid the risk of them changing their views and grades because other panel members have different or more loudly expressed opinions. UKRI observed improved panel discussion quality after introducing the above measures.

9 Summary and recommendations

Our headline findings are noted at the outset of this report. However, before we conclude with our list of recommendations resulting from our research, we briefly summarise our research in the table below. It shows how each of the 38 interventions relates to the 7 main aims posited at the start, as well as the main hazards of each intervention, and our evidence strength rating.

Table 3   Summary of aims, hazards and evidence strength

	Intervention	Save time	Increase relevance	Manage app volume	Reduce bias	Reduce burden	Fund high-risk	Increase review qual.	Hazards	Evidence strength rating
Pre-call	Assessment criteria definition		x						Reviewers may not follow guidance; too many criteria risk over-complicating discussions	***
	Demand management: individuals (1)			x	x				Shifts burden to other funders, savings are minimal	*
	Demand management: individuals (2)			x		x			May simply shift re-submission to other funders, somewhat controversial	**
	Demand management: institutions			x		x			Largely shifts burden to institutions; potential additional bias, depending on institutional processes	****
	Working with underrepresented groups				x				May take some time to show effect; may entail administrative burden	****
Application-design & parameters	Applicant behaviours				x				None known	*
	Expression of interest/pre-proposal		x			x			Longer time-to-grant, influx of out-of-scope EoIs, limits information to inform decision-making	***
	Reducing applications length/cutting sections	x				x			Limits information to inform decision-making, may not always save burden for applicants	***
Process design	'Sandpits'/Matching events						x		Problems for access, EDI issues; can be partially resolved through remote events	****
	2-stage application process		x			x			Slight danger of reduced levels of feedback	****
	Applicant anonymisation				x		x		Limited ability to judge feasibility of projects	****
	Automation-assisted reviewer allocation	x			x	x		x	Technology is not widely tested; some algorithms may have problems	***

Intervention	Save time	Increase relevance	Manage app volume	Reduce bias	Reduce burden	Fund high-risk	Increase review qual.	Hazards	Evidence strength rating
Dragon's den-style pitch		X				X		Favours applicants with sharp presenting skills; may present access-problems	*
External review only (no panel)	X				X	X		Reduced layers of risk control, potential lack of transparency	**
Group review							X	Group-bias	*
Changing the number of reviewers	X				X		X	Increase numbers: a single bad review can sink an application; labour intensive Decrease numbers: reduced robustness, potential for greater bias	***
Interviews							X	Resource-intensive, bias and/or disadvantage for certain groups	**
Moderation of reviews							X	Time-consuming for administrators; administrators may not have sufficient thematic expertise	*
Moderation panel							X	Not known	*
Panel only (no postal/external review)	X	X					X	Difficulty to cover the required expertise in a panel, may still need additional reviews, potential bias	***
Peer allocation	X				X			Possibly open to abuse/gaming, adds to applicant burden	***
Programme manager's discretion	X	X				X		Evidence that it may be under-used as programme managers themselves can be risk averse; lacks transparency, potentially a 'winners' game'	***
Standing panels vs. portfolio panels							X	Standing panels may potentially lead to institutionalised bias	***
Use of international assessors				X			X	May require more guidance/training for panellists	**
Use of metrics							X	Highly controversial: poor measure of excellence, open to bias and abuse, may contravene the DORA	***
Use of non-academic assessors (i.e. industry, policy & practice, patients, 'user' representatives)		X						May dilute notions of basic research, not recommended for such contexts	****
Virtual panels	X			X	X			Potentially less robust or detailed discussion, though this is unclear	***
Wildcard	X					X		Open to abuse if conflicts of interest are not monitored very well. Requires anonymised reviewing	***

	Intervention	Save time	Increase relevance	Manage app volume	Reduce bias	Reduce burden	Fund high-risk	Increase review qual.	Hazards	Evidence strength rating
	Partial randomisation				X	X	X		Reputational impact on applicants	***
	Scoring mechanisms		X					X	None confirmed but may disadvantage high-risk/ high-reward applications	****
	Sequential application of criteria (rather than simultaneous application of criteria)		X					X	None known	**
	Use of quotas				X				Very drastic approach	*
Training and feedback	**Bringing in reviewers from earlier careers & providing mentoring**				X			X	None known	**
	Embedding EDI in assessment				X				Ineffective training may install a false sense of confidence	**
	Expanding or reducing the amount/detail of feedback to unsuccessful applicants							X	Added burden; feedback may be of inconsistent quality	**
	Funder representation on review panels		X					X	None known	**
	Improving quality of reviews		X		X			X	None known	****
	Open review/rebuttal				X			X	Possibly increased burden for funder (and longer timelines depending how rebuttal works)	***

9.1   Recommendations

There are several recommendations stemming from our research. An initial draft of these was discussed and slightly refined at a validation meeting with UKRI in April 2023.

We note that the recommendations below are not specific to UKRI. They are intended as recommendations of good practice for any organisation involved in R&I funding.

Recommendations on how to use the interventions

Our headline recommendation is that **process design should always be a constituent part of scheme design**. The standard review process posited at the start of this report (submission, eligibility check, 2-3 external reviews, panel review, decision) should never be a 'default'. Every funding scheme has specific aims and characteristics, and so the design of the application, review and decision-making process should be considered for each individual funding opportunity.

We encourage funders to **make extensive use of the interventions studied here and to vary their assessment processes widely**. Our review shows that some highly effective interventions (e.g., two-stage processes, encouraging positive behaviours, interactive assessment processes) in achieving desired outcomes still require additional staff effort, which can be challenging in resource constraints. However, plenty of interventions also present opportunities for resource savings (e.g., using automation-assisted peer allocation, virtual panels, and partial randomisation). Therefore, funders can strategically review the mix of their funding portfolio and use interventions appropriate for the objectives of specific funding schemes and seek balanced use of interventions in terms of the resources required. For example, resources saved by introducing partial randomisation or panel-only approaches for smaller grants can be used to run two-stage processes and recruit non-academic reviewers in programmes that fund projects with extra-scientific objectives.

It is worth noting that such diversification may create a high cognitive load for both funder staff and researchers. In order to facilitate such diversification, it is therefore important that funders have the necessary resources and modernised systems needed to implement interventions as easily as possible. This likely constitutes an important confluence-point between this study and other recent work in the UK and beyond on research bureaucracy and research culture: there are many reasons to reduce bureaucracy and change research culture, and doing so will likely also create conditions where interventions to peer review processes can be implemented more easily.

Most critically, to ensure our recommended level of variation is possible, **IT systems need to have the necessary flexibility and function**: funders' application and review management systems (i.e. the IT underpinning the process) need to be designed in such a way that the interventions can easily be integrated into every bespoke scheme setup. While this is not a prerequisite for all 38 interventions studied here, it plays a part in many of them. Outdated, overly rigid IT systems may risk stifling funders' ability to vary and optimise their processes.

Critically, we note that the **judgement of experienced R&I funder staff is critical**. Almost every intervention we have considered has advantages as well as potential hazards and drawbacks. Our research can give extensive guidance on which interventions might suit a particular funding scheme, but scheme design is not a mechanical process with 'only one right answer'.

Most interventions studied here are suitable for specific contexts and should not be rolled out across all R&I funding opportunities. Indeed, a small number have extremely limited applicability (use of quotas, metrics, dragon's den pitches). However, **some interventions have**

the potential to become a 'new normal' in order to save burden and reduce bias across the board:

- **Providing additional support to groups unrepresented in the funder's portfolio to encourage them to apply and support them may be used by funders to improve diversity.** Out of interventions aiming to support greater inclusion, working with underrepresented groups is one with the highest demonstrated evidence strength. The actual implementation may vary from more sophisticated actions, including hands-on support, to less involved actions, like simply stating in the call document that the unrepresented groups are encouraged to apply. Both approaches are shown to be effective

- **Use of peer review colleges (and the training/briefing opportunities they entail) may be a good default practice** to improve review quality. Where the expertise represented on such colleges does not cover certain applications, there must however remain the possibility to recruit reviewers beyond the college. Funders should ensure the college membership is diverse (e.g., open to ECRs) and open to new participants

- On a related note, **automated reviewer allocation may become a genuine opportunity** for saving administrative burden, avoid conflicts of interest and increase reviewer response rates. Experience-sharing among funders will be important here, especially in relation to which systems have been proven to work. Peer review colleges combined with automation-assisted reviewer allocation would bring additional benefits

- There is a good case to **substantially expand use of anonymised reviewing**. Most funding schemes likely need at some stage to scrutinise the track record of applicants, but in multi-stage assessment processes and for smaller awards (where risk-levels are lower), having at least parts of the process anonymised would help reduce bias and inequitable outcomes

- While often seen as a 'radical' innovation in R&I funding, **there is a good case to mainstream an element of partial randomisation** across most R&I funding endeavours. This should not be extensive and should not cover all or even the majority of funding decisions: expert judgement through peer and panel review does well at identifying the very best applications, as well as the 'tail' of unsuitable ones. However, having partial randomisation as a consistently available option would enable some time-savings and counteract bias, both against underrepresented groups, but also against high-risk/high-reward ideas. As a minimum, randomisation should be used in cases where applications are of indistinguishable quality so as to avoid excessive and laboured discussion. Funders may however go further and randomise among a larger subset of high-quality applications where panels struggle to reach agreement

Recommendations on testing and further research

For some of the interventions covered in this report, there is limited evidence of their effectiveness simply because they have not been empirically studied to a sufficient degree. Virtual panels are potentially the most telling example. Many research funders have widely adopted virtual panels since the Covid-19 pandemic and report this has become a 'new normal' because of the time savings and associated improved ability to secure panel membership and diversity. While these gains are obvious and valuable, evidence of the impact on the discussion quality is scarce and requires further research. However, this should not necessarily discourage R&I funders from considering the interventions. For both well-tested and more embryonic interventions, we recommend that **funders monitor any interventions they undertake**, and where possible compare them to a pre-intervention baseline or to other funding schemes running in parallel. Importantly, funders should share good practice with their peers so that successes can be mainstreamed.

To counter the perceived risk that might accompany innovative use of the interventions, **we recommend that funders first test the intervention on a smaller scale via a pilot call and/or by commissioning accompanying process evaluations.** If funders introduce the intervention to an existing programme, then evaluation or simply review of monitoring data comparing the processes and outcomes pre- and post-intervention can be organised. The comparison allows for detecting the benefits (or lack of), improving the process and making a case for the decision-makers. Most evaluations of the interventions rely on programme monitoring data analysis, programme staff and stakeholder (applicants and reviewers) consultation and complete the evaluations during or right after the funding calls that introduce new interventions.

Our review shows that some interventions (demand management, shortening applications) can reduce the burden for the funder but not the system because the burden is simply shifted elsewhere, e.g. to the research community, to institutions, to other funders. Therefore, R&I funders should **follow up and assess the effects of the interventions on these wider constituencies.**

Recommendations beyond the interventions

Our review reveals that the assessment process can be improved with various interventions. However, procedural changes alone cannot fix wider systemic problems that may exist in research culture. Often interventions can go some way to enable improved outcomes, but wider problems of research culture may persist and even dampen the capacity for the interventions to achieve their greatest possible effect. We note this in particular because there have been great efforts by many funders and experts in recent years to assess and improve many elements of research culture, and our findings here should not be read as alternative 'quick fixes' to those important endeavours. **Investigations into wider research culture categorically need to continue alongside the process-interventions discussed in this report.**

Appendix A   Literature

A.1. Search strategy

Evidence of the use and outcomes of each of the 38 interventions was identified and reviewed. Documents include academic literature, 'grey' literature, as well as primary sources, such as websites and grant manuals produced by funding organisations.

Search

The documents were identified through 3 main routes, combining curation of known sources with an external search for additional documents:

1. UKRI shared material pertaining to the organisation's previous work on peer review with the project team. This included internal documents, PowerPoint presentations, published UKRI reports, and other relevant journal publications. UKRI also shared several documents listing external sources. 53 documents were shared with a total of 61 unique source references.

2. Another set of documents were identified through recommendations from external project advisers, experts within Technopolis, and interviewees. 41 unique documents were identified in this way, in addition to recommendations concerning specific programmes or funding organisations for which documentation was sought.

3. The project team then carried out an additional searches for each of the 38 interventions. Each search was carried out in Google Scholar (which captures academic literature but also has substantial coverage of grey literature, as well as working papers and other reports that may not be contained in other research information systems). The default search terms were "Grant peer review" combined with one or several terms related to the specific intervention using the Boolean operator "AND" and results were filtered to include only results from the last 10 years. For each search result, the top 25 results were scanned for relevance and the most relevant documents reviewed in full.

The coverage of these documents was somewhat uneven across the 38 interventions, with a larger number of sources on topics such as randomisation and EDI, and only few (if any) on other interventions.

Finally, our full resulting literature list was reviewed by one external and one Technopolis-internal expert, to potentially identify any significant gaps, i.e. omissions of any important sources known to either of the experts.

Analysis

The search produced a consolidated list of 176 references.[16] Each was reviewed and coded according to the one or more interventions they cover, key findings and type of evidence. On this basis, the team appraised the findings and strength of evidence available for each of the 38 interventions.

---

[16] Not all reviewed documents were included in this list: Some recommended sources were not relevant to the scope of the study and some UKRI documents were confidential and reviewed for background only.

Figure 2   Summary overview of literature review process
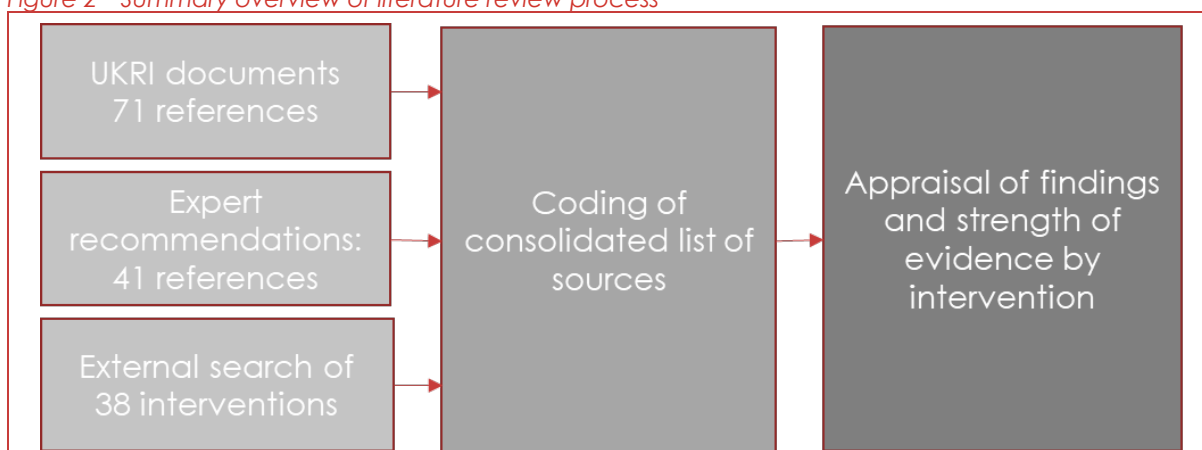
A.2. Bibliography

- 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). https://doi.org/10.1371/journal.pone.0046054

- American Academy of Arts & Sciences. (2008). *Advancing Research In Science and Engineering: Investing in Early-Career Scientists and High-Risk, High-Reward Research*. 69.

- Arnold, E., Lu, Y., & Xue, L. (2011). International evaluation of funding and management of the National Natural Science Foundation of China. Technopolis Group. http://www.technopolis-group.com/resources/downloads/reports/nsfc_evaluation_report.pdf

- Arnott, J. C. (2021). Pens and purse strings: Exploring the opportunities and limits to funding actionable sustainability science. *Research Policy*, *50*(10), 104362. https://doi.org/10.1016/j.respol.2021.104362

- Avin, S. (2019). Mavericks and lotteries. *Studies in History and Philosophy of Science Part A*, *76*(January 2018), 13–23. https://doi.org/10.1016/j.shpsa.2018.11.006

- Banal-Estañol, A., Liu, Q., Macho-Stadler, I., & Pérez-Castrillo, D. (2021). Similar-to-me effects in the grant application process: Applicants, panelists, and the likelihood of obtaining funds. *Economics Working Paper Series*, *1801*.

- Barnett, A. G., Herbert, D. L., Campbell, M., Daly, N., Roberts, J. A., Mudge, A., & Graves, N. (2015). Streamlined research funding using short proposals and accelerated peer review: An observational study. *BMC Health Services Research*, *15*(1), 2–7. https://doi.org/10.1186/s12913-015-0721-7

- Barr, J., Bryan, B., Kolarz, P., Potau, X., Punton, M., Simmonds, P., & Vogel, I. (2018). *GCRF Evaluation - Foundation Stage: Final Report*. Itad and Technopolis Group. https://www.gov.uk/government/publications/global-challenges-research-fund-gcrf-foundation-stage-evaluation

- Baždarić, K., Vrkić, I., Arh, E., Mavrinac, M., Marković, M. G., Bilić-Zulle, L., Stojanovski, J., & Malički, M. (2021). Attitudes and practices of open data, preprinting, and peer-review-A cross sectional study on Croatian scientists. *PLoS ONE*, *16*(6 June), 1–18. https://doi.org/10.1371/journal.pone.0244529

- Bendiscioli, B. S., Firpo, T., Bravo-biosca, A., Czibor, E., Garfinkel, M., Stafford, T., & Wilsdon, J. (2022). *The experimental Research funder's handbook*. Research on Research Institute (RoRI).

- Bendiscioli, S. (2019). The troubles with peer review for allocating research funding. *EMBO Reports*, *20*(12), 1–5. https://doi.org/10.15252/embr.201949472

- Bendiscioli, S., & Garfinkel, M. (2021). *Dealing with the limits of peer review with innovative approaches to allocating research funding*. EMBO Science Policy Programme.

- Bhattacharya, A. (2012). Science funding: Duel to the death. *Social Forces*, *488*, 20–52. https://doi.org/10.1093/sf/4.1.189

- Bieri, M., Roser, K., Heyard, R., & Egger, M. (2021). Face-to-face panel meetings versus remote evaluation of fellowship applications: Simulation study at the Swiss National Science Foundation. *BMJ Open*, *11*(5), 1–8. https://doi.org/10.1136/bmjopen-2020-047386

- Blackburn, M., Coutinho, K., & Suviste, H. (2022). *The Equality, Diversity and Inclusion Engagement Fellowship Pilot AHRC Funding Scheme Report*. The Open University.

- Bohannon, J. (2011) Meeting for Peer Review at a Resort That's Virtually Free. Science. 7 Jan 2011, Vol 331, Issue 6013, p. 27. https://www.science.org/doi/10.1126/science.331.6013.27

- Bol, T., de Vaan, M., & van de Rijt, A. (2022). Gender-equal funding rates conceal unequal evaluations. *Research Policy*, *51*(1), 104399. https://doi.org/10.1016/j.respol.2021.104399

- Bonetta, L. (2008). Enhancing NIH Grant Peer Review: A Broader Perspective. *Cell,* 135(2), 201-204. https://doi.org/10.1016/j.cell.2008.09.051

- Boudreau, K., Guinan, E., Lakhani, K., & Riedl, C. (2016). Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science. *Manage Science*, 62(10). https://pubsonline.informs.org/doi/10.1287/mnsc.2015.2285

- Cechlárová, K., Fleiner, T., & Potpinková, E. (2014). Assigning evaluators to research grant applications: the case of Slovak Research and Development Agency. *Scientometrics*, 99, 495–506.

- CGIAR IAES. (2023). Standing Panel on Impact Assessment (SPIA). CGIAR Independent Advisory and Evaluation Services (IAES). https://iaes.cgiar.org/spia (Accessed: 24 March 2023)

- Clarke, P., Herbert, D., Graves, N., & Barnett, A. G. (2016). A randomized trial of fellowships for early career researchers finds a high reliability in funding decisions. *Journal of Clinical Epidemiology*, 69, 147–151. https://doi.org/10.1016/j.jclinepi.2015.04.010

- Centre for Scientific Review (2023). CSR Overview. URL: https://public.csr.nih.gov/AboutCSR/Evaluations . (Accessed: 23 March 2023)

- CGIAR Independent Advisory and Evaluation Services (2019). Standing Panel on Impact Assessment. Available: https://iaes.cgiar.org/spia (Accessed: 24 March 2023)

- Comfort, N. (2021). Addressing Racial Disparities in NIH Funding. *Journal of Science Policy & Governance*, *18*(04). https://doi.org/10.38126/jspg180408

- Coyte, R., Rooney, J., & Phua, B. (2013). The Impossibility of Philanthropic Funding Decisions: The Australian Non-Government Funder Experience. *Financial Accountability & Management*, 29(4), 397–418. https://doi.org/10.1111/faam.12020

- Cruz-Castro, L., & Sanz-Menéndez, L. (2019). *Grant Allocation Disparities from a Gender Perspective: Literature Review. Synthesis Report*. *1*(November), 84. http://hdl.handle.net/10261/200024

- Cunningham, P. (2012). 06 The Impact and Effectiveness of Policies to Support Collaboration for R & D and Innovation. *NESTA Compendium of Evidence on Innovation Policy Intervention*, *March*, 1–53. http://www.innovation-policy.net/compendium/

- Curry, S., Gadd, E., & Wilsdon, J. (2022). *Harnessing the metric tide: indicators, infrastructures and priorities for responsible research assessment in the UK* (Issue December). https://rori.figshare.com/articles/report/Harnessing_the_Metric_Tide/21701624

- *D 5.1 Report on "Strategic advice for enhancing the gender dimension of Open Science and Innovation Policy"*. (2018). GENDer equality in the ERA Community To Innovate policy implementatiON (GENDERACTION). https://genderaction.eu/wp-content/uploads/2018/12/741466_GENDERACTION_D11_Report-on-Strategic-advice-for-enhancing-the-gender-dimension-of-OS-and-Innovation-Policy_approved-1.pdf

- Mazhindu D, Gregory S. Dragons' Den: promoting healthcare research and innovation. Nurs Manag (Harrow). 2015 Jul;22(4):28-33. doi: 10.7748/nm.22.4.28.e1375. PMID: 26135194

- *Demand management*. (2022). NERC. https://www.ukri.org/councils/nerc/guidance-for-applicants/types-of-funding-we-offer/discovery-science/demand-management/

- Department of Education (2013). National Institute on Disability and Rehabilitation Research: Long-Range Plan for Fiscal Years 2013-2017. Federal Register. Available: https://www.federalregister.gov/documents/2013/04/04/2013-07879/national-institute-on-disability-and-rehabilitation-research-long-range-plan-for-fiscal-years

- Derrick, G. E., Zimmermann, A., Greaves, H., Best, J., & Klavans, R. (2022). *Targeted, actionable and fair: reviewer reports as feedback and its effect on ECR career choices*. https://osf.io/preprints/socarxiv/a8psh/

- Dias Lopes, A., & Wakeling, P. (2022). *Inequality in Early Career Research in the UK Life Sciences*. University of York. https://www.ukri.org/wp-content/uploads/2022/11/BBSRC-301122-BBSRCInequalityInECRReport.pdf

- Dumenco, L., Engle, D., Goodell, K., Nagler, A., Ovitsh, R., & Whicker, S. (2017). Expanding Group Peer Review: A Proposal for Medical Education Scholarship. Academic Medicine, 92(2), 147–149.

- EPSRC. (2023). *Ethnicity and Race Inequity In Our Portfolio: Findings of our community engagement and actions for change*. https://www.ukri.org/publications/ethnicity-and-race-inequity-in-our-portfolio/ethnicity-and-race-inequity-in-our-portfolio-findings-of-our-community-engagement-and-actions-for-change/

- EPSRC (2022). Ethnicity and Race Inequity In Our Portfolio: Findings of our community engagement and actions for change. https://www.ukri.org/wp-content/uploads/2022/11/21112022-EPSRC-Ethnicity-and-Race-Inequity-In-Our-Portfolio_Nov-2022.pdf

- EPSRC (2020). *Understanding our portfolio: A gender perspective*. https://www.ukri.org/publications/epsrc-understanding-our-portfolio-a-gender-perspective/

- EPSRC (2022). *Gender Diversity in our Portfolio: Survey Findings and Interventions*. UKRI. https://www.ukri.org/publications/gender-diversity-in-our-portfolio-survey-findings/

- *ESRC – Large Grants Competition 2016/17 (Update)*. (n.d.). University of Lincoln: Research Blog. Retrieved February 20, 2023, from https://research.blogs.lincoln.ac.uk/2016/08/03/esrc-large-grants-competition-201617-update/

- ESRC. (n.d.). *Might randomised funding increase ESRC investigator diversity? An analysis of the 'ESRC Research Grants' call*.

- ESRC. (n.d.). *ESRC applicant and reviewer characteristics*.

- *ESPR Research Grant Programme 2017-2022*. (2022). European Society for Paediatric Research. https://www.espr.eu/funding/research_grant_programme.php

- ESRC (2022). *ESRC review request responses: rates and reasons* (Issue October). Presentation.

- ESRC (n.d.). *How long should we wait for reviewers to respond?* Presentation.

- European Science Foundation. (2011). *ESF Survey Analysis Report on Peer Review Practices*. 60. http://www.esf.org/fileadmin/Public_documents/Publications/pr_guide_survey.pdf

- Fang, F. C., & Casadevall, A. (2016). Research funding: The case for a modified lottery. *MBio*, *7*(2). https://doi.org/10.1128/mBio.00422-16

- Fang, F. C., Bowen, A., & Casadevall, A. (2016). NIH peer review percentile scores are poorly predictive of grant productivity. *ELife*, *5*, 1–6. https://doi.org/10.7554/eLife.13323

- Feliciani, T., Morreau, M., Luo, J., Lucas, P., & Shankar, K. (2022). Designing grant-review panels for better funding decisions: Lessons from an empirically calibrated simulation model. *Research Policy*, *51*(4), 104467. https://doi.org/10.1016/j.respol.2021.104467

- Fitzgerald, C., Martin, A., Berner, D. & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review. BMC Psychology. 7:29. https://bmcpsychology.biomedcentral.com/articles/10.1186/s40359-019-0299-7

- Foundation, E. S. (2011). European Peer Review Guide Integrating Policies and Practices. In *System*.

- Freel, S. A., Smith, P. C., Burns, E. N., Downer, J. B., Brown], A. J., & Dewhirst, M. W. (2017). Multidisciplinary Mentoring Programs to Enhance Junior Faculty Research Grant Success. *Acad Med.*, 92(10), 1410–1415. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5589470/

- Gallo, S. A., Carpenter, A. S., & Glisson, S. R. (2013). Teleconference versus Face-to-Face Scientific Peer Review of Grant Application: Effects on Review Outcomes. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071693

- Gallo, S. A., Sullivan, J. H., & Glisson, S. R. (2016). The Influence of Peer Reviewer Expertise on the Evaluation of Research Funding Applications. *PLoS ONE*, *11*(10), e0165147. https://doi.org/10.1371/journal.pone.0165147

- Gallo, S., et al. (2018). Participation and Motivations of Grant Peer Reviewers: A Comprehensive Survey of the Biomedical Research Community. bioRxiv., Available: https://www.biorxiv.org/content/10.1101/479816v2

- Gallo, S. A., Sullivan, J. H., & Croslan, D. J. R. (2022). Scientists from Minority-Serving Institutions and Their Participation in Grant Peer Review. *BioScience*, *72*(3), 289–299. https://doi.org/10.1093/biosci/biab130

- Gamage, D. D. K., & Sevilla, A. (2019). Gender Equality and Positive Action: Evidence from UK Universities. *AEA Papers and Proceedings*, *109*, 105–109. https://doi.org/10.2139/ssrn.3390198

- GAO. (2016). *Federal Research Grants: Opportunities Remain for Agencies to Streamline Administrative Requirements* (Issue June). United States Government Accountability Office. https://www.gao.gov/products/gao-16-573

- García-ruiz, J. M. (2019). I won a project! *Substantia*, *3*(1), 5–7. https://doi.org/10.13128/Substantia-182

- Ginther, D. K., & Heggeness, M. L. (2020). Administrative Discretion in Scientific Funding: Evidence from a Prestigious Postdoctoral Training Program. *Research Policy*, *49*(4).

- Gluckman, P., Ferguson, M., Glover, A., Grant, J., Groves, T., Lauer, M., & Ulfendahl, M. (2017). *International Peer Review Expert Panel report: A report to the Governing Council of the Canadian Institutes of Health Research - CIHR*. Canadian Institutes of Health Research. https://cihr-irsc.gc.ca/e/50248.html

- Graham, C. L. B., Landrain, T. E., Vjestica, A., Masselot, C., Lawton, E., Blondel, L., Haenal, L., Greshake Tzovaras, B., & Santolini, M. (2022). Community review: a robust and scalable selection system for resource allocation within

open science and innovation communities. *F1000Research*, *11*, 1440. https://doi.org/10.12688/f1000research.125886.1

- Grant, J., & Allen, L. (1999). Evaluating high risk research: An assessment of the Wellcome Trust's Sir Henry Wellcome Commemorative Awards for Innovative Research. *Research Evaluation*, *8*(3), 201–204. https://doi.org/10.3152/147154499781777513

- Grant preparation time: the law of diminishing returns prompts a rethink on the application and review process. (2013). *BMJ Open*. https://acquire.cqu.edu.au/articles/journal_contribution/Grant_preparation_time_the_law_of_diminishing_returns_prompts_a_rethink_on_the_application_and_review_process/13461869

- Green, T. & Hagiwara, N. (2020). The Problem with Implicit Bias Training. Scientific American. https://www.scientificamerican.com/article/the-problem-with-implicit-bias-training/

- Gregory-Smith, I. (2017). Positive action towards gender equality? Evidence from the Athena SWAN Charter in UK Medical Schools. *British Journal of Industrial Relations*, *56*(3), 463–483

- Grove, L. (2017). *The effects of funding policies on academic research*. Doctoral thesis, University College London. http://eprints.lse.ac.uk/88207/

- Gunashekar, S., Wooding, S., & Guthrie, S. (2017). How do NIHR peer review panels use bibliometric information to support their decisions? *Scientometrics*, *112*(3), 1813–1835. https://doi.org/10.1007/s11192-017-2417-8

- Guthrie, S. (2019). Innovating in the Research Funding Process: Peer Review Alternatives and Adaptations - Horizon Scan. Paradigm Project, AcademyHealth. https://www.academyhealth.org/sites/default/files/innovatingresearchfundingnovember2019.pdf

- Guthrie, S., Ghiga, I., & Wooding, S. (2017). What do we know about grant peer review in the health sciences? *F1000Research*, *6*(May). https://doi.org/10.12688/f1000research.11917.2

- Guthrie, S., Guérin, B., Wu, H., Ismail, S., & Wooding, S. (2013). Alternatives to peer review in research project funding. *RAND Europe Report*, 1–44. http://www.rand.org/content/dam/rand/pubs/research_reports/RR100/RR139/RAND_RR139.pdf

- Guthrie, S., Rodriguez-Rincon, D., McInroy, G., Ioppolo, B., & Gunashekar, S. (2019). *Design and development of an evaluation framework for the NHMRC: Review of international practice, overview of NHMRC's data and grant management systems, and an evaluation framework and metrics.* https://doi.org/10.7249/rr2984

- Heinze, T. (2008). How to sponsor ground-breaking research: A comparison of funding schemes. *Science and Public Policy*, *35*(5), 302–318. https://doi.org/10.3152/030234208X317151

- Herbert, D. L., Graves, N., Clarke, P., & Barnett, A. G. (2015). Using simplified peer review processes to fund research: A prospective study. *BMJ Open*, *5*(7), 1–7. https://doi.org/10.1136/bmjopen-2015-008380

- Hesselberg, J.-O., Fostervold, K. I., Ulleberg, P., & Svege, I. (2021). Individual versus general structured feedback to improve agreement in grant peer review: a randomized controlled trial. *Research Integrity and Peer Review*, *6*(1), 1–11. https://doi.org/10.1186/s41073-021-00115-5

- Hillier, E., Razack, S., Clary, V., & Münter, L. (2018). *The State of Women in Life Sciences at McGill: A Summary and Report of the Win4Science Forum.* https://www.mcgill.ca/pharma/files/pharma/final_report.win4science_report_-_nov_30_2018.pdf

- Höylä, T., Bartneck, C., & Tiihonen, T. (2016). The consequences of competition: simulating the effects of research grant allocation strategies. *Scientometrics*, *108*(263–288).

- Hug, S. E., & Aeschbach, M. (2020). Criteria for assessing grant applications: a systematic review. *Palgrave Communications*, *6*(1), 1–15. https://doi.org/10.1057/s41599-020-0412-9

- Hug, S. E., & Ochsner, M. (2022). Do peers share the same criteria for assessing grant applications? *Research Evaluation*, *31*(1), 104–117. https://doi.org/10.1093/reseval/rvab034

- Huntoon, L. R. (2019). Sham Peer Review : the Destruction of Medical Careers. *Journal of American Physicians and Surgeons*, *24*(4), 99–100. https://jpands.org/vol24no4/huntoon.pdf

- Institute of Education (2019). IES guide for grant peer review panel members: information, tips and instructions for members of IES scientific grant peer review panels. Available: https://ies.ed.gov/director/sro/pdf/IESGuideforPeerReviewersFY2019.pdf

- Ismail, S., Farrands, A., & Wooding, S. (2009). *Evaluating Grant Peer Review in the Health Sciences: A review of the literature.* RAND Europe.

- Ito, Y. (2004). Japanese research grants for young researchers. *Science and Public Policy*, *31*(3), 175–184. https://doi.org/10.4319/lo.1959.4.4.0503

- Jerrim, J., & de Vries, R. (2020). Are peer-reviews of grant proposals reliable? An analysis of Economic and Social Research Council (ESRC) funding applications. *The Social Science Journal*.

- Kalmukov, Y. (2020). An algorithm for automatic assignment of reviewers to papers. *Scientometrics*, *124*(3), 1811–1850. https://dl.acm.org/doi/abs/10.1007/s11192-020-03519-0

- Kelso, A. (2021). *Peer review for Ideas Grants in 2021*. CEO Communique, February 2021, National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/news-centre/peer-review-ideas-grants-2021

- Kolarz, P., Davé, A., Bryan, B., Urquhart, I., Rigby, J. & Suninen, L. (2023) *Organisational and Procress Review of the Human Frontier Science Program*. Technopolis Group. Published by HFSP: https://www.hfsp.org/node/74873#book/

- Kolarz, P., Arnold, E., Bryan, B., D'hont, J., Horvath, A., Simmonds, P., Varnai, P. & Vingre, A. (2022) Process review of UKRI's research and innovation response to Covid-19. UKRI. https://www.ukri.org/wp-content/uploads/2022/01/UKRI-180122-ProcessReviewUKRIResponseCOVID19-FinalReport.pdf

- Kolarz, P., Arnold, E., Cimatti, R., Dobson, C. & Seth, V. (2020). Process review of the IRC Laureate Award. Irish Research Council. https://research.ie/assets/uploads/2021/02/IRC-Laureate-review-FINAL.pdf

- Kolarz, P., Arnold, E., Davé, A., Andreasson, H., & Bryan, B. (2018). How Research Funders Ensure the Scientific Legitimacy of their Decisions: Investigation in support of the design of Formas (Issue November). Technopolis Group.

- Kolarz, P., Arnold, E., Farla, K., Gardham, S., Nielsen, K., Rosemberg, C., & Wain, M. (2016). *Evaluation of the ESRC Transformative Research Scheme*. Technopolis Group, July 2016. https://esrc.ukri.org/files/research/research-and-impact-evaluation/esrc-transformative-research-scheme-evaluation/

- Kolarz, P., Bryan, B., Farla, K., Krčál, A., Potau, X., & Simmonds, P. (2018). *UKRI Research and Innovation Funding Service (RIFS) visioning work - Final report*. Technopolis Group, 18 May 2018. http://www.technopolis-group.com/wp-content/uploads/2018/07/UKRI-RIFS-Final-report.pdf

- Kolehmainen, C., & Carnes, M. (2018). Who Resembles a Scientific Leader – Jack or Jill? How Implicit Bias Could Influence Research Grant Funding. *Circulation.*, *137*(8), 769.

- Kolev, J., Murray, Y., & Fuentes-Medel, F. (2019). Is Blinded Review Enough? How Gendered Outcomes Arise Even Under Anonymous Evaluation. *NBER Working Paper Series*. http://www.nber.org/papers/w25759

- Lal, B., Wilson, A., Jonas, S., Lee, E., Richards, A., & Peña, V. (2012). *An outcome evaluation of the National Institutes of Health (NIH) Director's pioneer award (NDPA) program, FY 2004-2006*. 204. https://www.ida.org/~/media/Corporate/Files/Publications/STPIPubs/ida-p-4899.ashx

- Lane, J. N., Teplitskiy, M., Gray, G., Ranu, H., Menietti, M., Guinan, E. C., & Lakhani, K. R. (2022). Conservatism Gets Funded? A Field Experiment on the Role of Negative Information in Novel Project Evaluation. *Management Science*, *68*(6), 4478–4495. https://doi.org/10.1287/mnsc.2021.4107

- Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science*, *31*(6), 820–841. https://doi.org/10.1177/030631201031006002

- Langfeldt, L., Ingeborgrud, L., Reymert, I., Svartefoss, S. M., & Borlaug, S. B. (2022). *Evaluation of the Spark pilot*.

- Langfeldt, L., Reymert, I., & Aksnes, D. W. (2021). The role of metrics in peer assessments. *Research Evaluation*, *30*(1), 112–126. https://doi.org/10.1093/reseval/rvaa032

- Lauer, M. (2020). *Anonymizing Peer Review for the NIH Director's Transformative Research Award Applications*. NIH Extramural Nexus. https://nexus.od.nih.gov/all/2020/05/27/anonymizing-peer-review-for-the-nih-directors-transformative-research-award-applications/

- Lee, C. J. (2015). Commensuration bias in peer review. *Philosophy of Science*, *82*(5), 1272–1283. https://doi.org/10.1086/683652

- Lodge, H. (2020). *SANDPIT METHODOLOGY: Results of a rapid literature search to inform a sandpit exercise for PETRA*. https://petranetwork.org/wp-content/uploads/2020/05/Sandpit-methodology-overview-March-2020.pdf

- Luo, J., Ma, L., & Shankar, K. (2021). Does the inclusion of non-academic reviewers make any difference for grant impact panels? *Science and Public Policy*, *48*(6), 763–775. https://doi.org/10.1093/scipol/scab046

- Luukkonen, T., Stampfer, M., & Strassnig, M. (2015). *Evaluation practices in the selection of ground-breaking research proposals*. fte Workiing Paper. https://www.etla.fi/en/publications/evaluation-practices-in-the-selection-of-ground-breaking-research-proposals/

- Marco, B., & Università, C. (2022). *Horizon Europe and Plan B research funding : Turning adversity into opportunity*. *September*, 1–8.

- Maxwell, K., & Benneworth, P. (2018). The construction of new scientific norms for solving Grand Challenges. *Palgrave Communications*, *4*(1). https://doi.org/10.1057/s41599-018-0105-9

- Maxwell, K., Benneworth, P., & Siefkes, M. (2018). Sandpits can develop cross-disciplinary projects, but funders need to be as open-minded as researchers. *LSE Impact Blog*. https://doi.org/10.1057/s41599-018-0105-9

- Mazhindu, D., & Gregory, S. (2015). Dragons' Den: promoting healthcare research and innovation. *Nursing Management*, 22(4), 28–33.

- Meadmore, K., Fackrell, K., Recio-Saucedo, A., Bull, A., Fraser, S. D. S., & Blatch-Jones, A. (2020). Decision-making approaches used by UK and international health funding organisations for allocating research funds: A survey of current practice. *PLoS ONE*, 15(11), 1–17. https://doi.org/10.1371/journal.pone.0239757

- Members, P., Selectors, E., & Administrators, P. (2013). *Peer review: options appraisal*. REPSET : 02/12/13.

- Mervis., J (2014) A radical change in peer review: A pilot project to ease pressure on NSF's vaunted peer-review system required grant applicants to review seven competing proposals. Science, 18 Jul 2014, Vol 345, Issue 6194, pp. 248-249. https://www.science.org/doi/full/10.1126/science.345.6194.248

- Mervis, J. (2016). NSF tries two-step review, drawing praise - And darts. *Science*, 353(6299), 528–529. https://doi.org/10.1126/science.353.6299.528

- Moreno, M. del C. C., Warta, K., Arnold, E., Tiefenthaler, B., Kolarz, P., & Skok, S. (2020). *Science Europe Study on Research Assessment Practices - Final Report*. Technopolis Group. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjJ27vTka_9AhXySfEDHRleBb0QFnoECAgQAQ&url=https%3A%2F%2Fwww.scienceeurope.org%2Fmedia%2Ffmdihoqy%2Fse-study-on-research-assessment-practices-report.pdf&usg=AOvVaw0fVC_jSsWgeSQVbBU8DN0M

- Morgan, B., Yu, L. M., Solomon, T., & Ziebland, S. (2020). Assessing health research grant applications: A retrospective comparative review of a one-stage versus a two-stage application assessment process. *PLoS ONE*, 15(3), 1–18. https://doi.org/10.1371/journal.pone.0230118

- Morgan, B., Yu, L.-M., Solomon, T., & Ziebland, S. (2020). Assessing health research grant applications: A retrospective comparative review of a one-stage versus a two-stage application assessment process. *PLoS One*, 15(3).

- Morris, M., Alencar, Y., Rachet, B., Stephens, R., & Coleman, M. P. (2020). Fleshing out the data: when epidemiological researchers engage with patients and carers. Learning lessons from a patient involvement activity. *BMJ Open*, 10(9), e036311. https://doi.org/10.1136/bmjopen-2019-036311

- Mutz, R., Bornmann, L., & Daniel, H. D. (2016). Funding decision-making systems: An empirical comparison of continuous and dichotomous approaches based on psychometric theory. *Research Evaluation*, 25(4), 416–426. https://doi.org/10.1093/reseval/rvw002

- Nagler, A., Ovitsh, R., Dumenco, L., Whicker, S., Engle, D., & Goodell, K. (2019). Communities of Practice in Peer Review: Outlining a Group Review Process. Academic Medicine, 94(19), 1437–1442.

- Nakamura, R., Mann, L. S., Lindner, M. D., Braithwaite, J., Chen, M. C., Vancea, A., Byrnes, N., Durrant, V., & Reed, B. (2021). An experimental test of the effects of redacting grant applicant identifiers on peer review outcomes. *ELife*, 10, 1–15. https://doi.org/10.7554/eLife.71368

- National Institute of Justice (2023). Becoming a Peer Reviewer for NIJ. Available at: https://nij.ojp.gov/funding/becoming-peer-reviewer-nij (Accessed: 23 March 2023)

- *NERC funding investments between 2014 to 2015 and 2019 to 2020*. (2022). UKRI. https://www.ukri.org/publications/nerc-funding-investments-2014-2015-and-2019-2020/nerc-funding-investments-between-2014-2015-and-2019-2020/

- Newton, P. & Feucht, T. (2012). Improving NIJ's Peer Review Process: The Scientific Review Panel Pilot Project. National Institute of Justice Journal. Available at: https://nij.ojp.gov/topics/articles/improving-nijs-peer-review-process-scientific-review-panel-pilot-project (accessed: 23 March 2023)

- NIH. (2023). CSR Data & Evaluations. National Institutes of Health, Center for Scientific Review. https://public.csr.nih.gov/AboutCSR/Evaluations (Accessed: 23 March 2023)

- Nous Group. (2018). *Consultation on the development of peer review for NHMRC's new grant program - Final Report*. https://doi.org/10.1111/j.1834-7819.1982.tb04109.x

- Owen, D., Featherstone, H., & Kerry, L. (2016). *The State of Play: Public Engagement with Research in UK Universities*. *September*, 52.

- Ovseiko, P. V, Taylor, M., Gilligan, R. E., Birks, J., Elhussein, L., Rogers, M., Tesanovic, S., Hernandez, J., Wells, G., Greenhalgh, T., & Buchan, A. M. (2020). Effect of Athena SWAN funding incentives on women's research leadership. *BMJ*, 371, m3975. https://www.bmj.com/content/371/bmj.m3975

- Park, H., Lee, J., & Kim, B. C. (2015). Project selection in NIH: A natural experiment from ARRA. *Research Policy*, 44(6), 1145–1159. https://doi.org/10.1016/j.respol.2015.03.004

- Pier, E. L., Raclaw, J., Kaatz, A., Brauer, M., Carnes, M., Nathan, M. J., & Ford, C. E. (2017). "Your comments are meaner than your score": Score calibration talk influences intra- and inter-panel variability during scientific grant peer review. *Research Evaluation*, 26(1), 1–14. https://doi.org/10.1093/reseval/rvw025

- *Powering Discovery: The Expert Panel on International Practices for Funding Natural Sciences and Engineering Research*. (2021). The Council 0f Canadian Academies. https://cca-reports.ca/wp-content/uploads/2021/05/Powering-Discovery-Full-Report-EN_DIGITAL_FINAL.pdf

- Pritlove, C., Juando-Prats, C., Ala-Leppilampi, K. & Parsons, J. (2019), The good, the bad, and the ugly of implicit bias. The Lancet. 393. https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(18)32267-0.pdf

- Publons. (2018). *Global State Of Peer Review 2018*. Publons and Clarivate Analytics. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjy8o-Tmq_9AhWMtKQKHY9-ASAQFnoECAQQAQ&url=https%3A%2F%2Fpublons.com%2Fstatic%2FPublons-Global-State-Of-Peer-Review-2018.pdf&usg=AOvVaw3xjM8vzC-KkdNCHrGDT3iv

- Pupella, V., Monteverde, M. E., Lombardo, C., Belardelli, F., & Giacomini, M. (2014). A semi-automatic web based tool for the selection of research projects reviewers. *Studies in Health Technology and Informatics*, *205*, 950–954. https://doi.org/10.3233/978-1-61499-432-9-950

- Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: a realist synthesis. *Research Integrity and Peer Review*, *7*(1), 1–28. https://doi.org/10.1186/s41073-022-00120-2

- Research Councils UK. (2006). *Report of the Research Councils UK Efficiency and Effectiveness of Peer Review Project*. http://www.rcuk.ac.uk/documents/documents/rcukprreport.pdf

- Rivard, J., O'Connell, M. & Wegman, D. (2012). Review of Disability and Rehabilitation Research: NIDRR Grantmaking Processes and Products. National Academy of Sciences. Available: https://www.ncbi.nlm.nih.gov/books/NBK189591/pdf/Bookshelf_NBK189591.pdf

- Rodgers, A., & Minchin, W. (2020). *Draft Expert Review Analysis - Full Report [Internal Use Only* (Issue August). UKRI.

- Rodriguez-Rincon, D., Feijao, C., Stevenson, C., Evans, H., Sinclair, A., Thomson, S., & Guthrie, S. (2021). *Study on the proposal evaluation system for the EU R&I framework programme*. European Commision. https://doi.org/10.2777/16211

- Roper, R. L. (2019). Does Gender Bias Still Affect Women in Science? *Microbiology and Molecular Biology Reviews*, *83*(3). https://doi.org/10.1128/mmbr.00018-19

- Roumbanis, L. (2019). Peer Review or Lottery? A Critical Analysis of Two Different Forms of Decision-making Mechanisms for Allocation of Research Grants. *Science Technology and Human Values*, *44*(6), 994–1019. https://doi.org/10.1177/0162243918822744

- Royal Society. (2007). *Response to the RCUK consultation on the Efficiency and Effectiveness of Peer Review*. RS Policy Document 04/07. https://royalsociety.org/~/media/royal_society_content/policy/publications/2007/8090.pdf

- Royal Society. (2021). *Submission to the independent review of research bureaucracy*. 15 October 2021. https://royalsociety.org/-/media/policy/Publications/2021/21-10-15-royal-society-response-to-the-independent-review-of-research-bureaucracy-consultation.pdf

- Rubel, O. E., Boulevard, F., & Boulevard, F. (2018). *Institute Of Market Problems And Economic & Ecological Research of the National Academy Of Sciences Of Ukraine Institute Of Market Problems And Economic & Ecological Research of the National Academy Of Sciences Of Ukraine THE CONCEPT OF THE IMPLEMENTATION OF THE PERFORMANCE-BASED RESEARCH Інститут проблем ринку та економіко - екологічних досліджень НАНУ aspects were covered also by Y . Garfield , S . Bradford , A . Pritchard , L . Resing , D . Hawkins . Quantitative reports are based on the basis of national case studies and academic academic literature . [3]. Qualitative analysis. 4*(69), 0–1.

- Samuel, G. N., & Derrick, G. (2021). Exploring the degree of delegated authority for the peer review of societal impact. *Science And Public Policy*.

- Sapple, P. (2019). *A Report on Public Engagement with NERC Environmental Science: Challenges & Opportunities*. *August*, 1–37.

- Sato, S., Gygax, P.M., Randall, J. *et al.* (2021). The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. *High Educ* 82, 145–162. https://doi.org/10.1007/s10734-020-00626-y

- Sattler, D. N., McKnight, P. E., Naney, L., & Mathis, R. (2015). Grant peer review: Improving inter-rater reliability with training. *PLoS ONE*, *10*(6), 1–9. https://doi.org/10.1371/journal.pone.0130450

- Schnädelbach, H., Sun, X., Kefalidou, G., Coughlan, T., Meese, R., Norris, J., & McAuley, D. (2016). Creativity Greenhouse: At-a-distance collaboration and competition over research funding. *International Journal of Human Computer Studies*, *87*, 1–19. https://doi.org/10.1016/j.ijhcs.2015.10.006

- Schouten, R. et al. (2023), Evaluation of research proposals by peer review panels: broader panels for broader assessments? Science and Public Policy, 00, 1-14.

- Science, O., & Papers, P. (2018). *Effective operation of competitive research funding systems*. *57*. https://www.oecd-ilibrary.org/industry-and-services/effective-operation-of-competitive-research-funding-systems_2ae8c0dc-en

- Sen, A. (2014). Totally radical: From transformative research to transformative innovation. *Science and Public Policy*, *41*(3), 344–358. https://doi.org/10.1093/scipol/sct065

- Severin, A., Martins, J., Delavy, F., Jorstad, A., & Egger, M. (2019). Potential Bias in Peer Review of Grant Applications at the Swiss National Science Foundation. *PeerJ Preprints*, 0–2. https://peerj.com/preprints/27587/%0Ahttps://www.nature.com/articles/d41586-019-01198-3

- Shepherd, J., Frampton, G. K., Pickett, K., & Wyatt, J. C. (2018). Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. *PLoS ONE*, *13*(5), 1–26. https://doi.org/10.1371/journal.pone.0196914

- Sinkjær, T. (2012). Fund ideas, not pedigree, to find fresh insight. *Nature*, https://doi.org/10.4135/9781412971959.n413

- Solans-Domènech, M., Guillamón, I., Ribera, A., Ferreira-González, I., Carrion, C., Permanyer-Miralda, G., & Pons, J. M. V. (2017). Blinding applicants in a first-stage peer-review process of biomedical research grants: An observational study. *Research Evaluation*, *26*(3), 181–189. https://doi.org/10.1093/reseval/rvx021

- STFC. (n.d.). *Public Engagement: Attitude, Culture and Ethos*

- Strang, V., & McLeish, T. (2015). Evaluating Interdisciplinary Research: a practical guide. *Report*, 1–20. https://www.dur.ac.uk/ias/news/?itemno=25309

- Swedish Research Council. (n.d.). *Several grants simultaneously*. Retrieved March 1, 2023, from https://www.vr.se/english/applying-for-funding/applying-for-a-grant/several-grants-simultaneously.html

- Swedish Research Council. (n.d.). *What requirements apply if I already have a grant from the Swedish Research Council*. Retrieved March 1, 2023, from What requirements apply if I already have a grant from the Swedish Research Council?

- Tamblyn, R., Girard, N., Qian, C. J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. *Cmaj*, *190*(16), E489–E499. https://doi.org/10.1503/cmaj.170901

- The Behavioural Insights Team (2020). Unconscious bias and diversity training – what the evidence says. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/944431/20-12-14_UBT_BIT_report.pdf

- Tickell, A. (2022) Independent Review of Research Bureaucracy: Final Report. Published by the department of Business, Energy and Industrial Strategy (BEIS), UK: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1094648/independent-review-research-bureaucracy-final-report.pdf

- Tinkle, F. S. S., Han, X. S., Thrift, E. M., Newell, L. P., Hazan, G. G., & Dinh, N. N. (2022). *Assessment of an NIH Anonymized Review Process*.

- Tough love. (2010). *Nature*, *464*(7288), 465. https://doi.org/10.1038/464465a

- Treem, J. W., Schneider, M., Zender, R. L., & Sorkin, D. H. (2018). Exploring the potential role of community engagement in evaluating clinical and translational science grant proposals. *Journal of Clinical and Translational Science*, *2*(3), 139–146. https://doi.org/10.1017/cts.2018.311

- UKRI. (2020). *Diversity results for UKRI funding data: 2014-15 to 2018-19*. UK Research and Innovation. https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-DiversityResultsForUKRIFundingData2014-19.pdf

- UKRI. (n.d.). *Comparing assessment methods : outline-to-full versus full proposal only processes.*

- UKRI. (2021). *How does the number of grant reviews affect the reliability of peer review decisions?* (Issue December).

- UKRI (2022). Responsive mode grant assessment process. Available: https://www.ukri.org/about-us/esrc/grant-assessment-panels/ (Accessed 24 March 2023)

- University of Florida College of Education (2015). An Overview: IES Procedures for Peers Review of Grant Applications. Available: https://education.ufl.edu/educational-research/5499 (Accessed 24 March 2023)

- van Arensbergen, P., van der Weijden, I., & van der Besselaar, P. (2014). Academic talent selection in grant review panels. *(Re)Searching Scientific Careers*, 25–54.

- Viner, N., Powell, P., & Green, R. (2004). Institutionalized biases in the award of research grants: a preliminary analysis revisiting the principle of accumulative advantage. *Research Policy*, *33*(3), 443–454. https://www.sciencedirect.com/science/article/abs/pii/S0048733303001471

- Viner, N., Green, R., & Powel, P. (2006). Segmenting academics: Resource targeting of research grants. *Science and Public Policy*, *33*(3), 166–178. https://doi.org/10.3152/147154306781779037

- Vingre, A., Kolarz, P., & Bryan, B. (2022). On Your Marks, Get Set, Fund! Rapid Responses to the COVID-19 Pandemic. *Fteval*, *53*, 8–17. https://doi.org/10.22163/fteval.2022.538

- VolkswagenStiftung. (n.d.). *Partially Randomized Procedure - Lottery and Peer Review*. https://www.volkswagenstiftung.de/en/partially-randomized-procedure-lottery-and-peer-review

- Wagner, C. S., & Alexander, J. (2013). Evaluating transformative research programmes: A case study of the NSF Small Grants for Exploratory Research programme. *Research Evaluation*, *22*(3), 187–197. https://doi.org/10.1093/reseval/rvt006

- Walsweer, T. (2018). *Give Chance a Chance*. VolkswagenStiftung, May 22, 2018. https://www.volkswagenstiftung.de/en/news/story/give-chance-chance

- Weber-Main, A. M., McGee, R., Boman, K. E., Hemming, J., Hall, M., Unold, T., Harwood, E. M., Risner, L. E., Smith, A., Lawson, K., Engler, J., Steer, C. J., Buchwald, D., Jones, H. P., Manson, S. M., Ofili, E., Schwartz, N. B., Vishwanatha, J. K., & Okuyemi, K. S. (2020). Grant application outcomes for biomedical researchers who participated in the National Research Mentoring Network's Grant Writing Coaching Programs. *PLoS ONE*, *15*(11 November). https://doi.org/10.1371/journal.pone.0241851

- *What happens after you submit your proposal*. (n.d.). UKRI. Retrieved February 20, 2023, from https://www.ukri.org/councils/epsrc/guidance-for-applicants/what-happens-after-you-submit-your-proposal/

- Wigginton, N., et al. (2021). Managing internal nomination and peer review processes to reduce bias. University of Michigan. Available at: https://research.umich.edu/wp-content/uploads/2021/11/guidance_for_equitably_managing_nomination_and_peer_review_process_to_reduce_bias_on_web_2-9-21_9_pm.pdf (Accessed: 23 March 2023)

- Wieczorkowska, G., & Kowalczyk, K. (2021). Ensuring sustainable evaluation: How to improve quality of evaluating grant proposals? *Sustainability (Switzerland)*, *13*(5), 1–11. https://doi.org/10.3390/su13052842

- de Winde, C. M., Sarabipour, S., Carignano, H., Davla, S., Eccles, D., Hainer, S. J., Haidar, M., Ilangovan, V., Jadavji, N. M., Kritsiligkou, P., Lee, T.-Y., & Ólafsdóttir, H. F. (2021). Towards inclusive funding practices for early career researchers. *Journal of Science Policy & Governance*, *18*(01). https://doi.org/10.38126/jspg180105

- Wyatt, J. (2017). *A systematic map and review of changes to peer review to help build the evidence base for peer review innovations Aims of the systematic map & review* (Issue June). University of Southampton. https://www.nwo.nl/en/international-conference-2017

- Yip, P. S. F., Xiao, Y., Wong, C. L. H., & Au, T. K. F. (2020). Is there gender bias in research grant success in social sciences? Hong Kong as a case study. *Humanities and Social Sciences Communications*, *7*(1), 1–10. https://doi.org/10.1057/s41599-020-00656-y

Appendix B   Survey details

We received 241 survey responses from UKRI staff members. For questions where we requested written responses (as opposed to simple point-and-click survey items), up to 61 respondents per question provided such information.

Our survey was distributed through UKRI's own communication channels rather through person-targeted e-mail invitations from us directly. This means we cannot fully track responses and any self-selection biases. However, we asked a number of questions in the survey to gather personal identifiers (research council, gender, role-type, years of experience), which allows us to check whether we have a response pool that is representative of UKRI as a whole. We can note:

- Around 20% of respondents are from MRC, 6.7% from AHRC and we had just one response from Research England. All other parts of UKRI are represented with in 8-12% of our response pool

- 67% report that they are or have been involved in the design of funding opportunities and 83% report that they are or have been involved in the implementation of funding opportunities

- 63% state that they are female and 30% state that they are male

- Respondents represent a broad range of experience levels, with 33% reporting 0-4 years' experience in research funding, and 13% reporting more than 20 years

The table below lists the dates and channels through which the survey was distributed. We are confident that all staff have been made aware of the survey and would therefore have at least had a chance to participate.

Table 4   List of survey distribution activities

Tracking survey distribution	Date	Reach
Organisation-wide distribution		
News Article - Source	02-Feb	Can be viewed by all staff
The Cascade	w/c 6-Feb	Received by Grade G's and above to cascade to team members
Wednesday Webinar - presentation slot, link in the chat	08-Feb	All staff invited, about 1,500 in attendance
The Stream	14-Feb	All staff
Internal Groups		
PCT Forum/Culture network	03-Feb	Approx. 90 people
FPOG	02-Feb	~40 people, asked to cascade
Talent Strategic Leadership Group	02-Feb	A small group
Behaviours and Incentive Strategic Leadership Group	03-Feb	~10 people
GRECON	03-Feb	~24 people
PAG	06-Feb	A small group

B.1. Survey script and raw data

Please select the Council you work for		
Answer Choices	Responses	
Arts and Humanities Research Council	6.67%	16
Biotechnology and Biological Sciences Research Council	11.25%	27
Economic and Social Research Council	12.08%	29
Engineering and Physical Sciences Research Council	11.25%	27
Innovate UK	9.58%	23
Medical Research Council	20.42%	49
Natural Environment Research Council	7.92%	19
Research England	0.42%	1
Science and Technology Facilities Council	9.58%	23
UKRI	10.83%	26
Other (please specify)	0.00%	0
	Answered	240
	Skipped	2

Which of the below best describe your role at UKRI?Tick all that apply.		
Answer Choices	Responses	
I am or have previously been involved in the design of funding opportunities	67.36%	161
I am or have been previously involved in implementing funding opportunities (e.g., publishing funding calls, advising applicants, grants administrators, peer review officers, panel secretariat, organising assessment processes, monitoring and evaluation etc.)	82.85%	198
None of the above. Please briefly describe your role:	6.28%	15
	Answered	239
	Skipped	3

How many years of experience in research funding (at UKRI or other public or private research and innovation funders) do you have?		
Answer Choices	Responses	
0-4 years	32.50%	78
5-9 years	26.25%	63
10-14 years	18.75%	45
15-19 years	9.17%	22
20 years or more	13.33%	32
	Answered	240
	Skipped	2

What is your gender?		
Answer Choices	Responses	
Female	62.76%	150
Male	30.96%	74

Non-binary	0.00%	0
Prefer not to say	5.86%	14
Prefer to self-describe:	0.42%	1
	Answered	**239**
	Skipped	**3**

Are you aware of any of the following activities being used (past or present) in peer review assessment processes at UKRI?		
Answer Choices	Responses	
None of the above	0.99%	2
Assessment criteria definition: Adding new assessment criteria; may involve a tiered system for assessment criteria, for example, essential vs. desirable	78.82%	160
Demand management: individuals: Stipulating the number of projects that an individual can be involved in as PI and/or Co-I, for a particular Opportunity	64.53%	131
Demand management: individuals: Limiting the number of applications an individual can submit, if the quality of their previous applications has been of lower quality over a certain time period (e.g. only one proposal allowed for next 12 months if repeatedly unsuccessful in the previous 24 months)	47.29%	96
Demand management: institutions: Limiting the number of applications accepted from a single institution	69.46%	141
Demand management: institutions: Limiting the number of re-submissions accepted from a single institution	22.17%	45
Positive action: working with underrepresented groups: Providing additional support to groups that are unrepresented in UKRI's portfolio to encourage them to apply and support them as they do, with the view to increasing diversity	38.92%	79
Applicant behaviours: Designing application forms and processes with a view to encouraging positive behaviours among applicants (e.g. removing hierarchies of applicants to encourage team work and collaboration)	48.28%	98
Expression of interest: A short document providing the information of the organisation and applicant interested in applying. Used as tool to understand interest in the funding opportunity	86.70%	176
Outline applications: A short proposal containing the key information that is assessed to understand the project's merit and determine whether it has sufficient potential for the applicant to be invited to develop the application further	85.71%	174
Reducing application length/cutting sections: Shortening application forms (page/word length) to reduce burden. Requiring only project description and not track, or cutting other sections	62.56%	127
'Sandpits'/Matching events: In UKRI, this might look like an interactive workshop, structured to drive lateral thinking and radical approaches to address research challenges, with the aim of producing research proposals. At the end of the process grants are agreed	55.17%	112
2-stage application process: Not all the information required to make the final decision is included in the first submission	75.37%	153
Applicant anonymisation: Reviewers or panels members or both do not see the identity of the applicant/s	35.47%	72
Automated reviewer allocation: Using algorithms/ AI / text recognition to allocate reviewers to applications	17.24%	35
Dragon's den-style pitch: Applicants are invited to pitch their proposal in front of a panel, and panels have an opportunity to ask questions. This differs from an interview in that no other form of evidence (e.g. written proposals or external expert review) is used in the assessment	18.72%	38
External review only (no panel): Proposals are only assessed by external reviewers and review scores are simply combined to give the final score	28.57%	58
Group review: The same reviewer comments on multiple applications	43.84%	89
Number of reviewers2-3 external reviews of applications is typical, but this number may be lowered to 1 or significantly increased	68.47%	139
Interviews: Lead applicant (or several application team members) may do a presentation (optional) and are then asked questions on their application by panel members, reviewers or funder representatives	82.76%	168
Internal assessment of reviews: Reviews are processed internally by funding organisation staff and are only passed to the external panel if of sufficient quality	59.11%	120
Assessment panels: Assessment panels use external reviews alongside their own expertise to assess the proposal	84.73%	172
Moderation panel: Moderation panels do not use their own expertise but can only use the reviews to inform their scores	58.13%	118

Panel only (no postal/external review): Proposals are only assessed by a panel of experts	76.85%	156
Peer allocation: The applicants are also the assessors, and review the proposals they are competing against to decide In UKRI, this may be known as "Pitch to Peers"	8.87%	18
Office decision: Applications go directly to the 'office' i.e. scheme manager/team /SRO/director, who can recommend funding or even decide to fund unilaterally. No peer and panel review involved.	37.44%	76
Standing panels: The same members year on year with some replacement due to retirement from the panel	63.55%	129
Portfolio panels: Assembled based on the proposals received and therefore will be comprised differently in each round of funding)	56.65%	115
Use of international assessors: Having quotas for assessors based in countries other than the funder's 'home' country. May extend to mandating all-international panels and/or reviewers	50.25%	102
Use of metrics: Use of metrics and bibliometrics as part of the evidence-base to inform decision-making	12.32%	25
Use of non-academic assessors (i.e. industry, policy & practice, patients, 'user' representatives): May extend to all-user panels and/or reviewers. May take the shape of consultation rather than direct decision-making. May or may not involve specific quotas	73.40%	149
Virtual panels: Convening panels online rather than in person	86.21%	175
Golden ticket/Joker[wildcard] Each panel members (or other decision-maker) is able to select one proposal (e.g. per call, per year, or similar) to guarantee funding, regardless of panel rankings or other decision-making processes	1.97%	4
Lottery [Partial randomisation]: Successful proposals are chosen at random. In most methodologies, randomisation is only partial. For example, proposals may be scored and sorted into bands, and only those on the border of being funded will be randomised.	15.76%	32
Scoring mechanisms: Including voting, weighting, variance-based scoring	60.59%	123
Sequential application of criteria (rather than simultaneous application of criteria)A proposal is scored for one set of criteria, ranked and a cut-off point determined. Then those above the cut-off point are assessed again for another set of criteria to determine the final funded this	15.27%	31
Use of quotas: After ranking, proposals are reviewed to ensure sufficient numbers in certain categories including positive action (quotas related to protected characteristics) or quotas related to place, themes, disciplines	26.11%	53
Bringing in reviewers from earlier careers & providing mentoring: Panels and reviewers tend to be very experienced researchers/innovators. Those early in their careers could be invited to review or be part of panels with additional training, bringing different perspectives and experiences. Previous calls' award winners may also be brought in as reviewers/panellists	41.38%	84
Embedding EDI in assessment: Training or support provided to make assessors aware of their unconscious biases and to encourage them to call each other out during the assessment process	73.40%	149
Expanding or reducing the amount/detail of feedback to unsuccessful applicants: Different levels of feedback may be provided on unsuccessful applications	60.10%	122
Funder representation on review panels: The funder is represented on the panel to guide discussion or provide briefing on programme aims. Their role is beyond a purely administrative function, they may even be in a chair-role or similar	48.28%	98
Improving quality of reviews: Through training/retaining good reviewers/recognition. May be done through peer review colleges	48.77%	99
Open review/rebuttal: Reviews are published and/or made available to the applicant before decisions are taken, so they can be viewed and responded to.	53.20%	108
Other activity/-ies to improve baseline peer review assessment process not listed above. We encourage you to report also small and incremental tweaks or experiments aiming to improve the assessment process. Please provide details on the rationale for the activity, a brief description of the activity and evidence on the effectiveness of the activity.	19.21%	39
	Answered	**203**
	Skipped	**39**

Please provide a brief description of the activity/-ies. Please provide details such as the title of the funding opportunity where the activity/-ies was introduced, funding opportunity objectives, a brief description of how the activity/-ies was introduced and how it differs from the baseline peer review process, and the challenges (if any) with introducing the activity/-ies. Please provide details on any other activities introduced to improve the baseline assessment process that were introduced to the same funding opportunity. Feel free to share anything you consider relevant:	
Answered	**61**
Skipped	**181**
[Free-text responses]	

Please share your insights on the rationale for the activity/-ies. This relates to the reasons why the activity was introduced. Examples of rationale might be encouraging wider and more diverse participation in the funding opportunity, encouraging high potential, disruptive research proposals, reducing the peer review burden, the need to assess non-research/innovation criteria or a mix of the above (and other). Please explain what were the expected outcomes of the activity/-ies (e.g., greater diversity, reduced burden, etc.).Feel free to share anything you consider relevant:	
Answered	**54**
Skipped	**188**
[Free-text responses]	

Please share your insights on the effectiveness of the activity/-ies. Please provide details on the evidence of the effectiveness; for example, has an evaluation of the activity/-ies or the funding opportunity been completed or is planned, and what other evidence on the effectiveness is available (e.g., staff observations, analysis of monitoring data, etc.). Please provide details on what the evidence on the effectiveness tells – did the activity/-ies achieve intended objectives, what worked well and less well, why and what are the lessons learned. Please also highlight any issues and weaknesses of the activity/-ies, unintended consequences, etc.Feel free to share anything you consider relevant:	
Answered	**50**
Skipped	**192**
[Free-text responses]	

Please provide a brief description of the activity/-ies. Please provide details such as the title of the funding opportunity where the activity/-ies was introduced, funding opportunity objectives, a brief description of how the activity/-ies was introduced and how it differs from the baseline peer review process, and the challenges (if any) with introducing the activity/-ies. Please provide details on any other activities introduced to improve the baseline assessment process that were introduced to the same funding opportunity. Feel free to share anything you consider relevant:	
Answered	**35**
Skipped	**207**
[Free-text responses]	

Please share your insights on the rationale for the activity/-ies. This relates to the reasons why the activity was introduced. Examples of rationale might be encouraging wider and more diverse participation in the funding opportunity, encouraging high potential, disruptive research proposals, reducing the peer review burden, the need to assess non-research/innovation criteria or a mix of the above (and other). Please explain what were the expected outcomes of the activity/-ies (e.g., greater diversity, reduced burden, etc.).Feel free to share anything you consider relevant:	
Answered	**31**
Skipped	**211**
[Free-text responses]	

Please share your insights on the effectiveness of the activity/-ies. Please provide details on the evidence of the effectiveness; for example, has an evaluation of the activity/-ies or the funding opportunity been completed or is planned, and what other evidence on the effectiveness is available (e.g., staff observations, analysis of monitoring data, etc.). Please provide details on what the evidence on the effectiveness tells – did the activity/-ies achieve intended objectives, what worked well and less well, why and what are the lessons learned. Feel free to share anything you consider relevant:	
Answered	**26**
Skipped	**216**
[Free-text responses]	

Please select an activity/-ies to improve the baseline peer review assessment process that you would like to be used more in the assessment process at UKRI.		
Answer Choices	Responses	
Assessment criteria definition: Adding new assessment criteria; may involve a tiered system for assessment criteria, for example, essential vs. desirable	39.78%	37
Demand management: individuals: Stipulating the number of projects that an individual can be involved in as PI and/or Co-I, for a particular Opportunity	27.96%	26

Demand management: individuals: Limiting the number of applications an individual can submit, if the quality of their previous applications has been of lower quality over a certain time period (e.g. only one proposal allowed for next 12 months if repeatedly unsuccessful in the previous 24 months)	19.35%	18
Demand management: institutions: Limiting the number of applications accepted from a single institution	23.66%	22
Demand management: institutions: Limiting the number of re-submissions accepted from a single institution	15.05%	14
Positive action: working with underrepresented groups: Providing additional support to groups that are unrepresented in UKRI's portfolio to encourage them to apply and support them as they do, with the view to increasing diversity	62.37%	58
Applicant behaviours: Designing application forms and processes with a view to encouraging positive behaviours among applicants (e.g. removing hierarchies of applicants to encourage team work and collaboration)	61.29%	57
Expression of interest: A short document providing the information of the organisation and applicant interested in applying. Used as tool to understand interest in the funding opportunity	33.33%	31
Outline applications: A short proposal containing the key information that is assessed to understand the project's merit and determine whether it has sufficient potential for the applicant to be invited to develop the application further	32.26%	30
Reducing application length/cutting sections: Shortening application forms (page/word length) to reduce burden. Requiring only project description and not track, or cutting other sections	27.96%	26
'Sandpits'/Matching events: In UKRI, this might look like an interactive workshop, structured to drive lateral thinking and radical approaches to address research challenges, with the aim of producing research proposals. At the end of the process grants are agreed	20.43%	19
2-stage application process: Not all the information required to make the final decision is included in the first submission	33.33%	31
Applicant anonymisation: Reviewers or panels members or both do not see the identity of the applicant/s	40.86%	38
Automated reviewer allocation: Using algorithms/ AI / text recognition to allocate reviewers to applications	21.51%	20
Dragon's den-style pitch: Applicants are invited to pitch their proposal in front of a panel, and panels have an opportunity to ask questions. This differs from an interview in that no other form of evidence (e.g. written proposals or external expert review) is used in the assessment	6.45%	6
External review only (no panel): Proposals are only assessed by external reviewers and review scores are simply combined to give the final score	3.23%	3
Group review: The same reviewer comments on multiple applications	22.58%	21
Number of reviewers2-3 external reviews of applications is typical, but this number may be lowered to 1 or significantly increased	22.58%	21
Interviews: Lead applicant (or several application team members) may do a presentation (optional) and are then asked questions on their application by panel members, reviewers or funder representatives	21.51%	20
Internal assessment of reviews: Reviews are processed internally by funding organisation staff and are only passed to the external panel if of sufficient quality	15.05%	14
Assessment panels: Assessment panels use external reviews alongside their own expertise to assess the proposal	35.48%	33
Moderation panel: Moderation panels do not use their own expertise but can only use the reviews to inform their scores	13.98%	13
Panel only (no postal/external review): Proposals are only assessed by a panel of experts	19.35%	18
Peer allocation: The applicants are also the assessors, and review the proposals they are competing against to decide In UKRI, this may be known as "Pitch to Peers"	4.30%	4
Office decision: Applications go directly to the 'office' i.e. scheme manager/team /SRO/director, who can recommend funding or even decide to fund unilaterally. No peer and panel review involved.	11.83%	11
Standing panels: The same members year on year with some replacement due to retirement from the panel	18.28%	17
Portfolio panels: Assembled based on the proposals received and therefore will be comprised differently in each round of funding)	22.58%	21
Use of international assessors: Having quotas for assessors based in countries other than the funder's 'home' country. May extend to mandating all-international panels and/or reviewers	23.66%	22
Use of metrics: Use of metrics and bibliometrics as part of the evidence-base to inform decision-making	4.30%	4
Use of non-academic assessors (i.e. industry, policy & practice, patients, 'user' representatives): Use of non-academic assessors (i.e. industry, policy & practice, patients, 'user' representatives)May extend to all-user panels and/or reviewers. May take the shape of consultation rather than direct decision-making. May or may not involve specific quotas	44.09%	41

Virtual panels: Convening panels online rather than in person	27.96%	26
Golden ticket/Joker [Wildcard] Each panel members (or other decision-maker) is able to select one proposal (e.g. per call, per year, or similar) to guarantee funding, regardless of panel rankings or other decision-making processes	6.45%	6
Lottery: Successful proposals are chosen at random. In most methodologies, randomisation is only partial. For example, proposals may be scored and sorted into bands, and only those on the border of being funded will be randomised.	22.58%	21
Scoring mechanisms: Including voting, weighting, variance-based scoring	17.20%	16
Sequential application of criteria (rather than simultaneous application of criteria): A proposal is scored for one set of criteria, ranked and a cut-off point determined. Then those above the cut-off point are assessed again for another set of criteria to determine the final funded this	12.90%	12
Use of quotas: After ranking, proposals are reviewed to ensure sufficient numbers in certain categories including positive action (quotas related to protected characteristics) or quotas related to place, themes, disciplines	9.68%	9
Bringing in reviewers from earlier careers & providing mentoring: Panels and reviewers tend to be very experienced researchers/innovators. Those early in their careers could be invited to review or be part of panels with additional training, bringing different perspectives and experiences. Previous calls' award winners may also be brought in as reviewers/panellists	61.29%	57
Embedding EDI in assessment: Training or support provided to make assessors aware of their unconscious biases and to encourage them to call each other out during the assessment process	50.54%	47
Expanding or reducing the amount/detail of feedback to unsuccessful applicants: Different levels of feedback may be provided on unsuccessful applications	24.73%	23
Funder representation on review panels: The funder is represented on the panel to guide discussion or provide briefing on programme aims. Their role is beyond a purely administrative function, they may even be in a chair-role or similar	17.20%	16
Improving quality of reviews: Through training/retaining good reviewers/recognition. May be done through peer review colleges	49.46%	46
Open review/rebuttal: Reviews are published and/or made available to the applicant before decisions are taken, so they can be viewed and responded to.	26.88%	25
None of the above	2.15%	2
Other activity/-ies to improve baseline peer review assessment process not listed above. Please provide details on the rationale for the activity, a brief description of the activity and evidence on the effectiveness of the activity.	17.20%	16
	Answered	**93**
	Skipped	**149**

Why would you like to see increased use of this activity/-ies to improve peer review?	
Answered	**56**
Skipped	**186**
[Free-text responses]	

Please feel free to share any further thoughts or reflections you have on the peer review assessment process and activities to improve it.	
Answered	**29**
Skipped	**213**
[Free-text responses]	

As part of this study, we plan to conduct a small number of follow-up interviews. May we contact you via e-mail to discuss your answers to this survey? If yes, please provide your e-mail address.	
Answered	**47**
Skipped	**195**
[Free-text responses]	

Appendix C   Interview details

C.1. Interviewees

Table 5   List of interviewees

Name	Organisation	Role	Interview date
UKRI staff (survey follow-up interviews)			
James Sundquist	UKRI, BBSRC	Senior Portfolio Manager	22/02/2023
Liam Blackwell	UKRI, EPSRC	Deputy Director for Cross Council Programmes	21/02/2023
Georgina Freeman	UKRI, STFC	Senior Programme Manager	27/02/2023
Stephen Meader	UKRI	Director, Future Leaders Fellowships	04/04/2023
Laura Bones	UKRI	Senior Programme Manager	02/03/2023
Other UK funders			
Sue Russel and Amy Bradburn	Cancer Research UK	Senior Policy & Governance Manager/Head of Grants Management	15/02/2023
Paul McDonald	Royal Society	Head of Grants	02/02/2023
Alyson Fox	Wellcome	Director of Research Funding	01/02/2023
Ken Emond	British Academy	Head of Research Awards	21/02/2023
Vicky Taylore	National Institutes of Health Research	Assistant Director for Applications and Funding	08/03/2023
International funders			
Ulrike Bischler	Volkswagen Foundation	Director of Grants	02/02/2023
Sylvia Jeney	Swiss National Science Foundation	Head of Open Research Data, previously Spark programme manager	08/02/2023
Kristin Oxley	Research Council Norway	Senior Adviser	03/02/2023
Kristin M. Kramer	National Institutes of Health	Director at the Office of Communications and Outreach, Center for Scientific Review	16/02/2023
Uwe von Ahsen	FWF	Head of Strategy Department	24/02/2023
Research and innovation policy experts			
James Wilsdon	Research on Research Institute	Director	13/02/2023
Adrian Barnett	Queensland University of Technology	Professor	20/02/2023
Jenny Gladstone	University of Oxford	Strategic Research Development Manager	14/02/2023

Gemma Derick	University of Bristol	Associate Professor	24/02/2023
Sector representatives			
Joanna Burton	Russel Group*	Policy Manager	14/02/2023
Nicola Eckersley-Waites	Confederation of British Industry	Head of Innovation	16/02/2023
Daniel Wake	Universities UK*	Policy manager	03/03/2023

Note: *organisations collected feedback on the study questions from their members and shared a summary of feedback. Russel Group representative summarised the feedback also in an interview.

C.2. Interview tool

Name	
Institution/organisation	
Role	
Interview date/time	
Interviewer	

Points to make before the start of the interview:

This interview is part of the Review of Peer Review Study commissioned by UKRI to Technopolis. The study will analyse evidence on the effectiveness of interventions in the peer review process. 'Interventions' refers to any form of deviation from the standard application assessment process used by research and innovation funders involving external peer review and panel review.

No attributable quotes will be used from these interviews. However, can we please note your name in the method annex to our final report? You have the right to withdraw your participation at any time.

Interview questions

- Please can you describe your role at your organisation?
- Could you please provide a brief description of your organisation's standard peer review assessment process?
- What (if any) problems have you identified with the standard peer review assessment process?

The following questions are not relevant to all interviewees. Please focus on the programmes and interventions that we know the funder has introduced and can reflect upon.

- Could you please briefly summarise what interventions to the standard peer review assessment process your organisation has introduced? 'Interventions' refers to any form of deviation from the standard application assessment process. Interventions can include various significant modifications and smaller process tweaks to the 'standard' peer review process for grant allocations. Examples could include:

- **pre-call interventions** such as use of quotas, specific eligibility requirements
- **interventions around application design** such as pre-applications, application time window variations
- **interventions around process design** such as the use of interviews, anonymised review, pitching
- **interventions around decision-making** such as lottery, wildcard and
- **interventions to support training and feedback** such as training of reviewers, applicant rebuttal and similar

- What was the rationale for the introduction of the intervention/s?

Prompt for:

- encouraging wider and more diverse participation in the funding programme
- encouraging riskier, disruptive research proposals
- reducing the peer review burden
- the need to assess non-academic criteria
- a mix of the above

- Could you please briefly describe the implementation of the intervention/s? Prompt for:
  - What were the objectives of the funding scheme where the intervention/s was introduced?
  - A brief description of the intervention
  - What (if any) were the practical challenges when introducing the intervention?
  - Did the introduction of the intervention require additional resources and specific staff competence?

- What type of evidence do you have on the effectiveness of the intervention/s? Prompt for:
  - Strong evidence like controlled experiments
  - Light-touch evaluation
  - Anecdotal staff observations

- What does the evidence on the effectiveness of the intervention/s tell? Prompt for:
  - Did the intervention achieve the intended objectives and how exactly?
  - Are there any unintended consequences?
  - Are there any lessons for what type of programme the intervention/s works better?

- Have you implemented multiple interventions to one programme? For example, a two-stage application process and unconscious bias training.
  - If 'yes', do you have any observation and lessons learned on the effectiveness of combinations of interventions?

- Have any interventions we discussed become (or will become) a 'new normal' in your organisation's funding process?

- Overall, is there an appetite in your organisation, organisations that oversee your work and wider academic community to address problems around standard peer review and experiment with interventions around peer review?

- Do you have any other thoughts on your experience and wider observations on the interventions around peer review that we have not covered yet?

Appendix D   State of play at UKRI

Our study is intended to be of use to the widest possible audience of R&I funders, and so while we draw many examples from UKRI, we do not reflect in the main body on UKRI itself. However, as consultation of UKRI staff was a major part of our data collection, we are able to describe in more detail the views and hopes within UKRI. We present these findings below.

We note at the outset that our study was in no way a review of practices at UKRI, let alone an evaluation of them. What is compiled below is a snapshot of views rather than of facts. They may provide some guidance on how UKRI might go about implementing interventions to peer review. Additionally, they present a sample case study of the current perceptions and attitudes within one funder – other funders may recognise themselves in some of these findings, or take them on board in their considerations of how best to evolve their processes.

Several UKRI staff consultees for this study pointed out that this study is perceived to be important as it will provide UKRI staff with a resource to inform their work. Some consulted UKRI staff members noted that too little sharing of experiences with implementation of the interventions is happening within the organisation. As a result, people often feel as though they are doing things for the first time when in fact they are not.

A total of 203 members of staff across seven research councils (AHRC, BBSRC, ESRC, EPSRC, MRC, NERC and STFC), Innovate UK and central UKRI responded to a survey question asking which of the 38 interventions to the peer review process they had heard of being used at UKRI. The responses showed organisation, intervention and intervention-type based variations.

Intervention-type awareness

Of the five intervention types, interventions in the 'application parameters' cluster are best known; on average, 69% of all respondents noted being aware of the interventions in this bracket being used (past or present) at UKRI. It is also the only bracket in which every intervention was known by at least a third of every part of UKRI.

The interventions at the decision-making stage appear to have the lowest profile at UKRI. The best-known intervention in the 'decision-making' cluster was scoring mechanisms, reported by 60% of all respondents to have been used at UKRI. The rest of the interventions in this bracket were considerably less well-known, with only 24% of respondents reporting familiarity of included interventions on average.

Intervention-level awareness

At a more granular level, the most widely known interventions are around elements in the application and process design stages. As shown in Figure 3, expressions of interest and outlines are reported to have been used at UKRI by 86% and 85% of respondents respectively. In the process design stage, virtual panels were familiar to 86% of respondents, followed by assessment panels (at 84%) and applicant interviews (at 82%). It is possible that virtual proceedings became considerably better known as a result of the restrictions related to the Covid-19 pandemic. Furthermore, these are interventions which can be considered to either provide more robust information (interviews, assessment panels, virtual panels) or expedite the process (outlines and EoIs). It can be argued that they do not significantly alter the overall process (compared to interventions like partial randomisation or matching events for instance, where the decision-making or applicant behaviours change entirely). This, in turn, could suggest that the threshold to experiment with these interventions is relatively low.

In line with the intervention-type level awareness at UKRI, three of the five least known interventions came from the decision-making cluster. Instances selecting awards via wildcards

or randomisation were reported by 2% and 16% of respondents respectively. In addition, sequential assessment criteria were reportedly familiar to 15% of respondents. Similar levels of familiarity were reported on the use of metrics and peer allocation of reviews in the process-design bracket. 12% of respondents reported knowledge of use of metrics, and 9% of peer allocation.

Figure 3   The five most and least well-known interventions to peer review by UKRI staff
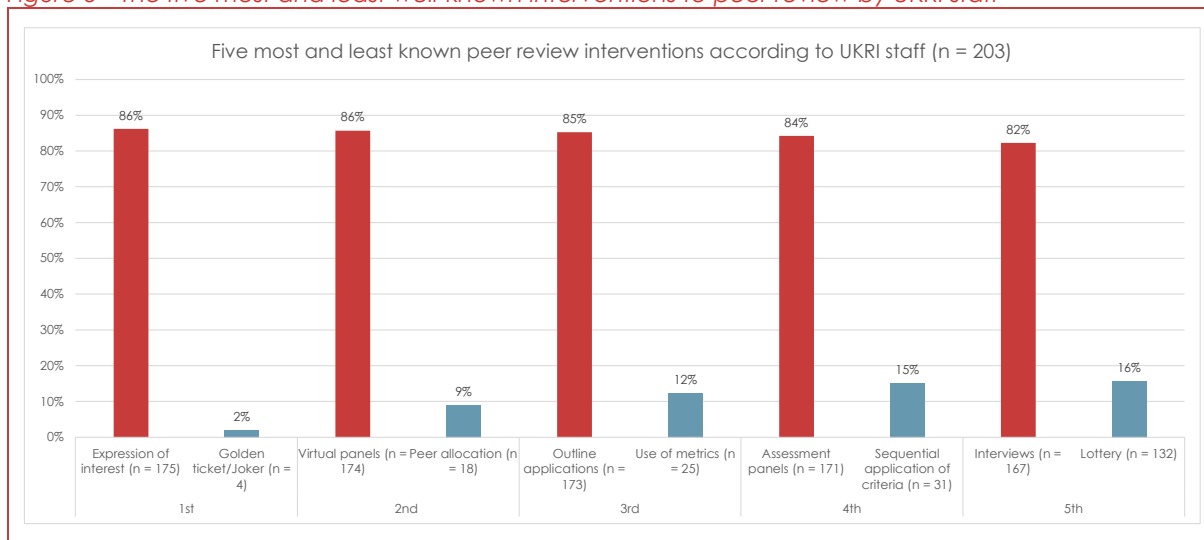
Source: Technopolis survey. Wording of survey question: "Are you aware of any of the following activities being used (past or present) in peer review assessment processes at UKRI?"

For many interventions, staff in one part of UKRI show far more awareness of them than in others, which likely indicates use or applicability of the intervention in one particular Research Council (or in a small number of them). Interventions with particular fluctuation in familiarity between councils included moderation panels and standing panels, individual demand management (based on previous performance) and early career reviewers. These are likely the instances where experience-sharing across the organisation will be of the greatest benefit. It is also notable that there were differences between Research Councils in the overall level of awareness of different interventions as shown in Figure 4. This may demonstrate general differences in the readiness to experiment with various interventions, but also in the accumulated knowledge.

There were 39 responses to the open 'other interventions' survey question which contained elements from a wide range of the 38 interventions. These occurred in specific combinations which had not been included (e.g., tweaking eligibility to enable early career researchers to apply as PIs), or in ways which were more specific in nature to (e.g., lived experience experts as reviewers rather than non-academic reviewers). The following points emerge from this, though we note that many are merely slight modifications from our original set of 38 interventions:

- Five respondents mentioned systematic training of, or guidance for reviewers or panels for embedded EDI or elevated quality of assessment
- Five respondents mentioned designing the application forms to capture demonstrations of EDI (e.g., via tweaking eligibility to include early career researchers), good practices or emphasis on critical issues to the call. A small subset of this was particularly aimed at removing information about applicant track

- Three respondents mentioned the right to reply. This is perhaps understood separately from open reviews / rebuttals, as emphasising written dialogue between applicants and reviewers

- Three respondents mentioned lived experience experts as either reviewers or panel members

- Four respondents mentioned ways to expedite the panel process or reduce workload per individual panellists. This would be done by triaging or banding applications to focus panel time, or by pre-excluding applications before panel stage where they fail to receive a sufficient number of high scores

Figure 4   Survey results: awareness of interventions being used at UKRI

	AHRC 12	BBSRC 26	ESRC 23	EPSRC 26	IUK 18	MRC 43	NERC 13	STFC 18	UKRI 21	Total
Assessment criteria	58%	88%	91%	92%	67%	72%	77%	78%	71%	78%
Demand management: individuals1	58%	81%	70%	92%	44%	58%	62%	39%	67%	64%
Demand management: individuals2	8%	50%	22%	92%	28%	44%	38%	56%	62%	47%
Demand management: institutions1	42%	73%	78%	92%	44%	58%	62%	72%	90%	69%
Demand management: institutions2	0%	15%	26%	27%	28%	16%	31%	11%	38%	22%
Positive action	25%	38%	17%	50%	39%	37%	38%	39%	62%	38%
Applicant behaviours	50%	62%	43%	62%	17%	51%	54%	33%	48%	48%
Expression of interest	92%	100%	91%	96%	67%	79%	92%	89%	81%	86%
Outline applications	58%	92%	91%	96%	44%	93%	92%	89%	86%	85%
Reducing application length	33%	73%	70%	77%	33%	58%	77%	56%	71%	62%
'Sandpits'/Matching events	33%	65%	52%	92%	33%	56%	46%	33%	57%	55%
2-stage application process	50%	77%	96%	85%	67%	72%	85%	56%	76%	75%
Applicant anonymisation	17%	31%	48%	88%	22%	21%	23%	28%	29%	35%
Automated reviewer allocation	8%	12%	17%	19%	22%	14%	23%	17%	24%	17%
Dragon's den-style pitch	8%	12%	22%	54%	28%	5%	0%	22%	19%	19%
External review only	25%	27%	30%	35%	61%	9%	23%	33%	33%	29%
Group review	17%	27%	65%	65%	33%	42%	23%	67%	38%	43%
Number of reviewers	75%	88%	70%	62%	39%	70%	62%	67%	76%	68%
Interviews	58%	92%	91%	96%	67%	81%	85%	72%	81%	82%
Internal assessment of reviews	50%	54%	65%	92%	61%	44%	85%	44%	48%	59%
Assessment panels	92%	88%	100%	85%	44%	86%	100%	78%	86%	84%
Moderation panel	92%	50%	43%	92%	50%	30%	92%	39%	76%	58%
Panel only	75%	96%	91%	88%	28%	67%	92%	78%	71%	76%
Peer allocation	0%	0%	30%	4%	6%	5%	0%	11%	19%	9%
Office decision	25%	46%	26%	77%	17%	19%	46%	33%	48%	37%
Standing panels	25%	62%	83%	62%	22%	77%	54%	83%	67%	63%
Portfolio panels	42%	69%	61%	77%	44%	49%	54%	50%	62%	57%
international assessors	42%	54%	43%	69%	28%	56%	31%	50%	57%	50%
Use of metrics	0%	12%	9%	12%	11%	16%	0%	22%	19%	12%
Non-academic assessors	58%	92%	96%	81%	67%	65%	62%	44%	81%	73%
Virtual panels	83%	92%	87%	96%	61%	88%	92%	89%	76%	86%
Golden ticket/Joker	17%	0%	4%	0%	0%	2%	0%	0%	0%	2%
Randomisation	17%	23%	9%	15%	6%	7%	38%	28%	14%	16%
Scoring mechanisms	50%	58%	52%	77%	44%	70%	62%	67%	48%	60%
Sequential application of criteria	0%	15%	4%	23%	11%	14%	23%	28%	14%	15%
Use of quotas	17%	15%	26%	42%	28%	23%	31%	17%	29%	26%
Early career reviewers	33%	38%	39%	85%	11%	28%	38%	39%	52%	41%
Embedding EDI in assessment	67%	77%	70%	88%	50%	77%	77%	72%	67%	73%
Expanding or reducing feedback	67%	62%	43%	77%	44%	77%	54%	56%	43%	60%
Funder representation on panels	67%	58%	57%	50%	44%	40%	46%	56%	29%	48%
Training	58%	23%	65%	73%	33%	40%	62%	28%	67%	49%
Open review/rebuttal	42%	58%	70%	62%	17%	51%	54%	67%	52%	53%
Other activity/-ies	17%	23%	22%	19%	17%	16%	8%	33%	19%	19%
None of the above	0%	0%	0%	4%	0%	0%	0%	6%	0%	1%

n=200. Wording of survey question: "Are you aware of any of the following activities being used (past or present) in peer review assessment processes at UKRI?" Note that the wording of some interventions was altered to better reflect standard terminology within UKRI. As the remainder of our research covers non-UKRI and international sources, we opt for more mainstream terminology in the rest of our study. *Demand management for individuals and institutions was broken down further into sub-categories and is now defined as follows: Demand management: individuals 1: Stipulating the number of projects that an individual can be involved in as PI and/or Co-I, for a particular Opportunity; Demand management: individuals 2: Limiting the number of applications an individual can submit, if their previous applications has been of lower quality over a certain time period (e.g. only one proposal allowed for next 12 months if repeatedly unsuccessful in the previous 24 months); Demand management: institutions 1: Limiting the

We also put a forward-looking question to UKRI staff, asking them to indicate which of the interventions to improve the baseline peer review assessment process they would like to be used more in the assessment processes at UKRI. We received 92 responses. The overall levels of approval (as measured by the wish to see interventions more) were somewhat lower than when asked what interventions respondents were already aware of being used at UKRI. This may be explained to some degree by the lower response rate with the same number of organisations making the impact of single responses (or lack thereof) larger. That said, it may also demonstrate a more conservative approach to interventions or their future increase.

Intervention-type level

Enthusiasm appears highest for interventions in the 'application design and parameters' (38% average) and 'training and feedback' (37%) clusters. Across all intervention types, 'application design and parameters' was the only bracket where all interventions received some degree of support for increased future use from all research councils.

Similarly to the level of awareness, the least supported bracket for increased use concerned changes to decision-making with the mean rating at 15%, and where each intervention received no support from at least one research council.

Intervention-level

Across the Research Councils, interventions promoting affirmative actions were generally most sought-after. Working with underrepresented groups[17] was selected by at least half of respondents from every represented organisation, and by as many as 86% of AHRC and 75% of UKRI representatives. It also had the highest overall approval level of 62%.

Other generally well-received interventions included designing application forms and processes with a view to encourage positive behaviours (particularly supported by representatives of UKRI, MRC and EPSRC; 75%, 73% and 70% of respondents respectively), bringing in early career researchers as reviewers (61% overall approval level), embedding EDI in the assessment (51% overall approval level) and review training (50% overall approval level).

Less thematic alignment was identified among the collectively least sought-after interventions. External review only, peer allocation, use of metrics, dragon's dens and wildcards, each were hoped for by less than 10% of the overall respondent population. However, four of the five least hoped for interventions are in the 'process design' cluster and one 'decision-making'.

Compared to the most well-received interventions, several of the five least supported ones propose a considerable shift in the way their respective process stages are carried out. Dragon's den style pitches rely primarily on oral presentation of proposed ideas instead of written proposals, peer allocation shifts some part of the assessment responsibility to applicants, and wildcards place a considerable selection power with individual selectors.

---

[17] In our original survey of UKRI staff, we used the term 'Positive action – working with underrepresented groups'. The term 'positive action' is not in use in the report, as it is too broad (and has multiple definitions) to be treated as a single intervention.

Figure 5 Most and least hoped interventions to peer review for the future by UKRI staff
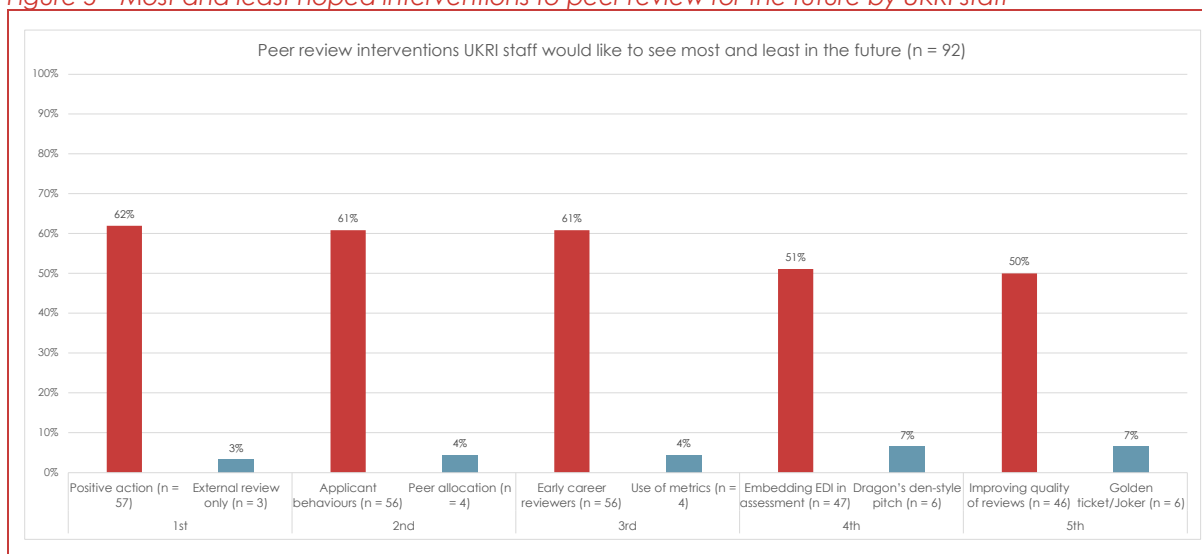
Technopolis survey

The representatives from UKRI (i.e. not attached to any particular Research Council) and NERC appeared positive about the largest number of interventions with at least 50% of UKRI and NERC representatives naming 10 interventions which they would like to see more.

There was also a notable degree of variance between research councils in terms of supported interventions. For instance, working with underrepresented groups was the only intervention which at least 50% of each Research Council rated as something they would like to see increased. Conversely, the use of non-academic assessors was selected by 71% of one research council (AHRC) and 14% of another (STFC). A similarly high level of inter-council variance was identified with applicant anonymisation (75% at most at Innovate UK, 33% at least at STFC). As with the responses to general awareness, these elements indicate differences in perceived suitability between research councils.

Finally, there were 16 answers to the open 'other' question. Again, as with the same question about awareness, there were answers which may have fit an existing intervention but included further detail, and answers which combined elements from different interventions. Some answers also recommended caution rather than a particular intervention per se.

- Four respondents recommended the inclusion of various non-academic populations which, at times, also aligned with the 'working with underrepresented groups' intervention. These populations were lived experience experts (e.g., patients), public engagement specialists, non-academics in academic settings (e.g., technical specialists) and non-clinical end-users (typically form industry)

- Four respondents recommended ways to enable a more robust understanding of applications. Two of the respondents recommended opportunities for discussions for panels ahead of the official panel meeting (e.g., online discussion boards). One respondent recommended co-reviewing, especially for interdisciplinary applications to ensure that the proposed ideas are rounded, while the fourth hoped for review prompts to assess EDI-related outcomes. Training was also mentioned by several of the four.

- Two respondents approached applicant track from different perspectives; one hoped for visibility of past research to ensure that selections are not made based on application-writing skills, while another recommended caution with the use of metrics in responsible research funding

Figure 6   Survey results: Appetite for interventions / increased use of interventions at UKRI

	AHRC 7	BBSRC 13	ESRC 9	EPSRC 10	IUK 4	MRC 22	NERC 5	STFC 14	UKRI 8	Total
Assessment criteria	29%	54%	56%	20%	50%	27%	80%	43%	38%	40%
Demand management: individuals1	0%	31%	0%	30%	25%	27%	40%	50%	38%	28%
Demand management: individuals2	0%	15%	0%	30%	25%	27%	20%	29%	13%	20%
Demand management: institutions1	14%	23%	11%	20%	50%	23%	40%	14%	50%	24%
Demand management: institutions2	0%	23%	0%	0%	25%	18%	20%	14%	38%	15%
Positive action	86%	62%	56%	60%	50%	59%	60%	57%	75%	62%
Applicant behaviours	57%	46%	56%	70%	50%	73%	60%	50%	75%	61%
Expression of interest	43%	54%	22%	20%	25%	27%	60%	29%	38%	34%
Outline applications	14%	54%	11%	10%	25%	45%	40%	36%	25%	33%
Reducing application length	14%	31%	33%	10%	25%	27%	40%	21%	50%	27%
'Sandpits'/Matching events	29%	38%	0%	10%	50%	23%	0%	14%	25%	21%
2-stage application process	14%	38%	44%	10%	25%	45%	60%	29%	25%	34%
Applicant anonymisation	43%	46%	33%	50%	75%	41%	60%	14%	38%	40%
Automated reviewer allocation	0%	38%	11%	30%	0%	32%	0%	21%	13%	22%
Dragon's den-style pitch	14%	0%	22%	0%	0%	5%	20%	0%	13%	7%
External review only	0%	8%	0%	0%	25%	0%	0%	7%	0%	3%
Group review	0%	15%	22%	40%	25%	32%	20%	21%	13%	23%
Number of reviewers	29%	15%	33%	0%	25%	32%	40%	21%	13%	23%
Interviews	0%	23%	0%	0%	25%	41%	20%	21%	38%	22%
Internal assessment of reviews	0%	23%	11%	0%	0%	27%	0%	21%	13%	15%
Assessment panels	29%	38%	33%	0%	50%	45%	60%	29%	50%	36%
Moderation panel	14%	8%	11%	10%	25%	5%	0%	14%	63%	14%
Panel only	14%	31%	22%	0%	0%	23%	20%	14%	38%	20%
Peer allocation	14%	0%	0%	0%	0%	5%	0%	7%	13%	4%
Office decision	14%	15%	0%	0%	0%	18%	20%	7%	25%	12%
Standing panels	0%	15%	22%	0%	0%	36%	20%	21%	13%	18%
Portfolio panels	14%	38%	22%	30%	25%	14%	0%	21%	38%	23%
international assessors	43%	38%	11%	10%	0%	32%	20%	0%	50%	24%
Use of metrics	14%	8%	0%	0%	25%	0%	0%	7%	0%	4%
Non-academic assessors	71%	46%	44%	30%	25%	59%	40%	14%	63%	45%
Virtual panels	43%	46%	22%	20%	25%	23%	20%	21%	38%	28%
Golden ticket/Joker	0%	15%	0%	20%	25%	0%	20%	0%	0%	7%
Randomisation	14%	46%	0%	30%	0%	18%	20%	21%	38%	23%
Scoring mechanisms	14%	8%	0%	20%	50%	23%	20%	14%	13%	16%
Sequential application of criteria	14%	15%	0%	0%	25%	14%	20%	14%	25%	13%
Use of quotas	29%	23%	11%	10%	0%	0%	20%	0%	13%	10%
Early career reviewers	57%	62%	67%	50%	25%	68%	80%	57%	63%	61%
Embedding EDI in assessment	71%	54%	44%	40%	25%	50%	60%	50%	63%	51%
Expanding or reducing feedback	14%	31%	11%	10%	50%	27%	60%	21%	25%	25%
Funder representation on panels	29%	8%	0%	30%	25%	27%	0%	14%	13%	17%
Training	57%	38%	33%	30%	75%	68%	20%	57%	50%	50%
Open review/rebuttal	0%	31%	22%	20%	0%	41%	40%	29%	25%	27%
Other activity/-ies	29%	15%	22%	0%	25%	5%	20%	36%	25%	17%

n=92. Wording of survey question: "Please select an activity/-ies to improve the baseline peer review assessment process that you would like to be used more in the assessment process at UKRI." Note that the wording of some interventions was altered to better reflect standard terminology within UKRI. As the remainder of our research covers non-UKRI and international sources, we opt for more mainstream terminology in the rest of our study. *Demand management for individuals and institutions was broken down further into sub-categories and is now defined as follows: Demand management: individuals 1: Stipulating the number of projects that an individual can be involved in as PI and/or Co-I, for a particular Opportunity; Demand management: individuals 2: Limiting the number of applications an individual can submit, if their previous applications has been of lower quality over a certain time period (e.g. only one proposal allowed for next 12 months if repeatedly unsuccessful in the previous 24 months); Demand management: institutions 1: Limiting the number of applications accepted from a single institution; Demand management: institutions 2: Limiting the number of re-submissions accepted from a single institution

Sharing of good practices

We find that the use of interventions varies across UKRI councils, and there might be parts of the organisation that have more to share. Some examples of the potential for wider sharing across the organisation are EPSRC's use of demand management, the recent introduction of discussion boards at BBSRC, and lessons from the first use of randomisation at NERC. Based on our consultation, we conclude that these interventions (with some exceptions of demand management) are not used yet by other councils. Our survey reveals that UKRI staff would most

like to see wider use of interventions around working with underrepresented groups. The survey demonstrates various interventions in this area across councils; many are minor tweaks.

Our consultation reveals that for some interventions, wider knowledge sharing is a must to ensure a well-functioning review process. For example, councils have different demand management approaches, and no UKRI-wide system exists. Our consultations indicate that currently there may be a risk of applicants playing the system by re-submitting applications to other councils. Thus, an organisation-wide approach or oversight of demand management might be necessary.

Furthermore, increasing calls for cross-council programmes are made, and closer coordination across the councils is generally called for by individuals who submitted views to our study. The operation of cross-council programmes over the past years since the establishment of UKRI has revealed some lessons for the assessment processes in these programmes. For example, our consultation with UKRI staff reveals that programmes that fund cross-council areas and use panel members from different councils with different experiences and previous guidelines can be problematic if the differences are properly accounted for. The panel members might rely on their previous experiences and not the procedures of the cross-council panel. The applications from certain disciplines can be disadvantaged because certain panel members treat the panel differently than others. Therefore staff running cross-council investments should understand the differences between the councils and their review processes to provide a proper briefing to the panel members at cross-council panels and mitigate any problems arising because of different previous experiences of panel members.

The above demonstrates a need for sharing good practices across the organisation. Several survey respondents also expressed a need for this and a willingness to engage. In practical terms, this can take various forms, and the outputs of this study can serve as a starting point to organising a further collection of organisational intelligence and exchange of experience. One option might be to focus on specific interventions, such as those rated as most relevant in the survey. UKRI (and other funders) would benefit from having an organised list of tested interventions or a toolkit with options indicating when the specific options could be appropriate (for what funding objectives). That is essentially the end product of this study. The tool would be most valuable if regularly updated as the study demonstrates that while UKRI might not have yet implemented very radical interventions, significant effort is regularly invested in improving the assessment process.

Appendix E   Conceptual framework

Figure 7   Conceptual framework part 1: from change drivers to interventions

Drivers for change	Intervention type	List of interventions
Proactive • Addressing societal needs • Encouraging wide participation • Wide range and combinations of disciplines • Disruptive and transformational as well as routine research • Occasionally: react at speed to emergencies	**Pre-Call**	Involving researchers and civil society in call specification design
		Use of quotas
		Specific eligibility requirements
		[etc…]
	Application design & parameters	Adding/removing specific sections to application form
		Short pre-application / letter of intent / expression of interest
		Expand/contract application time window
		[etc…]
Reactive • Peer review burden • Risk of bias/cronyism • Problems for MIDRI/High-risk research • Arbitrary outcomes • Difficulty to consider non-academic criteria	**Process design**	Interviews
		Double-blind reviewing
		Pitching ('Dragon's Den') style events
		[etc…]
	Decision-making	Lottery
		'Wildcard'
		[etc…]
	Training & feedback	Applicant rebuttal
		Unconscious bias training
		[etc…]

At its core, our conceptual framework has a tabular approach. For each intervention, we will synthesise evidence from our three strands of data collection – literature review, interviews and survey – and provide information for each in the following categories:

- Definition(s): what exactly does the intervention involve? Are there relevant differences in how different funders practise the intervention?

- Why to do it: what is the envisaged benefit of the intervention? What problems/issues is it supposed to solve? What, therefore, might be measures of its success?

- Why not to do it: does the intervention have any weaknesses or drawbacks? Are these especially problematic under certain circumstances (i.e. for particular scheme types)?

- Evidence verdict and strength of evidence: is there evidence to show that this intervention has (or has not) worked? What is the strength of the evidence (e.g. controlled experiments, light-touch evaluation, anecdotal)? Besides a written verdict, we will add a ranking of evidence strength on a scale to provide an at-a-glance view on which interventions have been well explored by funders and academics, and which ones are still at experimental stage (meaning future schemes looking to use them ought to consider a pilot/trial first)

- Schemes and sources: list of sources used for each intervention for reference

Populating the table below is the core task of this study. It will form the basis of our reporting, and also for the infographic to be developed at the end of this study. We note that the format of the table below is for illustrative purposes only. The quantity of information yielded will, at

least at initial analysis stages, far outstrip the capacity of the format shown below. We do however aim to also arrive at a simplified summary version that can be presented in such tabular form, to act as a basis for an infographic.

Figure 8   Conceptual framework part 2: evidence matrix

Intervention type	Possible interventions	Definition(s)	Why do this? (desired objectives/ outcomes)	Why not do this? (potential hazards)	Evidence verdict (has it been shown to work/not work? Strength of evidence?)	Schemes and sources
Pre-Call						
	[etc…]					
Application design & parameters						
	[etc…]					
Process design						
	[etc…]					
Decision-making						
	[etc…]					
Training & feedback						

This core evidence table will be populated with synthesised information, i.e. combinations and summaries of evidence from multiple sources and (usually) from multiple existing funding schemes. As an intermediate step, each individual piece of evidence will be assessed. Here we have two fundamental approaches, depending on the specific piece of evidence in question.

- **Evidence by intervention:** this will likely be the less common but more straightforward approach. Some items in our literature review will focus specifically on one intervention (e.g. an academic study involving controlled experiments to analyse that specific intervention). A small number of interviewees may also be experts on one specific intervention type and be able to make robust claims on its pros and cons. These cases can unproblematically be assigned to the relevant intervention in our evidence table

- **Evidence by programme:** Often, evidence sources will not treat a specific intervention across many programmes. Instead, there will be evaluations of specific programmes that used an intervention, or interviewees/survey respondents who designed or supervised such programmes. Evidence in these cases will not always be sufficient to fully determine the effect of one specific intervention in the assessment process. This is especially the case for programmes where multiple interventions have been implemented. For example, a programme may involve anonymised reviewing, a two-stage application submission and unconscious bias training for reviewers. Unless an evaluation specifically looked at each of these elements individually, it will not be possible to fully attribute any observed outcomes to one specific intervention. There are gradations here of course and we will consider the strength of each piece of evidence on its own merit. Nevertheless, it is important to acknowledge that in many pieces of evidence, the effect of one specific intervention may

not always be possible. Looking across a range of evidence pieces from several programmes using the same intervention, strength of evidence will of course increase

Figure 9   Data collection framework part 1: Evidence by intervention
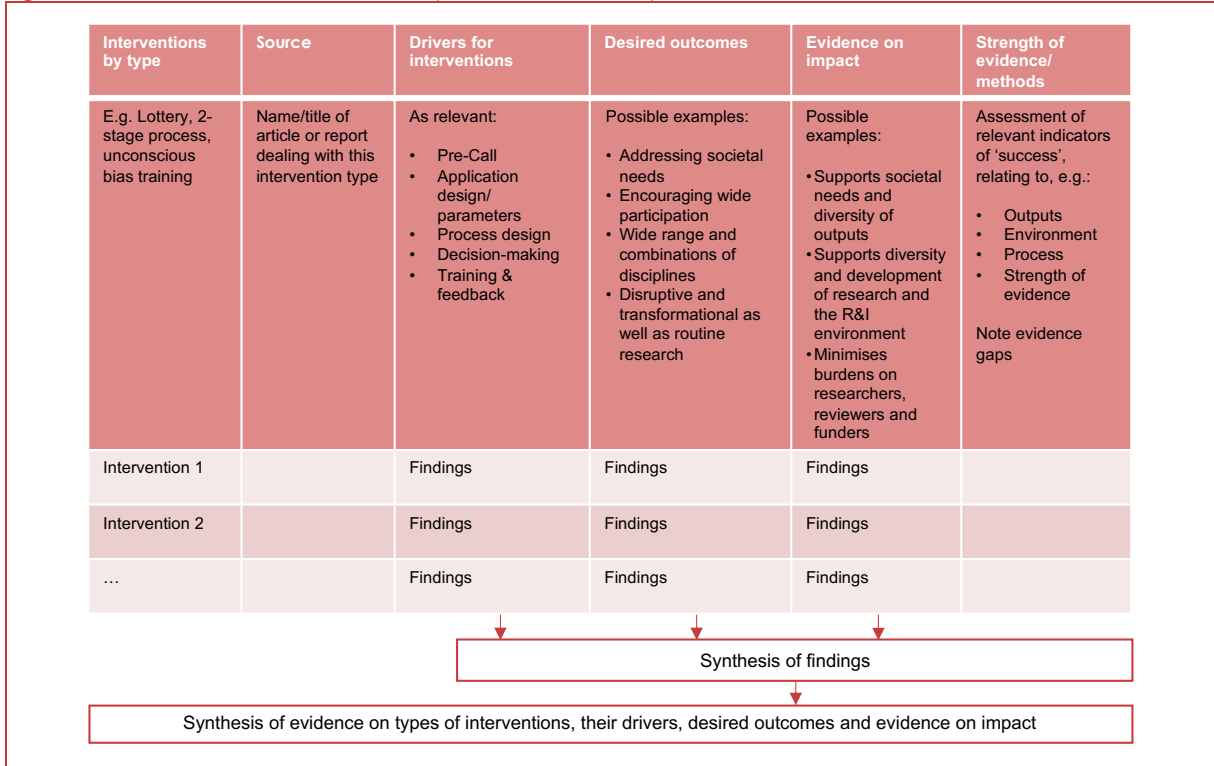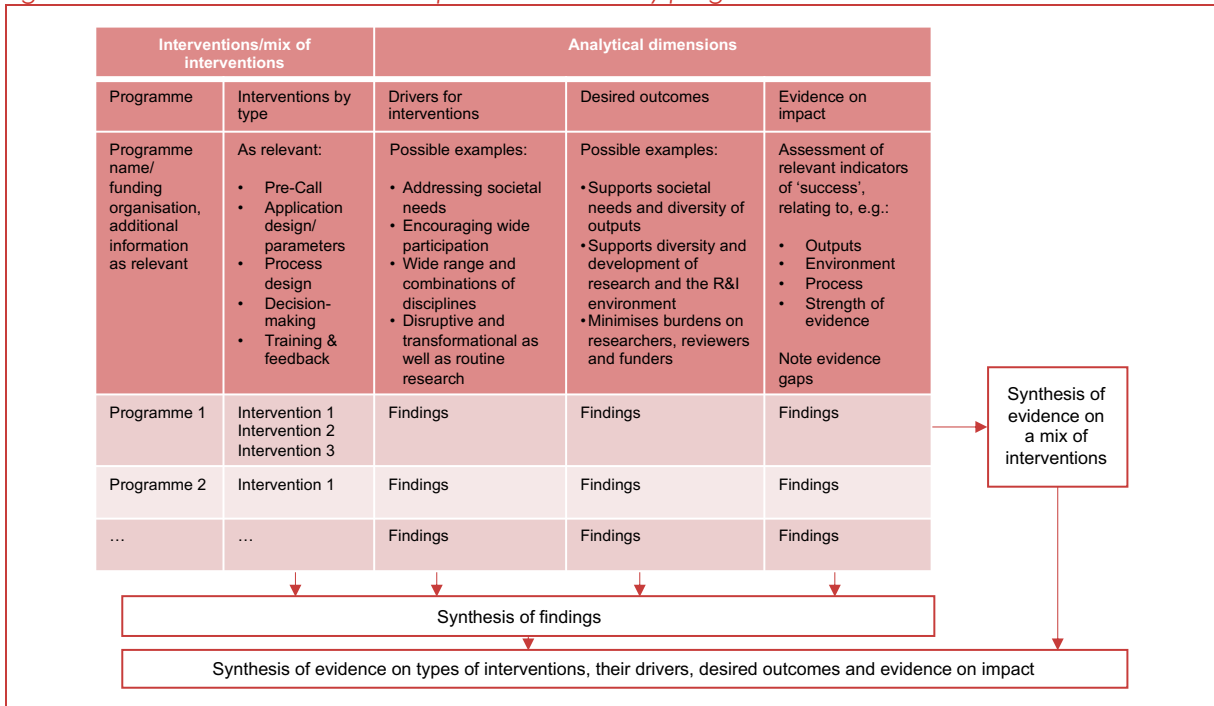
Interventions by type	Source	Drivers for interventions	Desired outcomes	Evidence on impact	Strength of evidence/ methods
E.g. Lottery, 2-stage process, unconscious bias training	Name/title of article or report dealing with this intervention type	As relevant: • Pre-Call • Application design/ parameters • Process design • Decision-making • Training & feedback	Possible examples: • Addressing societal needs • Encouraging wide participation • Wide range and combinations of disciplines • Disruptive and transformational as well as routine research	Possible examples: • Supports societal needs and diversity of outputs • Supports diversity and development of research and the R&I environment • Minimises burdens on researchers, reviewers and funders	Assessment of relevant indicators of 'success', relating to, e.g.: • Outputs • Environment • Process • Strength of evidence Note evidence gaps
Intervention 1		Findings	Findings	Findings	
Intervention 2		Findings	Findings	Findings	
…		Findings	Findings	Findings	

Synthesis of findings

Synthesis of evidence on types of interventions, their drivers, desired outcomes and evidence on impact

Figure 10 Data collection framework part 2: Evidence by programme

Interventions/mix of interventions		Analytical dimensions		
Programme	Interventions by type	Drivers for interventions	Desired outcomes	Evidence on impact
Programme name/ funding organisation, additional information as relevant	As relevant: • Pre-Call • Application design/ parameters • Process design • Decision-making • Training & feedback	Possible examples: • Addressing societal needs • Encouraging wide participation • Wide range and combinations of disciplines • Disruptive and transformational as well as routine research	Possible examples: • Supports societal needs and diversity of outputs • Supports diversity and development of research and the R&I environment • Minimises burdens on researchers, reviewers and funders	Assessment of relevant indicators of 'success', relating to, e.g.: • Outputs • Environment • Process • Strength of evidence Note evidence gaps
Programme 1	Intervention 1 Intervention 2 Intervention 3	Findings	Findings	Findings
Programme 2	Intervention 1	Findings	Findings	Findings
…	…	Findings	Findings	Findings

Synthesis of evidence on a mix of interventions

Synthesis of findings

Synthesis of evidence on types of interventions, their drivers, desired outcomes and evidence on impact

technopolis |group

www.technopolis-group.com