Steven De Peuter, PhD

Faculty of Psychology and Educational Sciences, University of Leuven, Leuven, Belgium

Address for correspondence:

Steven De Peuter

University of Leuven

Methods, Individual and Cultural Differences, Affect and Social Behavior (MICAS)

Tiensestraat 102 box 3727

3000 Leuven, Belgium

steven.depeuter@kuleuven.be

Title

When randomness is intrinsic, formalize it

Abstract

Competition for research funds has, in the recent decade, become hypercompetitive. Commonly, to determine which proposals receive funding, a system of peer review is used, which is broadly accepted, easily understood and broadly trusted among researchers. It is often considered the best system in use, but it suffers from important shortcomings and adaptations to overcome these shortcomings have small and often short-lived effects. Hence, the preference for peer review does not mean it *necessarily* outperforms all other systems. In fact, it is time for an open discussion about alternative allocation mechanisms. Random allocation of research funding may be a viable alternative to the current peer review system. In particular the “organized randomness” of a *modified lottery* is interesting, combining the benefits of randomization with some of the most valuable aspects of peer review. Still, many questions remain and this is certainly not a plea to allocate all research funds using lotteries without further research. But we need to be prepared to consider alternatives, *even though they are not perfect*, and modified lotteries should be part of the solution.

Keywords

Grant peer review, lottery, randomness, research funding, research integrity

1. Introduction

Competition for research funds has, in the recent decade, become hypercompetitive. Applications for grant money have multiplied, whereas research funding itself remained more or less constant (Fang & Casadevall, 2016a). Success rates are falling. This dramatically increases the time and effort individual researchers invest in attracting funding; and it requires funding bodies to differentiate the very best proposals from the barely weaker (Pier et al., 2018).

Currently, peer review is most often used to select proposals: scientists with relevant expertise review, judge, and score proposals; those scores are then used to rank the proposals and a cut-off (often based on the available funding) determines which proposals are funded. This system is widely accepted, easily understood and broadly trusted among researchers: more than three out of four researchers agreed that peer review is the best way to allocate funding to the strongest proposals and the best research (Hayes & Hardcastle, 2019). However, to say the least, it is far from perfect (Barnett, 2016; Bendiscioli, 2019). For an extensive criticism of the grant peer review system, I refer the reader to the RAND report by Guthrie, Ghiga and Wooding (2018). The most prominent problems are reviewed next.

1. Why the current peer review system is problematic
	1. Peer review is expensive

Grant peer review has outgrown its function of controlling how public money is spent (Baldwin, 2018). The majority of the time invested in drafting, writing, submitting, and re-writing applications for research funding (Bendiscioli, 2019; Hayes & Hardcastle, 2019; NIH, 2008) ‘goes to waste’ due to shrinking success rates (Roumbanis, 2019). These futile efforts, mainly at the expense of actually *doing* research, bring frustration and often even discouragement. It has been argued that writing proposals assists scientists to reflect on their past, current, and future research and peer review can provide constructive criticism that is under threat in the publish-or-perish climate (Reinhart & Schendzielorz, 2020, but see Roumbanis, 2020 for a strong counter-argument). However, using a mathematical model, Gross and Bergstrom (2019) show that as budgets are tighter and success rates decline, the value of the science researchers forbear by writing funding applications easily exceeds the funding budget.

Importantly, not only applicants invest a significant proportion of their (research) time in the quest for funding. Literally thousands of reviewers are also devoting substantial time to read, review, judge, discuss and select the applications (Adam, 2019; Bendiscioli, 2019). That more than 40% of reviewers declines an invitation to act as a reviewer can probably be interpreted as a sign that also reviewers are overburdened (<https://publons.com/community/gspr>).

* 1. Peer review is unreliable
		1. Peer review lacks precision “behind the comma”

Peer reviewers typically assign one global and/or several sub scores to proposals. However, due to the disbalance between the number of proposals and budgets, proposals have to be differentiated based on decimal scores and raw reviewer scores don’t allow using average scores with such precision (Fang & Casadevall, 2016b). Using differences in average scores as small as 0.01 on the NHS one-to-five scoring scale would require more than 38,000 reviewers to achieve reliability (Kaplan, Lacetera and Kaplan, 2008). Even reliably differentiating proposals at the level of *one* decimal would require each proposal to be scored by 384 reviewers (Kaplan, Lacetera and Kaplan, 2008).

Furthermore, there is a more fundamental objection to averaging reviewers’ scores: it assumes the scores not just represent an *ordinal* scale – “sufficient” is worse than “good”, “good” is worse than “excellent” – but an *interval* scale – i.e., there are similar distances between the scores (Sattler et al., 2015). However, considering that scoring criteria are often interpreted at reviewers’ own discretion (despite training, cf. 2.2.2 *Peer review scores are ill defined and inconsistent*), it is very unlikely that reviewer scores represent an interval scale.

* + 1. Peer review scores are ill defined and inconsistent

Typically, reviewers score proposals – and applicants – on several criteria, such as the quality of the proposal, its predicted scientific impact, societal relevance, and whether applicants have the expertise to successfully perform the proposed research. Increasingly, funders provide detailed instructions on how to interpret these criteria, but this does not avoid substantial interpretation from the part of the reviewers – not just what each criterium entails, but also how to weigh the different criteria in their final quotation. For example, reviewers differ strongly in how they take strengths and weaknesses into account; and weaknesses influence final scores stronger and more often than strengths (Pier et al., 2018). Experience does not necessarily attenuate this bias: whereas inexperienced reviewers have insufficient knowledge and understanding of both criteria and scores, experienced reviewers tend to overestimate their understanding and deviate deliberately from scoring criteria (Sattler et al., 2015).

Furthermore, and contrary to publication peer review, grant peer reviewers have relevant expertise in a field, but are most often not *the* experts. However, in particular for outstanding proposals, scores from field experts are more meaningful than an average score from groups of less specialist reviewers (Boudreau et al., 2016). Also in this respect does reviewer fatigue “water down” the meaning of reviewer scores.

Avin (2019) adds that although scientific peers may admittedly be best placed to estimate the scientific merit of proposals, *by definition* they cannot reliably do so, because what they need to estimate is in the future. What reviewers *can* do, is base their prediction on their experience with previous, similar projects and publication records, but this is problematic due to prediction error and regression to the mean.

As a result, there is limited agreement between reviewer scores (Graves, Barnett and Clarke, 2011; Pier et al., 2018). For example, more than three quarters of the variance in ‘scientific merit’ scores can be attributed to a combination of the specific reviewer, the interaction between reviewer and application, and random noise (Gallo, Sullivan and Glisson, 2016). Consequently, receiving favorable reviewer scores involves *luck*: adding or removing a single reviewer from a panel can lead to large differences in the final score, and, hence, ranking of a proposal (Kaplan, Lacetera and Kaplan, 2008). Consider the following illustrations: variability in reviewer scores could shift about one third of the funding decisions by the National Health and Medical Research Council of Australia from ‘fund’ to ‘not fund’ (or vice versa; Graves, Barnett and Clarke, 2011); and statistically correcting for the uncertainty included in the National Institutes of Health peer review scores could push up to 25% of the proposals from ‘not funded’ to ‘successful’ (and vice versa; Johnson, 2008). As such, the evidence is firmly against Reinhart and Schendzielorz’s (2020) defense of *legitimacy* as one of peer review’s strongest assets. Peer review may rely on “the result of critical deliberation according to scientific criteria” (p. S27), but if its results are unreliable, its legitimacy becomes null (Roumbanis, 2020).

* + 1. Peer review is biased

Peer review is biased (Fang & Casadevall, 2016b), even despite efforts to prevent it from being. Although probably only a minority of peer reviewers is deliberately or consciously biased, peer review is nevertheless favoring certain types of research and researchers.

* + - 1. The ‘usual suspects’

Peer review has been shown to disadvantage researchers based on race and ethnicity (Ginther et al., 2011); gender – even in blinded review because males tend to use more ‘vague’ language, which increases chances of success (Kolev, Fuentes-Medel and Murray, 2019; Pohlhaus et al., 2011; but see Sato et al., 2020 for an interesting discussion of gender bias in grant review); seniority; and the institute of the applicant (Daniels, 2015; Pier et al., 2018). Also cronyism, or rewarding grants to collaborators or researchers from your institution (or scientific field) influences funding decisions (Guthrie, Ghiga and Wooding, 2018; Jang et al., 2017; Mom, Sandström and van den Besselaar, 2018).

Furthermore, previous productivity – an applicant’s publication record, citations and grants – and reputation generate unfair advantages (Guthrie et al., 2019; for an in-depth discussion of these metrics as “perverse incentives”, see e.g. Nosek, Spies and Motyl, 2012). This can create a growing divide between initially equally skilled and talented researchers: due to intense competition, the same researchers will compete for funding from the same sources. But the variability in, and inconsistency of, reviewer scores may put one of those researchers just above the funding line and the other below, whereas *in reality*, their proposals were equally strong. The simple fact of getting this (first) funding may give the funded researcher the aura of being “more successful” or “worth funding”, increasing their future chances of being funded (the so-called Matthew effect; Bol, de Vaan and van de Rijt, 2018). The less fortunate researcher will not only lack this aura, but will have to keep investing time in new funding applications, which will keep them from actually doing research – and building the necessary track record to prove their scientific merit. This can easily result in a downward spiral.

It should be noted that *some* researchers seem to benefit from an initial setback in their career, working even harder and becoming more successful than researchers who received funding early in their career, but that is a small minority (Wang, Jones and Wang, 2019).

* + - 1. Innovative research is disadvantaged

Innovative research – ‘blue sky’ research; ‘edge science’ – is strongly disadvantaged by peer review (Adam, 2019; Hayes & Hardcastle, 2019; Li & Agha, 2015; Packalen & Bhattacharya, 2020) because it is inherently more risky. As it treads new ground its outcomes are more difficult to predict than those of safe, incremental research. As such, the current peer review system strongly promotes low-risk research. Packalen and Bhattacharya (2020) show, for example, that NIH funding is mostly reserved for ‘established’ research: there is a 7 to 10 year-gap between the emergence of new ideas and those ideas receiving maximal NIH funding.

Moreover, scoring applicants’ experience is tricky for innovative science and can create a double handicap: innovative research is novel by definition, hence relevant experience is often lacking. Reviewers could then take the track record of applicants into account – their previous successes may suggest they know what they are doing – but this obviously disadvantages courageous early career scientists.

Furthermore, reviewers’ definition of “excellent” may be at odds with “paradigm-changing”. Precisely because the exact methods to answer innovative questions and their expected results are difficult to predict, innovative research is at risk of being regarded as less rigorous (Gallo et al., 2018). Again, as more scientists decline invitations to act as peer reviewer, this problem is aggravated because it may “take one to know one”: a reviewer who has only limited knowledge of a field will have to make a guess how sound the proposed methods are – maybe an educated guess, but a guess nevertheless. In particular when prevailing paradigms are challenged, reviewers’ resistance can be fierce, as is illustrated by the anecdotal stories of Nobel prize winners and their struggles to get their research published (cf. “John Snow’s Grant Application”, Rothman, 2016).

* + - 1. Interdisciplinary research is disadvantaged

Peer review often underestimates the scientific merit of interdisciplinary research (Pluchino et al., 2019). Bromham, Dinnhage and Hua (2016) used the number of Field of Research codes from more than 18.000 applications to the Australian Research Council’s Discovery Programme as a proxy of interdisciplinarity and report a negative correlation between their interdisciplinarity measure and the chance to get funded.

Furthermore, reviewers tend to be wary to review interdisciplinary proposals because they don’t feel qualified, lacking expertise in *every* field of the application. Unfortunately, this results in proposals being reviewed by scientists with possibly even less relevant expertise (Hayes & Hardcastle, 2019).

In addition, review criteria and reviewer scores are inconsistent between disciplinary fields, which renders interdisciplinary applications’ scores difficult to combine into a global score (Roumbanis, 2020).

* 1. Peer review leads to perversions

One perverse outcome of the current system of grant peer review is “grantsmanship”: the skill of writing attractive projects, which has little or no relation to a researcher’s ability to do sound research. In fact, grantsmanship distracts from the content of a proposal – the very thing reviewers should be evaluating – and adds noise to an already noisy system (Dinov, 2020). The paradoxical importance of grantsmanship is reflected in the administrative resources universities devote to it, the manuals that are offered – even by the Office of Research Integrity (<https://ori.hhs.gov/education/products/wsu/writing_gra.html>) – and by commercial parties that advertise substantial higher success rates than what is common.

Another questionable outcome of the current peer review system is that it puts pressure on scientific integrity, because researchers may choose to ‘optimise’ their curricula by inflating their publication list (Dinov, 2020).

Importantly, the problems with peer review become even more problematic considering that research funding is a zero-sum game. When one researcher secures funding, this is *always* at the expense of other researchers.

1. Not just doom and gloom

It is only fair to admit that the peer review system is not entirely unreliable or useless: there is consensus that excellent projects are effectively identified (e.g. the top 20%; Fang & Casadevall, 2016b) and that peer reviewers’ added value is exactly in their ability to identify the strongest applications (Li & Agha, 2015) – even without reviewer training (Sattler et al. 2015). Similarly, low-quality projects are consistently weeded out by review panels (Pier et al., 2018; see also Gallo, Sullivan and Glisson, 2016).

Furthermore, within projects receiving R01 funding (NIH’s primary grant mechanism), higher peer review scores are associated with more publications, more high-impact publications, citations, and patents stemming from the projects; even after controlling for applicants’ previous accomplishments (publications and citations) and institution (Li & Agha, 2015; but see Fang, Bowen and Casadevall, 2016).

Again, this may justify the claim that peer review is probably the best system currently in use. It does not, however, imply that there are no equally good or even better alternatives. Remember that hypercompetition for funding requires funders to decide which projects are the best *within* the pool of outstanding projects (Fang, Bowen and Casadevall, 2016). However, above a certain quality threshold, the peer review process becomes completely random (Pier et al., 2018) and when sufficient funding is available to finance projects beyond the absolute ‘top selection’, projects in the ‘middle bracket’ cannot be accurately ranked (Adam, 2019; Fang, Bowen and Casadevall, 2016).

1. Improving peer review

A list of changes and modifications have been suggested and tested to deal with the aforementioned problems of peer review (for an overview, cf. NIH, 2008), with varying degrees of success.

4.1 Solving inconsistencies

First of all, reviewer training in rating and weighing proposals’ various aspects has been suggested to increase inter-rater reliability (Sattler et al., 2015). Similarly, training has been suggested to better detect flaws and weaknesses in applications (Schroter et al., 2004). However, despite reviewers’ enthusiasm to receive training, its effects are limited (Pier et al., 2018) and short-lived (Schroter et al., 2004). One might assume that reviewer *experience* may improve peer review. However, experienced reviewers can extract more information from an application, which increases the chances that they discover weaknesses and possible flaws. In addition, they tend to weigh those weaknesses heavier than possible strengths (Boudreau et al., 2016).

Having each proposal reviewed by more reviewers would increase the reliability of the average scores and rankings, but as mentioned above this would not only substantially increase the necessary resources, achieving the precision to reliably rank projects requires an unrealistically large pool of reviewers (Kaplan, Lacetera and Kaplan, 2008).

To avoid a single bad score ruining an applicant’s chances, Daniels (2015) suggested to remove the lowest scores from review panels. This would also even the odds for more daring proposals that cannot convince every single reviewer. Relatedly, Kaplan, Lacetera and Kaplan (2008) have suggested to use the range in reviewer scores to detect innovative projects.

Holliday and Robotin (2010) even suggest to implement a reverse strategy: instead of selecting the best applications, they suggest to remove the weakest projects, resulting in the ‘survival’ of only the very best projects after several rounds.

An interesting, although for the time being theoretical, proposal is from Bedessem (2020): instead of comparing (average) reviewers’ scores for individual projects, he proposes to have reviewers select the projects they want to see funded out of a pool of applications. Applications that are unanimously selected could then receive funding. Bedessem also suggests expanding reviewer panels to include the target audience: e.g. patients, politicians, farmers,… to make scores on ‘societal relevance’ more tangible. However, this would obviously increase the cost of the review process.

4.2 Reducing the cost/burden

Shortening applications has been suggested to alleviate some of the burden of writing proposals, but applicants may invest as much time in the shorter applications as they did before – trying to “just get it right” (Shepherd et al., 2018). A two-step review process, only inviting applicants to submit a full proposal after screening brief initial applications is used by e.g. the European Research Council (ERC) and Cancer Research UK (CRUK). In a study evaluating UK’s National Institutes for Health Research’s (NIHR) Research for Patient Benefit (RfPB) Programme, a two-step process was shown to be 30% cheaper and more efficient. Moreover, applicants who are not selected receive the decision sooner in the process – although the final decision about the full proposals were delayed in time (Morgan et al., 2020).

Online reviewer panels replacing on-site meetings have been shown to reduce the burden on reviewers, to increase efficiency, and to save substantial time – and travel costs (Shepherd et al., 2018).

4.2.1 Abandoning peer review altogether: Egalitarian distribution of funding

A system that could eradicate the cost of peer review altogether is egalitarian distribution of funding (Ioannidis, 2011). It would ascertain every researcher of – at least minimal – resources (Vaesen & Katzav, 2017). However, even this system would require a body of some sort to check whether researchers meet the demands to receive funding – and, hence, some sort of application. Furthermore, the cost of doing research differs hugely between, and even within, scientific disciplines, which would require a well-considered allocation mechanism to remain somewhat efficient. As such, even this ‘administration-free’ system would also have to rely on substantial resources.

4.3 Funding researchers, not research

Financing researchers instead of projects has been proposed (Ioannidis, 2011). The Howard Hughes Medical Institute (HHMI[[1]](#footnote-1)) assigns funding to individuals and awards them great freedom to experiment. Furthermore, and importantly, HHMI is failure-tolerant, rewarding long-term success instead: HHMI researchers can use their funding flexibly and shift resources if their initial attempts are unsuccessful.

As a result, HHMI funded researchers produce more high-impact publications than a group of similarly accomplished researchers receiving NIH funding (Azoulay, Zivin and Manso, 2011). Not surprisingly, HHMI researchers also produce a larger number of ‘flops’ than their NIH-colleagues, but this illustrates the efficacy of the program encouraging them to try out new research lines.

Bollen et al. (2014) take this proposition to the extreme and suggest to divide research funding equally among (qualified) researchers, with the obligation to assign a fixed proportion of that money to their fellow researchers. This would ascertain each researcher of minimal funding, still dividing the majority of funding through a “bottom-up” mechanism. Each year, funding would be redistributed (and those who get the most would also be redistributing the most). Although this is certainly a more “bottom-up” (or “decentralised”) way of distributing research funding, it requires substantial resources to control the finances.

1. Introducing organised randomness

In sum, the peer review system suffers from important shortcomings and suggested solutions have small, if any, and often short-lived effects (Fang & Casadevall, 2009) – as Roumbanis (2020) puts it: “some of the dilemmas inherent in peer review are almost impossible to neutralize” (p. S130). Hence, although peer review may be the best system in use, this does not mean it necessarily outperforms all alternatives (Li & Agha, 2015). In fact, it is time for an open discussion about alternative allocation mechanisms and we could start by admitting that there is little evidence that the current system is the best (Guthrie, Ghiga and Wooding, 2018).

Recently, there are calls to formalise the ‘chance’ aspect that is inherent to peer review by allocating research funding *randomly*, i.e. using a lottery (Graves, Barnett and Clarke, 2011). This could save applicants, reviewers and funders considerable time they would otherwise devote to write, re-write, read, evaluate, and select proposals. The time saved could be devoted to actual science, the money saved could be distributed by funders: it would sharply increase the process’ cost-efficiency. As a positive side effect, applicants would receive decisions sooner.

* 1. Mathematical models

The randomness of the peer review system has been modelled mathematically. Using his agent-based model, Avin (2015) has shown that the system of grant peer review is equivalent to the “proverbial man searching for his keys under the lamp-post because “that is where the light is”” (p. 153). Funding by peer review is *identical* to allocating funding based on an initial triage (separating outstanding from good proposals) and subsequently allocating funding randomly in the latter group. The crucial differences between both procedures are that random allocation is cheaper; and more fair because proposals that cannot be statistically differentiated from one another are being treated as equal.

Building on Avin’s model, Harnagel (2019) created a mid-level model between ‘theoretical’ and ‘empirical’ by integrating empirical data on bibliometric parameters. Using real-world publication and citation counts, she showed that peer review was not the best model to generate significant science; and randomness would benefit scientific creativity.

Gross and Bergstrom (2019) likewise showed that randomly allocating funding to proposals that pass a minimum quality threshold would decrease the costs of the process. Because applicants would cease to overly invest in their attempts to write the ‘best’ application, they would also sacrifice less of their valuable and costly research time. The model showed that a lottery in which 45% of applicants qualify is as efficient as peer review selecting the 45% ‘best’ proposals (independent of the chance of getting funded in the lottery). Furthermore, with the current emphasis on individual achievements for evaluation and promotion, some researchers may increasingly apply for funding *even though* *this hampers* their scientific productivity, further reducing the efficiency of the peer review system. Although Gross and Bergstrom admit that their model did not (yet) include the possibility to resubmit partial revisions after first rejections, they are resolute: the smaller the success rate, the larger the benefits of a lottery.

With a similar model, Pluchino, Biondo and Rapisarda (2018) and Pluchino et al. (2019) showed that although a certain amount of talent and skill is *necessary*, it is not a *sufficient* condition for success. Contrary to common belief, it also takes luck to be successful[[2]](#footnote-2). In fact, moderately talented individuals have a better chance to “make it”: because there are more moderately talented researchers than super talented, the former’s chances to be lucky are higher. These findings, importantly, also call into question the meritocratic model and the metrics that are used to evaluate applicants’ track record (Osterloh & Frey, 2020) – and drive the Matthew effect.

* 1. The case for a modified lottery

Considering these findings, random allocation of research funding may be a viable alternative to the current peer review system. In particular a *modified lottery* is interesting, combining the benefits of randomisation with some of the most valuable aspects of peer review. Specifically, a modified lottery includes a first *triage* phase, minimally removing proposals with obvious shortcomings (Fang & Casadevall, 2016b; Liu et al., 2020). Ideally, outstanding proposals should be allowed to bypass the lottery. However, if budgets are so tight that there are more outstanding proposals than the payline allows, Fang, Bowen and Casadevall (2016) suggest to organise a lottery in the top tier nevertheless.

The resources for projects in the middle tier can be assigned purely random, or through what Boyle (1998) called a *graduated lottery*. That is: after a rudimental ranking of proposals, for example 50% of the proposals in the lowest quartile is randomly selected. These selected proposals are added to the second-lowest quartile and again 50% of the proposals in this new pool is selected. These are added to the second-highest quartile, etc. This way, proposals with a higher ranking would have an increasing chance of being funded, without relying on peer review scores.

Gross and Bergstrom (2019) suggest to use a ‘multitiered’ lottery with *shortened* and/or *simplified* proposals as an interim solution between peer review and a pure lottery: it reduces the writing effort, whereas review panels would still have the opportunity to reward excellence.

A modified lottery further allows to ‘tweak’ some of the selection criteria – levelling the playing field for minority groups; or counteracting reviewer bias without imposing explicit quota (Adam, 2019). Importantly, by defining those selection criteria, they become explicit, instead of being intertwined with others and obscured in (sub)scores. The most important advantage of this approach is, no doubt, that relevant aspects can be accounted for without pretending to have accurate estimates.

Furthermore, specific ‘tracks’ could be created for specific groups. For example, the NIH has funding streams for early career scientists[[3]](#footnote-3) (Daniels, 2015) and the University of Leuven has a specific track for “bold interdisciplinary bottom-up research projects” (ID-N grants). Tracks could run in parallel, or be organised alternately. The University of Leuven, for example, has alternate calls for early career scientists and collaborative research.

In addition, a modified lottery would create a more realistic image of the merit of achieving funding (Gross & Bergstrom, 2019) and prevent others from experiencing a rejection as a personal failure (Osterloh & Frey, 2020). It would create a ‘substantiated meritocracy’ and reduce competition – which would promote scientific integrity and counteract the Matthew effect. A system that uses simplified, brief applications can relief some of the hyper competition (Roumbanis, 2019); knowing the selection is based on a lottery can also to some extent restore the balance between the time spent writing grant proposals and doing research (Gross & Bergstrom, 2019).

Finally, a lottery could render visible the vast amount of relevant and valuable research that cannot secure funding in the current system (Fang & Casadevall, 2016a).

Evidently, modified lotteries also require substantial resources to write, read, evaluate and rank the applications. Nevertheless, its cost would presumably be (substantially) lower than those of the current system. Furthermore, a lot of important questions remain unanswered. Should eligible but unsuccessful applications be automatically added to the next round (Barnett et al., 2014), or should they be resubmitted? Should applicants be allowed to resubmit immediately, or should there be a waiting period? If unsuccessful proposals are automatically added to the next round: how long do they remain in the pool? Should the number of proposals per applicant be limited? Should there be a ‘cooling-off’ period after several unsuccessful applications (Graves, Barnett and Clarke, 2011; Barnett et al., 2014)? Or, on the contrary: should there be a separate lottery for researchers who are eligible for several subsequent draws but never get funded? Or should a review panel decide (Fang & Casadevall, 2016b)? And what about researchers who are successful repeatedly: can they submit additional proposals?

The list of unanswered questions is long. However, these same questions have to be answered when setting up grant peer review. It seems it is only fair to conclude that, regardless of the practical details, a modified lottery is more fair than peer review.

* 1. Real-life case studies

For the time being, only a few funders have implemented randomness: the Health Research Council of New Zealand with its Explorer Grant funds[[4]](#footnote-4) and the New Zealand’s government’s Science for Technological Innovation National Science Challenge (SfTI); the Volkswagen Foundation’s Experiment! grants[[5]](#footnote-5); and the Swiss National Science Foundation (SNSF) postdoctoral fellowships for early-career scientists (Adam, 2019).

Due to their novelty, little is known about the effects of these programs, although a recent study showed that New Zealand’s researchers generally evaluate the random allocation of Explorer Grant funds positively. It should be noted that predominantly funded researcher were enthusiastic: 78% thought random allocation of funding is acceptable, compared with 44% for declined applicants (Liu et al., 2020). In general, 63% of those surveyed said that random allocation of funding is acceptable for Explorer Grant funds, but only 40% favoured the idea for other grant types (Liu et al., 2020).

1. Conclusion

This may be the time to admit that the current grant allocation system may have been performant when it was installed, but it doesn’t fit the current highly competitive research arena. We need to be prepared to consider alternatives, *even though they are not perfect* (Guthrie, Ghiga and Wooding, 2018).

Introducing “organised randomness” into the allocation of research funds could be a way to overcome some of the problems of the current grant peer review system. Importantly, it would not mean we give into despair about the current system, admit we don’t know what we are doing, be a sign of reluctance to thoroughly evaluate the research we are financing (often with tax money), or “the acceptance of a regime of indifference” (Reinhart & Schendzielorz, 2020; p. S28). In contrast, it would demonstrate our commitment to admitting and tackling the shortcomings and limits of the current system (Barnett, 2016).

Modified lotteries could well be the best – i.e. most *fair* – alternative to peer review. As Stone (2009) argued, because random allocation avoids decisions being based on reason(s), it side-lines *bad* reasons (i.e., bias). However, because it also precludes *good* (i.e., relevant) reasons from guiding the decision, its use must be carefully considered. Roumbanis (2019) recently argued that a lottery additionally would make the process of research funding more dynamic; and due to reduced bias also more divers, fair and impartial. Gildenhuys (2020) additionally argues that lotteries make science itself, as an institution, fairer.

Organised randomness could come with reduced costs, and even if the initial triage phase and subsequent lotteries (and the administration of resubmissions etc.) would not reduce the costs and funder resources, *decentralising* the decision could avoid reviewer bias (cf. Bedessem, 2020; but see also Barnett et al., 2017; Bollen et al., 2014).

Still, many questions remain and this paper is not intended as a plea to allocate all research funds using lotteries without further research. Excellent research should get funded and some research fields can only thrive with long-term funding, for example when large or expensive equipment is necessary (Stone, 2009; see also Roumbanis, 2019). Needless to say also modified lotteries have loopholes and may be subject to bias. But many of the shortcomings of modified lotteries are part of the peer review system as well, so it wouldn’t be fair to use the argument to dismiss only modified lotteries. Modified lotteries are part of the solution, not *the* solution (Roumbanis, 2019).

ACKNOWLEDGEMENTS

The author thanks Prof. Dr. G. Storms for his comments on the paper and for relentlessly propagating scientific integrity. The “Denkgroep Optimalisering Onderzoeksmiddelen” of the University of Leuven provided the inspiration for the paper and their report was most insightful.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Adam D. 2019. Science Funders Gamble on Grant Lotteries. Nature 575, 574-75. <https://doi.org/10.1038/d41586-019-03572-7>

Avin S. 2015. Breaking the grant cycle: on the rational allocation of public resources to scientific research projects (Doctoral thesis). <https://doi.org/10.17863/CAM.16172>

Avin S. 2019. Centralized Funding and Epistemic Exploration. The British Journal for the Philosophy of Science 70, 629-56. <https://doi.org/10.1093/bjps/axx059>

Azoulay P, Zivin J, Manso G. 2011. Incentives and creativity: Evidence from the academic life sciences. The RAND Journal of Economics 42, 527-54. <https://www.jstor.org/stable/23046811>

Baldwin M. 2018. Scientific Autonomy, Public Accountability, and the Rise of “Peer Review” in the Cold War United States. Isis 109, 538-58. <https://doi.org/10.1086/700070#d300098e656>

Barnett AG. 2016. Funding by lottery: political problems and research opportunities. mBio 7, e01369-16. <https://doi.org/10.1128/mBio.01369-16>

Barnett AG, Clarke P, Vaquette C, Graves N. 2017. Using Democracy to Award Research Funding: An Observational Study. Research Integrity and Peer-Review 2, 1-9. <https://doi.org/10.1186/s41073-017-0040-0>

Barnett A, Herbert D, Clarke P, Graves N. 2014. The research lottery: the pressures on the Australian grant system. AQ: Australian Quarterly 85, 4-9.

Bedessem B. 2020. Should we fund research randomly? An epistemological criticism of the lottery model as an alternative to peer review for the funding of science. Research Evaluation 29, 150-7. <https://doi.org/10.1093/reseval/rvz034>

Bendiscioli S. 2019. The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decision-making. EMBO Reports 20, e49472. <https://doi.org/10.15252/embr.201949472>

Bol T, de Vaan M, van de Rijt A. 2018. The Matthew effect in science funding. Proceedings of the National Academy of Sciences 115, 4887-90. <https://doi.org/10.1073/pnas.1719557115>

Bollen J, Crandall D, Junk D, Ding Y, Börner K. 2014. From Funding Agencies to Scientific Agency. EMBO Reports 15, 131-3. <https://doi.org/10.1002/embr.201338068>

Boyle C. 1998. Organizations selecting people: how the process could be made fairer by the appropriate use of lotteries. Journal of the Royal Statistical Society: Series D (The Statistician) 47, 291-321. <http://www.jstor.org/stable/2988669>

Bromham L, Dinnhage R, Hua X. 2016. Interdisciplinary research has consistently lower funding success. Nature 534, 684-7. <https://doi.org/10.1038/nature18315>

Boudreau KJ, Guinan EC, Lakhani KR, Riedl C. 2016. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. Management Science 62, 2765-83. <https://doi.org/10.1287/mnsc.2015.2285>

Daniels RJ. 2015. A generation at risk: young investigators and the future of the biomedical workforce. Proceedings of the National Academy of Sciences 112, 313-8. <https://doi.org/10.1073/pnas.1418761112>

Dinov ID. 2020. Flipping the grant application review process. Studies in Higher Education 45, 1737-45. <https://doi.org/10.1080/03075079.2019.1628201>

Else H. 2019. Male Researchers' 'vague' Language More Likely to Win Grants. Nature News. <https://doi.org/10.1038/d41586-019-01402-4>

Fang FC, Bowen A, Casadevall A. 2016. NIH peer review percentile scores are poorly predictive of grant productivity. eLife 5, e13323. <https://doi.org/10.7554/eLife.13323>

Fang FC, Casadevall A. 2009. NIH peer review reform--change we need, or lipstick on a pig? Infection and immunity 77, 929-32. <https://doi.org/10.1128/IAI.01567-08>

Fang FC, Casadevall A. 2016a. Research funding: the case for a modified lottery. mBio 7, e00422-16. <https://doi.org/10.1128/mBio.00422-16>

Fang FC, Casadevall A. 2016b. Grant funding: Playing the odds. Science 352, 158. <https://doi.org/10.1126/science.352.6282.158-a>

Gallo SA, Sullivan JH, Glisson SR. 2016. The Influence of Peer Reviewer Expertise on the Evaluation of Research Funding Applications. PLOS One 11, e0165147. <https://doi.org/10.1371/journal.pone.0165147>

Gallo S, Thompson L, Schmaling K, Glisson S. 2018. Risk evaluation in peer review of grant applications. Environment Systems and Decisions 38, 216-29. <https://doi.org/10.1007/s10669-018-9677-6>

Gildenhuys P. 2020. Lotteries make science fairer. Journal of Responsible Innovation 7:sup2, S30-S43. <https://doi.org/10.1080/23299460.2020.1812485>

Ginther DK, Schaffer WT, Schnell J, Masimore B, Liu F, Haak LL, Kington R. 2011. Race, Ethnicity, and NIH Research Awards. Science 333, 1015-9. <https://doi.org/10.1126/science.1196783>

Graves N, Barnett A, Clarke P. 2011. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. British Medical Journal 343, d4797. <https://doi.org/10.1136/bmj.d4797>

Gross K, Bergstrom CT. 2019. Contest models highlight inherent inefficiencies of scientific funding competitions. PLOS Biology 17, e3000065. <https://doi.org/10.1371/journal.pbio.3000065>

Guthrie S, Ghiga I, Wooding S. 2018. What do we know about grant peer review in the health sciences? An updated review of the literature and six case studies. RAND report, RAND Corporation. <https://www.rand.org/pubs/research_reports/RR1822.html>

Guthrie S, Rodriguez Rincon D, McInroy G, Ioppolo B, Gunashekar S. 2019. Measuring bias, burden and conservatism in research funding processes [version 1; peer review: 1 approved, 1 approved with reservations]. F1000Research 8, 851. <https://doi.org/10.12688/f1000research.19156.1>

Harnagel A. 2019. A mid-level approach to modeling scientific communities. Studies in History and Philosophy of Science Part A 76, 49-59. <https://doi.org/10.1016/j.shpsa.2018.12.010>

Hayes M, Hardcastle J. 2019. Grant review in focus. Global state of peer review series. Publons report. Downloaded from <https://publons.com/community/gspr/grant-review>

Holliday C, Robotin M. 2010. The Delphi process: a solution for reviewing novel grant applications. International Journal of General Medicine 3, 225-30. <https://doi.org/10.2147/IJGM.S11117>

Ioannidis JPA. 2011. More time for research: fund people not projects. Nature 477, 529-31. <https://doi.org/10.1038/477529a>

Jang D, Doh S, Kang G-M, Han D-S. 2017. Impact of Alumni Connections on Peer Review Ratings and Selection Success Rate in National Research. Science, Technology, & Human Values 42, 116-43. <https://doi.org/10.1177/0162243916665466>

Janosov M, Battiston F, Sinatra R. 2020. Success and luck in creative careers. EPJ Data Science 9, 9. <https://doi.org/10.1140/epjds/s13688-020-00227-w>

Johnson VE. 2008. Statistical analysis of the National Institutes of Health peer review system. Proceedings of the National Academy of Sciences 105, 11076-80. <https://doi.org/10.1073/pnas.0804538105>

Kaplan D, Lacetera N, Kaplan C. 2008. Sample Size and Precision in NIH Peer Review. PLOS One 3, e2761. <https://doi.org/10.1371/journal.pone.0002761>

Kolev J, Fuentes-Medel Y, Murray F. 2019. Is blinded review enough? How gendered outcomes arise even under anonymous evaluation. National Bureau of Economic Research Working Paper No. 25759. <https://www.nber.org/papers/w25759>

Li D, Agha L. 2015. Research funding. Big names or big ideas: do peer-review panels select the best science proposals? Science 348, 434-8. <https://doi.org/10.1126/science.aaa0185>

Liu M, Choy V, Clarke P, Barnett A, Blakely T, Pomeroy L. 2020. The acceptability of using a lottery to allocate research funding: a survey of applicants. Research Integrity and Peer Review 5, 3. <https://doi.org/10.1186/s41073-019-0089-z>

NIH. 2008. Enhancing peer review: A self-study by the NIH in partnership with the scientific community to strengthen peer review in changing times. Available from <http://enhancing-peer-review.nih.gov/meetings/EnhancingPeerReviewACD2-21-08.pdf>

Mom C, Sandström U, van den Besselaar P. 2018. Does cronyism affect grant application success? The role of organizational proximity. Proceedings of the 23rd International Conference on Science and Technology Indicators, 12-14 September 2018, Leiden, The Netherlands. <https://openaccess.leidenuniv.nl/bitstream/handle/1887/65246/STI2018_paper_263.pdf>

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, e0230118. <https://doi.org/10.1371/journal.pone.0230118>

Nosek BA, Spies JR, Motyl M. 2012. Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. Perspectives on Psychological Science 7, 615–31. <https://doi.org/10.1177/1745691612459058>

Osterloh M, Frey BS. 2020. How to avoid borrowed plumes in academia. Research Policy 49, 103831. <https://doi.org/10.1016/j.respol.2019.103831>

Packalen M, Bhattacharya J. 2020. NIH funding and the pursuit of edge science. Proceedings of the National Academy of Sciences 117, 12011-6. <https://doi.org/10.1073/pnas.1910160117>

Pier EL, Brauer M, Filut A, Kaatz A, Raclaw J, Nathan MJ, Ford CE, Carnes M. 2018. Low agreement among reviewers evaluating the same NIH grant applications. Proceedings of the National Academy of Sciences 115, 2952-7. <https://doi.org/10.1073/pnas.1714379115>

Pluchino A, Biondo AE, Rapisarda A. 2018. Talent vs luck: The role of randomness in success and failure. Advances in Complex Systems 21(03n04), 1850014. <https://doi.org/10.1142/S0219525918500145>

Pluchino A, Burgio G, Rapisarda A, Biondo AE, Pulvirenti A, Ferro A, Giorgino T. 2019. Exploring the role of interdisciplinarity in physics: Success, talent and luck. PLOS One 14, e0218793. <https://doi.org/10.1371/journal.pone.0218793>

Pohlhaus JR, Jiang H, Wagner RM, Schaffer WT, Pinn VW. 2011. Sex differences in application, success, and funding rates for NIH extramural programs. Academic Medicine 86, 759-67. <https://doi.org/10.1097/ACM.0b013e31821836ff>

Reinhart M, Schendzielorz C. 2020. The lottery in Babylon—On the role of chance in scientific success. Journal of Responsible Innovation 7:sup2, S25-9. <https://doi.org/10.1080/23299460.2020.1806429>

Rothman KJ. 2016. John Snow’s Grant Application. Epidemiology 27, 311-3. <https://doi.org/10.1097/EDE.0000000000000453>

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, & Human Values 44, 994-1019*.* <https://doi.org/10.1177/0162243918822744>

Roumbanis L. 2020. Two dogmas of peer-reviewism. Journal of Responsible Innovation 7:sup2, S129-S133. <https://doi.org/10.1080/23299460.2020.1855806>

Sato S, Gygax PM, Randall J, Schmid Mast M. 2020. The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. Higher Education. <https://doi.org/10.1007/s10734-020-00626-y>

Sattler DN, McKnight PE, Naney L, Mathis R. 2015. Grant Peer Review: Improving Inter-Rater Reliability with Training. PLOS One 10, e0130450. <https://doi.org/10.1371/journal.pone.0130450>

Schroter S, Black N, Evans S, Carpenter J, Godlee F, Smith R. 2004. Effects of training on quality of peer review: Randomized control trial. British Medical Journal 328, 673-5. <https://doi.org/10.1136/bmj.38023.700775.AE>

Shepherd J, Frampton GK, Pickett K, Wyatt JC. 2018. Peer review of health research funding proposals: A systematic map and systematic review of innovations for effectiveness and efficiency. PLOS One 13, e0196914. <https://doi.org/10.1371/journal.pone.0196914>

Stone P. 2009. The logic of random selection. Political theory 37, 375-97. [https://doi.org/10.1177%2F0090591709332329](https://doi.org/10.1177/0090591709332329)

Vaesen K, Katzav J. 2017. How much would each researcher receive if competitive government research funding were distributed equally among researchers? PLOS One 12, e0183967. <https://doi.org/10.1371/journal.pone.0183967>

Wang Y, Jones BF, Wang D. 2019. Early-career setback and future career impact. Nature Communications 10, 4331. <https://doi.org/10.1038/s41467-019-12189-3>

1. https://www.hhmi.org/programs/biomedical-research/investigator-program [↑](#footnote-ref-1)
2. Janosov et al. (2020) use a similar model to show that the role of luck in science is no different from other ‘creative’ domains such as music, literature, and film. [↑](#footnote-ref-2)
3. New Innovator Awards (DP2) for exceptionally creative early stage investigators, the Early Independence Awards (DP5) for young scientists to pursue independent research immediately after a terminal degree, and the Pathway to Independence Awards (K99/R00), which combine a mentored research phase with later, independent research support. [↑](#footnote-ref-3)
4. http://www.hrc.govt.nz/funding-opportunities/researcher-initiated-proposal/explorer-grants [↑](#footnote-ref-4)
5. https://www.volkswagenstiftung.de/en/funding/our-funding-portfolio-at-a-glance/experiment [↑](#footnote-ref-5)