

Research Culture

Strategic Plan
2023-2028

A partially randomised approach to internal funding allocation

A pilot by the University of Leeds' Research Culture team

Written by

Prof Cat Davies, Dean for Research Culture

Holly Ingram FHEA, Research Culture Project Manager

November 2023

<https://researchculture.leeds.ac.uk>



UNIVERSITY OF LEEDS

Summary

This report presents the rationale, methods, and preliminary data emerging from a 2023 trial of partially randomised funding allocation (PRA) used to internally distribute Research England’s Enhancing Research Culture fund at the University of Leeds. Of 26 eligible applications, six fell in the upper midfield on quality assessment and were randomised. Of this subset, one received the available funding.

Qualitative data from applicants, reviewers, and moderators in the trial suggest modest gains regarding the **reduction of bias** and **efficiency of peer review**. The benefits of the **feedback** that PRA affords are variable.

This report presents strategic and operational recommendations for colleagues interested in adopting or adapting PRA as a way of improving research culture. For example, we highlight the importance of a) sharing data on the role of chance in traditional peer review, and b) considering the benefits and risks of PRA at both group and individual applicant levels.

This project is part of University of Leeds’ [Research Culture Strategy](#) 2023-2028.

Contents

Summary	2
Acknowledgements	2
1. Introduction	3
Precedents for partially randomised allocation	3
Our rationale for change	4
2. Process	5
Launch	5
Reviewing.....	5
Randomisation	5
Outcome.....	5
Diversity monitoring.....	6
Notification	6
3. Stakeholder reception and feedback	6
4. Discussion	12
5. Conclusions	13
6. Recommendations	14
References.....	15
Further reading	15
Appendix 1: Comparative quantitative data on call process.....	16
Appendix 2: Questions used to elicit feedback from stakeholders on PRA	17
Appendix 3: Diversity monitoring data	18

Acknowledgements

The authors thank all the applicants, reviewers, and moderators involved in this pilot, who provided generous feedback. We acknowledge the wider research culture team for supporting the production of this report, as well as Helen Pickard from the University of Leeds who provided additional peer feedback on an earlier draft. Thanks also go to the research community at the UoL who are developing unique areas of research culture practice to make meaningful, lasting change. We are grateful to Research England for funding this work.

1. Introduction

In the move towards an improved research culture, several funding bodies have started to reconsider traditional approaches to peer review to minimise bias and ensure greater inclusion of disadvantaged groups, to improve review quality, to enable more transformative research, and to save time/reduce burden. UKRI's recent review of peer review (Kolarz et al., 2023) evaluates 38 intervention-types designed to optimise peer review processes. One of these approaches is **partially randomised allocation (PRA)** of funding, aiming 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.

The review justifies PRA thus:

[...] 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.

Kolarz et al. (2023: 44).

The use of PRA is relatively new and localised. Its principles and implications are not without controversy, as summarised in several recent thought-pieces (e.g. Golberg, 2022; Harford, 2023; Nature Editorial, 2022), as well as an emerging body of evaluative literature by researchers and metaresearchers (Kolarz et al., 2023; Woods and Wilsdon, 2021a; 2021b).

As part of the research culture programme at the University of Leeds, in 2023-24 we adopted PRA for our internal [open call for research culture projects](#) funded by Research England's Enhancing Research Culture fund. This meant that we effectively used a partial lottery system to make funding decisions for proposals considered equal on core quality criteria. Similar to other trials of randomisation in research funding, we opted for its partial form, by which the middle tranche of applications rated on core quality criteria go through to random selection for funding.

Due to the relatively small number of applications to our call, we were in a position to solicit detailed feedback on the experience. To add to the growing body of investigations into PRA, we have now compiled our processes and emerging data in this report to:

- share the outcomes of our trial,
- evaluate the impacts of the PRA pilot, and
- make recommendations for future adoption and adaptations.

Contributory data includes indicative quantitative measures from the PRA round relative to the traditional peer review process used in our previous open call, as well as stakeholders' reflections on their experience of PRA.

Precedents for partially randomised allocation

Kolarz et al.'s 2023 review identifies PRA as being used by at least six research funding bodies. Among the first funders trialling in the UK are [NERC](#) (*Exploring the Frontiers* and *Pushing the Frontiers* schemes) and the [British Academy](#) (BA/Leverhulme Small Research Grants). NERC's stated aims were to increase the diversity of award recipients and streamline the application process, while the BA's rationale was to improve research culture through a transparent and simplified system, remove human bias and partiality to achieve equity, enable limited feedback to applicants, and ease the burden on applicants and research officers without impacting the quality of applications and assessment. In summer 2023, Wellcome used PRA to allocate funding to applicants to its

Institutional Fund for Research Culture ([IFRC](#)), to work towards removing biases and support fairness, “recognising that there is no “right” answer for research culture but plenty of solutions” (Lewis-Wilson & Towers, 2023).

PRA is also used in charity funding, e.g. [Nesta](#)’s Explorations Initiatives. In their justification, Nesta describe funding as “a nightmare of unconscious biases, popularity contests, conservative decision-making and trying to slide a cigarette paper between two applications that are equally as good [...]”. Nesta states similar aims to the funders referenced above, i.e. to save time, reduce bias, improve diversity, and give space for more unconventional ideas.

Although PRA is still in its infancy, early impacts are emerging through funders’ assessments. Kolarz et al. (2023: 44) report that at least two funders were found to have diversified their awardee pool. Applications to three more were found to increase in response to the introduction of PRA, reportedly due to a perceived higher chance of success among applicants. In Woods and Wilsdon’s (2021a) small-scale qualitative study with six research funding bodies, the key driver for PRA was found to be fairness – both of decision-making and to applicants regardless of background or field. The study also revealed challenges regarding the communication of the system to stakeholders, with concomitant reputational risk.

The British Academy has recently published promising [interim findings](#) of its three-year trial of PRA. From the two initial rounds (of six), successful candidates have come from a wider range of institutions, many in receipt of their first BA Small Grant, and more from Scotland and Northern Ireland than previously. There has been a notable increase in successful BAME applicants (27% vs. 18% under the previous system). Notably, the BA have seen a 32% increase in applications to the scheme, up from the preceding year when the traditional application process was used.

Our rationale for change

One of our internal schemes funds projects to improve research culture at the University of Leeds. The scope of the fund is broad and elicits an extremely wide range of research questions, methodologies, activities, and team structures. This heterogeneity makes ranking more challenging than calls with a narrower focus. Second, having observed the external trials in partially randomised allocation, we wanted to investigate the anticipated benefits for research culture more locally. These were:

1. **Reduction of bias.** Among strong applications that are deemed equally fundable, randomisation should reduce conscious or unconscious bias against people, thus addressing inequities that may be experienced by e.g. early-career researchers or those from under-represented groups. Randomisation should also reduce bias against particular research ideas, e.g. towards safer options at the expense of more radical proposals.
2. **Efficiency.** Partial randomisation should ease the burden on reviewers as they need only to provide a simple but rigorous threshold judgment.
3. **Feedback.** Partial randomisation should allow us to provide brief feedback to applicants who do not pass the quality threshold, and to those who have passed the threshold but lost out during the randomisation process.

PRA was thought to be well-suited to our 2023-24 call based on the previous year’s response of high quality, small-scale applications for exploratory studies typical of this scheme, together with the relatively small and time-poor review panel, and a tight reviewing timeframe.

The previous year’s call used a more traditional reviewing process, requiring a panel of 20 reviewers to allocate 0-2 points against each of the 10 quality criteria, offer brief qualitative comments, then attend a full panel review session to come to final decisions.

2. Process

Launch

We provided the three-part rationale for adopting PRA within the call guidance distributed to potential applicants, as above.

Reviewing

Fourteen peer reviewers were invited from the University's research culture governance groups and two cohorts of the University's 100 Black Women Professors NOW programme. The reviewers were randomly paired and asked to reach agreement on approximately four applications. Each pair was allocated a moderator from the research culture team to adjudicate disagreements (this support was only called on from one of the seven reviewer teams). Successful applications were required to pass an initial quality threshold: applications were deemed fundable if they satisfied the first two criteria plus a minimum of two others. This was determined via a traditional but light-touch peer review process involving six binary criteria:

1. Does the proposal persuasively articulate the research culture problem or challenge that it aims to address? (required)
2. Are its aims clear and achievable within the given timescale? (required)
3. Is the methodology appropriate?
4. Are the likely impacts of the project identified, and are the outcomes measurable?
5. Are the roles and responsibilities of all team members and any partners clearly defined?
6. Are the costs requested appropriate?

The threshold and criteria provided were designed to eliminate proposals which were out of scope, unclear, unfeasible, without impact, and/or poor value for money.

Randomisation

Applications that received a 'yes' response from each reviewer pair on all six criteria clearly passed the quality threshold and were entered into the random allocation process. Those that came at the top of the randomised list down to the total funding limit would be recommended for funding.

If there were remaining funds after this first stage of randomised allocation, proposals that received one 'no' response from the reviewer team on criteria 3-6 would be randomised and those at the top of the list would be offered funding.

This process would be repeated for proposals that received two 'no' responses. Proposals receiving a 'no' on more than two criteria, or on criteria 1 and 2 at any stage failed the quality threshold.

NB. A subset of six applications were not randomised. These concerned projects that had been awarded pilot funding in the previous year's call, and which had then been submitted as follow-up applications. Five of these met the initial quality criteria and were funded without randomisation.

A simple R script was written to generate a random list of numbers, which was then used to sort the applications. This generated a priority list for funding.

Outcome

After the removal of the five successful follow-up applications, there were 26 remaining applications. Fourteen of these did not pass the initial quality threshold (in addition to the unsuccessful follow-up application).

Of the remaining 12, six passed all six quality criteria and were funded. They were not randomised since the funding cut-off was below the total amount applied for.

Six additional applications passed the quality threshold, passing five out of the six criteria. These were randomised. The application at the top of the randomised list was funded. The funding allocation limit was then reached. Therefore, five applications were unsuccessful due to randomisation. An explanation and demonstration of the randomisation process was recorded and is available [here](#). A total of £382k was allocated to 12 funded teams. One of these was awarded using PRA. A detailed breakdown of comparative data of both open calls for this scheme can be found in [Appendix 1](#).

Diversity monitoring

Diversity information was requested and provided for each team member on the application form, alongside the option to withhold this information ('prefer not to say'). After randomisation and funding recommendations, the research culture team manually collated and reviewed the gender, race, disability, and career stage diversity among project teams to verify that the PRA process had not generated any preferential biases among successful vs. unsuccessful applications (see [Appendix 3](#)). The research culture team presented the outcome of diversity monitoring to the reviewer panel. If there had been concerns regarding bias, we would have re-run the randomisation.

Notification

Applicants were notified of the review outcomes. Unsuccessful applicants received feedback indicating whether their proposals had been deemed fundable but not selected for funding (i.e. via randomisation) or simply not selected for funding. This was accompanied by written reviewer comments where provided.

3. Stakeholder reception and feedback

As [Appendix 1](#) shows, when comparing the 2023 (PRA) and 2022 (traditional) calls, a similar number of applications (31 vs. 35, respectively) and successful applications (12 vs. 13) were generated, though both the total amount requested and the mean cost per application was around 20% more in 2023, reflecting the higher value of the overall fund. On the reviewer side, the 2023 call required fewer reviewers and yielded significantly fewer incomplete reviews (0 vs. 15).

We elicited feedback on the PRA process from applicants, reviewers, and moderators (see table 1) using a Microsoft Form questionnaire. This was completed by reviewers and moderators during the reviewing panel session. Applicants received the questionnaire via email approximately a fortnight after the release of reviewing outcomes. See [Appendix 2](#) for the full questionnaires.

Number and type of stakeholders polled	Number of responses	Breakdown of respondent role type	Breakdown of respondent career stage
31 applicants	12	3 professional services 8 academic 1 did not specify	1 early career 4 mid-career 4 senior members of staff 3 did not specify
14 reviewers	7	3 professional services 2 academic 1 technical 1 did not specify	2 senior members of staff 2 mid-career 3 did not specify
5 moderators	6 (duplicate response by 1 moderator)	4 professional services 2 academic	3 senior members of staff 3 mid-career

Table 1. Frequency summary of responses from each stakeholder group, and their role type and career stage.

Below is a summary of qualitative responses by question and theme, illustrated by data provided.

Did the PRA approach influence your decision to apply? If so, in what way?

The majority of applicants reported no influence. Of those who did report an influence, this was negative, citing that the explicit chance factor was a deterrent.

- No x 7
- Yes x 3
- No response x 2
- *Yes, I almost did not apply as I was unsure if PRA was the right approach.*
- *It made me slightly more nervous to apply and also a little frustrated that, technically, a proposal that would have scored less than ours using a standard scoring criteria could have been funded, and ours not.*
- *On the whole, I found it a bit off-putting because it seemed that part of the process would be governed by chance rather than merit. I also didn't think the information on process provided beforehand was sufficient for me to understand what sort of PRA method was employed. We live in the age of AI - what model of randomisation was used and how was this model trained? How biased might the model be?*

Did the PRA approach affect the way you reviewed the applications? If so, in what way?

Although the response rate was low, some reviewers said that PRA gave them a heightened sense of the importance of their decisions. They also found the binary rating unsatisfactory. Other reviewers reported no influence of PRA on their behaviour.

- Increased awareness of power as a reviewer:

Ultimately it did not change anything but I was aware that the more applications that I scored as able to proceed to randomisation, the less chance each application had of success.

It did make me think about the role I played as a reviewer and how my scores would affect chances of success. I think it impacted my scoring - when I was not 100% sure about my score, as I struggled to give either a 100% yes or no answer, I found myself hoping that some bids would still get a chance to be selected through the 2nd stage random selection.

- Indirectly: the use of binary judgments was felt by some to lack nuance. This was mitigated by many reviewers by adding voluntary qualitative comments to their binary judgments (which they intended to be developmental for applicants).

It forced collective binary decisions which might have influenced the outcomes - if the reviewers had slightly different views there was a tendency to discuss and resolve to a single view.

No - but in combination with a binary approach I think it did, so I found myself wanting a third 'partially' option where it wasn't clearly a Y or a N as I was more aware of the consequences

I find the binary scoring unhelpful and unsatisfying. It was the same for my co-reviewer. I would have felt more comfortable giving scores from 1-5; as in very rare cases was it as clear cut as a yes or no. That made it really hard to review as I feel the responsibility to give deserving applications a chance for success. I felt under more pressure to get the scoring right, but the only way to showcase my thinking was providing comments. I also think the comments are helpful for the recipients - both those successful and unsuccessful.

I found the binary score challenging in some cases and perhaps a 3 (or 4?) point score would have allowed more nuance.

No x 2

I still reviewed the applications as I would have - no change in my behaviour as still entrenched in the traditional approach

Are you convinced by our rationale for trialling PRA, e.g. reduction of human bias, efficiency, provision of feedback, innovation of processes?

The data suggest that stakeholder role was a factor in how convinced stakeholders were of the rationale for PRA. Applicants were the most sceptical (while also representing the full range of scepticism). Reviewers and moderators were more convinced, with the caveat that moderators may have been positively biased: they were all from the research culture team, who tended to champion PRA.

Applicants:

- Not at all convinced x 4 (2 successful applications; 2 eliminated on quality)
- Somewhat convinced x 3 (2 successful; 1 eliminated on quality)
- Mostly convinced x 2 (1 success; 1 n/a)
- Totally convinced x 2 (1 success; 1 eliminated on quality)
- Other x 1, commenting: *I don't think the rationale was explained as such - just presented as 'this is how we're doing it'*

Reviewers:

- Somewhat convinced x 1
- Mostly convinced x 2
- Totally convinced x 3
- Other x 1, commenting: *I can see the argument, but I am not convinced it worked or resulted in a fairer selection: the proposals I/we chose to go forward into the random selection pool, the ones I thought had greatest merit did not get selected, whilst those I felt were good enough but not as strong as my/our top choices got selected. This feels really dissatisfying and wrong.*

Moderators:

- Somewhat convinced x 1
- Mostly convinced x 2
- Totally convinced x 3

Do you have any concerns about the move to PRA?

This question elicited many comments. Some applicants reported being uncomfortable with the perceived element of chance, citing unfairness that careful work should be reduced to chance. Others felt that PRA was an overly blunt instrument, unable to merit certain ideas or applicants with particular profiles. The partial nature of randomisation raised some concerns with some reviewers feeling uncomfortable that some applications (previous pilots and the highest scoring) did not undergo randomisation. Two applicants were concerned that the binary scoring system may threaten quality (i.e. the bar for a 'yes' decision may be lower than a continuous scoring system). Another applicant felt that the wording feedback did not clarify the reasons for rejection. Eight respondents reported no concerns.

- Dumb luck / lack of fairness:

The reason why I put 'not at all convinced' is because of the pot luck of this funding allocation after the threshold has been met. I do think it is fantastic to trial innovative ways but when I saw this as a

method on the guidelines it did feel me with anxiety. I am not sure whether it is fair to reduce the amount of work and consideration that is taken on the proposals to a lottery chance of success. Could each application that is peer reviewed be given a grade and then you have a rank order in terms of quality and the top ones could then be discussed at a panel (like the AHRC peer review college). (applicant)

I've yet to see a convincing argument for PRA in any area. However much it is claimed to reduce bias and be more 'efficient' (which is a managerial weasel word that is blighting academia), the bottom line is that it makes life simpler for reviewers (abdicating difficult decisions about funding) and considerably more unfair for applicants. Funding bodies should be able to justify their decisions, not leave them to chance. (applicant)

If I had not been successful I am sure I would be very negative about the PRA. As it turns out, I was funded, but I am still unsure PRA is appropriate. (applicant)

Funding more on luck than merit - is it fair to cite these successes in promotions, etc.? (moderator)

- Blunt instrument / inequity:

PRA may not fully account for the diversity of projects proposed, and may not be able to strategically target important areas. (applicant)

Yes, while I appreciate the gesture towards equality in the move to PRA, and take your point about the drawbacks of human bias, I believe that we're increasingly as a University moving in the direction of equity, which is the (welcome) recognition that some come with greater disadvantage, and therefore may require case-by-case assessment which a PRA process is incapable of providing. Efficiency may come at the cost of equity in PRA. For example, there may be two applications of equal merit, but one which comes with more disadvantage factors/ from a project team with protected characteristics. How would PRA account for this? In the same way that we're starting to recruit more equitably as a University, we should strive to [fund] projects more equitably as well. (applicant)

I agree that it will remove potential bias, unconscious and conscious, and this is very important. However, I think it could disadvantage earlier career researchers, who sometimes benefit from positive discrimination by reviewers, e.g. they are given the benefit of the doubt more often to account for their relative inexperience. (reviewer)

For PRA we need to understand how to pair reviewers up to assess applications: is two reviewers sufficient? Consider a 3rd for each application. (reviewer)

A post-outcome review of diversity data is key to successful implementation: without it we risk inadvertently making things worse. (moderator).

- Concerns about the partial nature of randomisation

I do wonder about moving to include the previously funded applicants and the top scorers in the randomisation: it might be fairer and more cost effective if RE funding is reduced. I appreciate that we solicited pilots last year with the option of getting more but just not totally convinced we should fund any without the randomisation. (moderator)

It is interesting that 6 of the applications were awarded 6/6 so went through without undergoing randomisation. As the quality and number of applications increases we need to consider how we avoid rewarding those better at writing an application, for example an application may score 5/6 but be a really good idea not quite written as well as another less good idea written very well scoring 6/6

- Threat to quality:

I totally get why the BA would do it for their small research grant scheme as it's so massively popular. PRA reduces the amount of time they have to spend deliberating on cases they really can't decide

between and the metrics so far prove that it has encouraged more applications from previously under-represented groups. I'm less certain about its use for relatively low volumes of applications to decide between. I guess that if I can be reassured that the quality threshold is high enough then I think it's a good idea. (applicant)

The British Academy/Leverhulme format, of scoring proposals and then them going into a ballot for funding should they meet a certain score threshold, seemed a slightly better approach than a simple yes/no approach as used for this call. To me, it seemed quite easy to meet the 'yes' criteria, but I'm not sure this binary approach would reflect the standard of the proposals which meet the 'yes' criteria. (applicant)

- Lack of transparency:

The process wasn't very transparent. My outcome/feedback email was completely unclear whether our application had been part of the lottery or whether it didn't meet the threshold for consideration. This isn't particularly helpful. Information needs to be explicit and clear. (applicant)

- No concerns:

It think it's a fantastic and very fair approach (applicant)

I think it's a really good move! I've had so many funding applications rejected by large schemes without knowing why. It's helpful to know that a rejection might not be because of my application (or chosen co-applicants) but luck of the draw. (applicant)

No, it's the way forward (moderator)

I think the benefits are strong especially as this can work to reduce bias (moderator)

+ 2 reviewers (no comments)

+ 2 moderators (no comments)

Did the PRA approach affect the typical burden involved in reviewing? If so, in what way?

Reviewers' responses to this question went against our anticipated reduction in burden, with most reviewers reporting that they spent the same amount of time and attention on the reviewers as they would using a traditional system. This was also reflected in the generous provision of optional comments in the reviews.

- No
- Not sure
- Probably not, as the panel meeting took time, but was actually really beneficial.
- No efficiencies garnered in terms of time and effort as only appropriate to review each in detail
- It did not impact - still read the applications in detail and reflected on each criteria before responding to Y/N
- I don't feel that I have enough experience of reviewing to comment.
- Not sure as I can't compare

Do you have any suggestions to help us improve our system of PRA?

Most comments reflected the need for greater transparency of feedback, a return to full peer review panels on equity grounds, a more nuanced scoring range, or a right to reply.

Be clear about whether an application failed at the quality threshold or at randomisation (applicant)

I'm not yet convinced that PRA is a sound replacement for human processes, and would be in favour of either a return to full peer review with unconscious bias observers or the introduction of some means of equity monitoring at the PRA stage of the process. This would of course confound the 'random' part of the process, but would be more equitable. (applicant)

Ditch it. But if you are going to make life easier for reviewers and things more 'efficient', then the least that could happen is that the feedback could be more constructive and clear on why the reviewers have made their recommendations. At present, the system seems to work against applicants in both the transparency of selection (if recommended for funding; obviously not an issue for my team!) but also in terms of the quality of feedback. (applicant)

If there are projects that are borderline above/below the threshold perhaps a peer group panel makes the final decision. (applicant)

Have an opportunity to respond to reviewer comments before deciding which proposals meet the quality threshold. (applicant)

I would apply a different, non-binary scoring system; I would also not rely on two criteria only to decide whether the bids make it through to the next round but all of the criteria; I would also introduce a criterion on return in investment as I think this is missing. (reviewer)

Thought needs to be given to the reapplication/resubmission policy for those proposals that passed a given quality threshold but have missed out on funding due to randomisation - are they permitted to resubmit? Where is the cut off? (reviewer)

Need to have some contingency plan on if more applications pass the quality review than you have funding. In this case would we revert to full randomisation? (moderator)

4. Discussion

The anticipated benefits of PRA stated in our rationale were reduction of bias; efficiency; and feedback. Here we reflect on the extent to which our trial achieved these.

Reduction of bias

As randomisation was only used for applications passing five of the six quality criteria, this reduction could only be partial. That is, reviewer bias for certain topics could have influenced initial binary decisions, leaving all those with six 'yes' responses with a straightforward route to funding. Bias should have been reduced for applications scoring at the next level down, though this only concerned six of the 25 applications. Thus, any reduction of bias was modest.

A system in which the most highly rated applications receive funding (i.e. **partial** randomisation) is inevitably subject to a degree of reviewer bias towards certain ideas. On balance, this is preferable to a fully randomised or lottery system which does not use a quality threshold.

Because qualitative feedback was not based on detailed knowledge of outcomes, it could not conclusively address the question of bias. However, some comments revealed concerns about the removal of **positive bias** towards disadvantaged groups, e.g. ECRs or researchers of colour.

Diversity monitoring indicated that randomisation had not generated any notable bias according to gender, ethnicity, disability or job type, and demonstrated an even distribution of unsuccessful and successful applicants across the areas listed (see [Appendix 3](#)). Due to other aspects of the call management differing from the traditional approach in the previous year, it is not possible to compare diversity outcomes.

Efficiency

Surprisingly, most reviewers reported that they spent the same amount of time and attention on their reviews as they would using a traditional system. This may have been compounded by the internal nature of the funds: institutional colleagues were keen to provide developmental feedback. If this finding is replicated more widely, it presents implications for the case for PRA. It also nullifies some of the scepticism about the rationale for PRA, i.e. it being solely to reduce reviewer burden.

Provision of feedback

Although we provided feedback on reasons for rejection (randomisation / quality), some applicants found this confusing. Wider discussion with colleagues also suggested that being rejected on the grounds of randomisation is both a) easier to accept and b) frustrating in that no improvements can be made when resubmitting.

Aside from an evaluation based our original rationale for trialling PRA, we must also consider a dominant theme in our data: the perceived element of **chance** and associated feelings of unfairness. Although randomisation seems to bring this theme to the fore, it is important to highlight the arbitrariness of traditional peer review process, where evidence from higher-volume calls demonstrates that the ranking of close-scoring applications is scarcely better than chance (Fang et al., 2016; Fang & Casadevall, 2016; Jerrim & de Vries, 2020): "It seems like a formalisation of what already happens between the strongest applications" (Golberg, 2022). It will be important to make the somewhat 'random' or chance nature of traditional peer review clear to stakeholders when considering potential benefits and risks of PRA (as well as the challenge of ranking heterogenous

proposals). This quote from Stafford et al. (2023) summarises the gains that PRA can bring to fairness of the process:

A recent study (Woods & Wilsdon, 2021b) found that the strongest motivator for funding institutions to use partial randomisation is fairness: a fairer decision-making process when peer review had reached its limits; fairer to applicants, as it is blind to institution, geographical location, race, gender, discipline and methodology; and also a transparent process and therefore easier to communicate and understand funding decisions. Other organisational motivators are the desire to break deadlocks in, or reduce time spent on panel decision making, and to ameliorate risk aversion or other concentrations of awards so as to facilitate the funding of a greater plurality of research topics and methodological types.

The chance factor also throws up concerns about reputational impact from both rejections and successes in PRA. Some applicants may worry about crediting their awards (if successful) to randomisation, or being judged negatively if they lost out due to randomisation. Where traditional academic CVs are still required, a brief line to clarify when randomisation was used in a grant scheme may help to mitigate this concern.

The qualitative data suggests that trialling PRA had no effect on the volume of applications submitted, though it was correlated with a smoother reviewing process (note that the 2023 round used pair-reviewing which may have been a stronger causal factor in reviewer compliance).

5. Conclusions

Based on feedback to our trial of PRA, only modest gains were made regarding the **reduction of bias** and **efficiency**. In addition, the loss of positive bias towards marginalised groups is a drawback, though mitigations could be introduced to the process to address this. The benefits of the **feedback** that PRA affords are subject to individual differences, so are inconclusive.

Our trial was based on immediate experiences of the **allocation process** rather than the longer-term impact of PRA on the nature and outcomes of the funded research. Downstream effects on reduction of bias, gains in efficiency, or trust in funding allocation should be monitored.

It is important to consider findings of PRA evaluations on two levels. At the group level, positive effects are emerging, e.g. The British Academy's diversification of successful applications. However, at the individual level, the process can be frustrating, e.g. perceptions of powerlessness by applicants.

This case study includes several limitations. First, it analyses effects only on stakeholders rather than on the wider research or funding system. These effects are also very early-stage. Second, the case study does not purport to be a controlled trial comparing PRA to the traditional peer review process used in our previous call: the two calls differed in several other aspects, e.g. reviewing criteria, concurrent industrial action in 2023, and a different applicant cohort. Third, relating to the stated aim to reduce bias, some reviewers may have been able to identify applicants despite anonymisation, due to belonging to the same University community (NB. Any conflicts of interest on the part of reviewers were expressed before the review process).

6. Recommendations

We will carefully consider the value of PRA in future funding schemes based on the data and findings above. We note several recommendations for ourselves and our partners interested in trialling PRA.

Strategic

1. Consider exemption from randomisation for marginalised groups to restore equity.
2. Reconsider exempting follow-on proposals from randomisation due to questions of fairness.
3. In future rounds of PRA for research culture projects, communicate with stakeholders the evidence demonstrating the influence of chance in traditional peer review, as well as the specific challenges of assessing heterogeneous projects.
4. Consider benefits and risks of PRA at both group and individual applicant levels.
5. Solicit longer-range feedback on the outcomes of PRA from successful applicant teams.
6. Re-run the trial in the subsequent round of this scheme to investigate whether the marginal gains this year are replicated.
7. Share the findings of this case study with colleagues within and beyond the institution to facilitate discussion about the merits of PRA for different funding schemes, and to encourage consideration of alternative approaches to peer review.

Operational

1. Include in outcome notifications a clear explanation of whether randomisation was used in the allocation process for a particular application.
2. Reconsider binary criteria, which some reviewers found more challenging to use than graded scoring (while for others it reduced the work required greatly). The granularity of the scale should be considered alongside the volume of applications: a finer-graded scale may be warranted in higher-volume calls to sufficiently discriminate quality.
3. Ensure consistency in the way that diversity monitoring data is provided by teams.

References

- 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>
- Golberg, A. (2022, November 15). The (partial) rise of (partial) randomisation. *Research Professional News*. <https://www.researchprofessionalnews.com/rr-news-uk-views-of-the-uk-2022-11-the-partial-rise-of-partial-randomisation/>
- Harford, T. (2023, September 29). Sometimes, a random solution is best. *Financial Times*. <https://www.ft.com/content/61f67834-c228-44bb-b1e9-9a60d6b57d27>
- 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>
- Kolarz, P. et al., (2023, July 3). *Review of Peer Review*, UKRI. <https://policycommons.net/artifacts/4385753/ukri-050723-reviewofpeerreview/5182293/>
- Lewis-Wilson, S. & Towers, S. (2023) The luck of the draw: Wellcome's Institutional Fund for Research Culture [version 1; peer review: awaiting peer review]. *Wellcome Open Res*, 8:525, <https://doi.org/10.12688/wellcomeopenres.20057.1>
- Nature Editorial (2022, September 20). The case for lotteries as a tiebreaker of quality in research funding. *Nature*. 609(7928): 653. <https://doi.org/10.1038/d41586-022-02959-3>
- Stafford, T., Rombach, I., Hind, D., Mateen, B., Woods, H. B., Dimario, M., & Wilsdon, J. (2023). Where next for partial randomisation of research funding? The feasibility of RCTs and alternatives. *Wellcome open research*, 8, 309. <https://doi.org/10.12688/wellcomeopenres.19565.1>
- Woods, H.B. & Wilsdon, J. (2021a). Experiments with randomisation in research funding: scoping and workshop report (RoRI Working Paper No.4). *Research on Research Institute*. <https://doi.org/10.6084/m9.figshare.16553067.v1>
- Woods, H.B. & Wilsdon, J. (2021b). Why draw lots? Funder motivations for using partial randomisation to allocate research grants (RoRI Working Paper No.7). *Research on Research Institute*. <https://doi.org/10.6084/m9.figshare.17102495.v2>

Further reading

- Bendiscioli, S., Firpo, T., Bravo-Biosca, A., et al. (2022). *The experimental research funder's handbook* (Revised edition, June 2022, ISBN 978-1-7397102-0-0). Research on Research Institute. <https://doi.org/10.6084/m9.figshare.19459328.v3>
- Gladstone, J., Schipper, L., Hara-Msulira, T., Casci, T. (2023). Equity and Inclusivity in Research Funding: Barriers and Delivering Change. <http://dx.doi.org/10.5287/bodleian:KZjBY77pO>
- Liu, M., Choy, V., Clarke, P. et al. (2020). The acceptability of using a lottery to allocate research funding: a survey of applicants. *Res Integr Peer Rev* 5, 3. <https://doi.org/10.1186/s41073-019-0089-z>

Appendix 1: Comparative quantitative data on call process

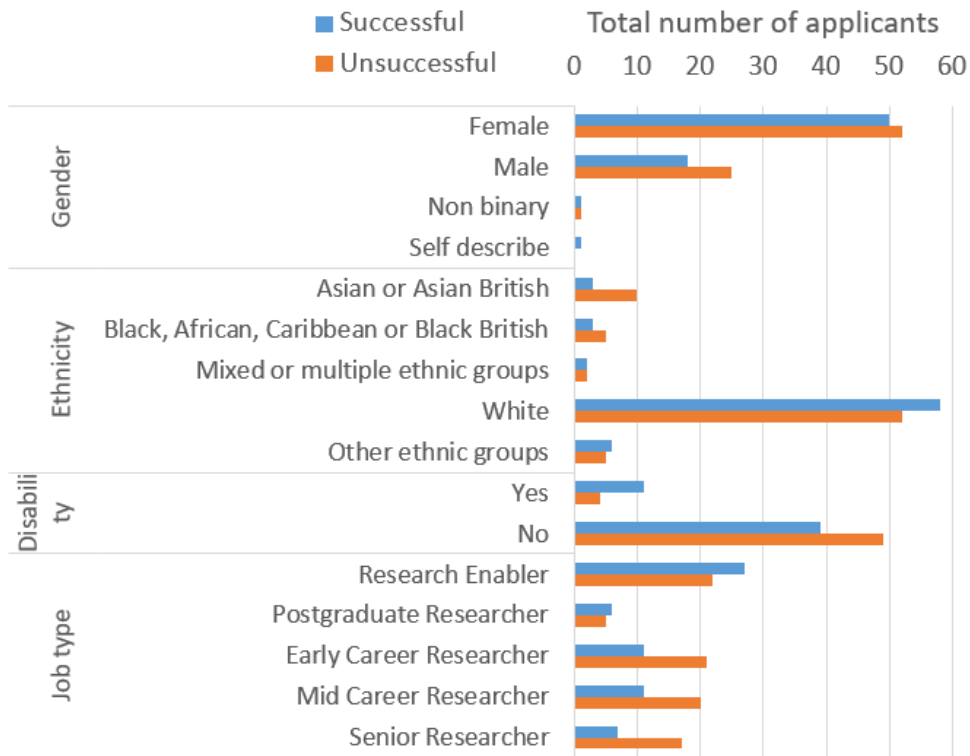
	Dec 2022	Jul 2023
Number of applications submitted	35	31 (25 new; 6 follow-up)
£ value of applications	Total: £853,713 Range: £6K - £50K Mean: £24,392 SD:	Total: £1,053,000 Range: £10K - £50K Mean: £30,086 SD:
No. applications randomised (i.e. passing quality threshold)	N/A	17 (5 were follow up)
No. applications awarded funding	13	12
Final success rate (x/y and %)	13/35 = 37%	12/31=37%
Total fund available	£280,000	£400,000
Total funding awarded	£267,783 (99% of funding pot)	£382,000 95% of funding pot
Profile of all applicants (i.e. gender, race, disability, career stage, Faculty)	Did not collect	See Appendix 3
Profile of unsuccessful applicants	Did not collect	See Appendix 3
Profile of successful applicants	Did not collect	See Appendix 3
Applicant team size	Smallest: 2 Largest: 14	Smallest: 2 Largest: approx. 35
Range / consensus of reviewer scores	Large SD	Six out of seven reviewer pairs were able to agree on all scores before joint score submission – one pair called on the moderator to come to an agreement on two applications.
Number of reviewers	<ul style="list-style-type: none"> • 20 total • 3 per app (approx. 5 apps per reviewer) 	<ul style="list-style-type: none"> • 14 total • 2 per app (approx 4 apps per reviewer)
Profile of reviewers (i.e. gender, race, disability, career stage, Faculty)	Did not collect	2 senior members of staff 2 mid-career 3 did not specify 3 professional services 2 academic 1 technical 1 did not specify
Number of reviewer pairs calling on the moderator	N/A	1
Volume (and broad type?) of queries from applicants	Approx 15 queries, about: Application form, project type, buy out, costings, deadline, eligibility, signatures,	Approx 7 queries, about: Future calls, external collaborators, project type, signatures, deadline extensions due to industrial action.

	future calls, application support.	
Volume (and broad type?) of queries from reviewers	Approx 3 queries, about: Availability and scoring process	2 queries, about: Binary judgements and HR timescales
Number of incomplete reviews	3 reviewers did not complete, which amounted to 15 incomplete reviews.	0

Appendix 2: Questions used to elicit feedback from stakeholders on PRA

Applicants	Reviewers	Moderators
Was the information on partially randomised allocation (PRA) in the call guidance adequate / clear?	Was the information on partially randomised allocation (PRA) in the call guidance and reviewer guidance adequate / clear?	Was the moderator guidance clear?
	Did the PRA approach affect the way you reviewed the applications?	Did the PRA approach affect the way you moderated the reviews?
Did the PRA approach influence your decision to apply? If so, in what way?	Did you provide comments with your binary judgments? Why / why not?	
Are you convinced by our rationale for trialling PRA, e.g. reduction of human bias, efficiency, provision of feedback, innovation of processes?	Are you convinced by our rationale for trialling PRA, e.g. reduction of human bias, efficiency, provision of feedback, innovation of processes?	Are you convinced by our rationale for trialling PRA, e.g. reduction of human bias, efficiency, provision of feedback, innovation of processes?
Do you have any concerns about the move to PRA?	Do you have any concerns about the move to PRA?	Do you have any concerns about the move to PRA?
If you applied for the ERC open call in Dec 2022, how did the 2023 experience compare?	If you reviewed for the ERC open call in Dec 2022, how did the 2023 experience compare?	
Were you successful in this year's call?		
What is your opinion on the feedback you received?	Did the PRA approach affect the burden for you as a reviewer? If so, in what way?	
	In general, did your reviewing process proceed in line with the guidance? If not, how did it differ? Do you have any suggestions on how to improve the reviewing process next time?	Was moderation support required? What did you do? Do you have any suggestions on how to improve the moderation process next time?
Do you have any suggestions to help us improve our system of PRA?	Do you have any suggestions to help us improve our system of PRA?	Do you have any suggestions to help us improve our system of PRA?

Appendix 3: Diversity monitoring data



NB. Applicants provided diversity data on their applications in varying formats. This representation is as accurate as possible from the data provided.



© University of Leeds 2023 DOI: <https://doi.org/10.48785/100/194> This work is licensed under a Creative Commons (CC BY-SA 4.0).