



Evidence from ten years of research contract management

Dr Jez Weston & Dr Jason Gush, Royal Society of New Zealand

The Royal Society of New Zealand regularly reviews the efficiency and effectiveness of our research contract management with the goal of improving these funding schemes. Here we present a summary of evidence from our last ten years of experience in research funding to inform a paper on peer review recently released by the Prime Minister's Chief Science Advisor¹, and to contribute to the wider debate about the evolution of the research system.

The Society holds responsibility for approximately \$65 million of public investment as part of the Government's larger investment portfolio. By size, this is dominated by the Marsden Fund (\$54 million) but we also administer a range of smaller funds, including the Rutherford Discovery Fellowships and a range of international travel grants. The Society's evaluation of these investments provides information from an operational level that should feed back into the design of research and innovation policy and processes. The data we present here addresses many of the recurring questions about how best to fund research in a manner that is both effective and efficient, where effective means reliable, fair, and accountable while delivering high quality, relevant, and timely research outcomes and efficient means the lowest possible transaction costs and burden upon researchers.

Funding process possibilities

There is a range of funding scales - from a travel grant of a few thousand dollars to a research platform of tens of millions; equally there is a range of processes that can be used to make research funding decisions. The Royal Society of New Zealand has long experience over the majority of the range of funding scales, from the smallest grants up to Marsden grants of over a million dollars. Similarly, we use peer review in a range of processes for awarding that funding, from those that emphasise the quality of the individual applicant to those that emphasise the quality of the research idea. This experience provides information about several funding schemes that represent distinct combinations of those possibilities. We regularly review our processes to ensure that we match particular decision processes with the particular combination of circumstances that each funding scheme represents.

We suggest that a key question for any discussion of research funding processes is not just "how does peer review support funding processes to deliver the best possible outcomes?" but "how does peer review contribute to each specific kind of funding decision?" Answering this later question should involve considering how the role of peer review changes with the scale of grants and as decisions focus more or less upon individuals rather than ideas.

There is no best process for research funding decision-making. Each funding process must take into account the particulars of each scheme.

¹ Gluckman, "[Which science to fund: time to review peer review?](#)", 2012

Marsden – Attempting to choose excellence

The Marsden fund is one of New Zealand's few funds that target investor-initiated excellent research. The fund uses a two-stage process to reduce effort for all involved. In the first round, preliminary one-page proposals are assessed by panels of New Zealand-based peers. In the second round, five-page proposals are scored by three independent reviewers and then recommended by the panels on the basis of those scores and panel discussions.

Applications to the Marsden Fund have a low success rate, down to 7.7% this year. This low rate does not reduce the numbers of applications received each year, due to the prestigious nature of the fund and the lack of other funding sources for much of the possible research.

While it is considered the best research fund in terms of research quality, the awards are described as being for excellent research rather than for the best research, recognising the role of uncertainty in the research assessment process. The low success rate requires the assessment process to try to identify the top 7.7% of the proposals. Of the applications that we receive, around 15% of those that reach the second stage are described by reviewers as excellent, implying that the fund is a lottery with even odds – an excellent application has a 50% chance of being funded despite that application being assessed to be as good as other applications that are funded.

The Royal Society of New Zealand's best estimate of total dollar cost of Marsden application process, including the cost to applicants and reviewers is \$10-\$20 million for a fund size of \$54.6 million. This cost breaks down into over 80% to the applicants in writing proposals, 10% to the reviewers and panellists, and less than 10% to the Society in running the fund.

Basic decision-making theory – inherent limits and the beneficial impact of more reviewers

Decision making theory must necessarily begin with Arrow's impossibility theorem — there is no rank-order voting system that can preserve the ranked preferences of individuals into a group-wide ranking whilst meeting the complete criteria set of:

- If every voter prefers alternative X over alternative Y, then the group prefers X over Y.
- If every voter's preference between X and Y is unchanged, then the group's preference between X and Y will also be unchanged (even if voters' preferences between other pairs like X and Z change).
- There is no single "dictator" with the power to always determine the group's preference.

Therefore, whilst a ranked voting system offers certain advantages, as with other systems, it does not obey all criteria that are commonly associated with 'fairness'.

Libby and Glass investigate optimal number of judges on a panel.² They find that this number depends on the evaluation rule, the accuracy of the judges and the (cost per judge)/(cost per error) ratio. Whilst increasing the number of accurate judges on a panel improves panel accuracy, increasing the number of less accurate judges actually decreases accuracy.

Impact of reviewer choice – nationality produces no bias, personal contact does

Concerns over the impact of the choice of peer reviewers have been an ongoing topic of discussion by the Marsden Fund Council, so in 2011 the Society was asked to investigate possible biases. Two concerns were raised, that New Zealand reviewers were biased against domestic research and that investigator nominated reviewers were biased in favour of their nominators.

Considering the two thousand referee reports from 2008-2010, 95% of reviewers are overseas. There is no significant difference in scores received from Oceanic reviewers when set against those from Asia, the Americas, or Europe.

² Libby & Glass, "[The Calculus of Committee Composition](#)", 2010

Investigator-nominated reviewers were found to give more positive scores than independently sought reviewers. They are now no longer used and all reviewers are independently sought.

Avoiding comparing across disciplines

Marsden accepts proposals across all research disciplines. Any particular process may be biased in favour of particular disciplines, for instance if citation-based metrics are used, then humanities researchers may be penalised. To avoid this effect, the ten subject-based panels limit their comparisons to those between similar disciplines. However, this issue is a concern for wider comparisons such as for the Centres of Research Excellence and potentially for the National Science Challenges.

Is the process biased against multi-disciplinary proposals?

Analysis of applications from 2008-2010 suggests that multi-panel proposals (as a proxy for multidisciplinaryity) had a higher chance of being funded than proposals that were assessed by a single panel. Proposals are now encouraged to choose a primary assessment panel, because of the assessment work load involved of being considered by multiple panels. Only two proposals had to be assessed by more than one panel this year, with the majority of proposals matched to one panel.

Is the process biased against “intellectually edgy” proposals?

While it is difficult to predict which proposals will turn out to be ground-breaking, one possible proxy is that the most challenging proposals will be more difficult to assess and their value will be less clear to one or more researchers. Thus the variability of reviewer scores could be used as a proxy for edginess. If this is the case, then the Marsden process does penalise such work. The low success rate combined with the capped scoring used by reviewers to assess proposals means that a proposal needs to be scored at the maximum by all reviewers to have a high chance of being funded. If one reviewer misunderstands or fails to see the value of a proposal and marks it at less than the maximum score, then that proposal is much less likely to be funded. Thus the Marsden Fund process could be criticised for rewarding risk aversion.

The value of Marsden assessments

The benefit of the Marsden Fund is not just the funding given out. The assessment process delivers information about the quality of research proposals, and the independence and credibility of these assessments raises the value of that information to other decision-makers. Thus Marsden scores and entry into the second round are used to award internal grants by several research organisations. Additional internal funding is provided to applicants who just fall short of obtaining a Marsden to improve their chances in the next year. Success in gaining Marsden funding is also used as a measure of merit in PBRF assessments.

Schemes for Emerging Researchers – Comparing ideas with people

The Rutherford Discovery Fellowships are an example of a funding process that is focused on individuals rather than on ideas. Both Marsden Fast-Start and the Rutherford Discovery Fellowships (RDF) focus on early-stage researchers but Fast-Start assesses on the basis of the idea, whereas RDF assesses primarily the individual.

Both schemes use similar processes with some differences: Fast Start uses a blind peer review of proposals and assesses solely on research merit, whereas RDF considers personal qualities such as leadership potential and uses an interview-based selection stage. RDF provides five years of full time research funding, whereas Fast Start provides less support for a three year period.

There is some correlation between individuals who are awarded the Fast Start grants and RDFs but given the limited numbers, this connection may not be significant.

Rutherford Discovery Fellowships – Looking for future leaders

Applicants to the RDF must be within 3-8 years from their Ph.D. (with allowances made for parental leave). The Fellowship funds ten fellows per year for five years with a success rate over the past two years of 9%. A two-stage process is used. In the first stage, applicants and personal reports from three applicant-chosen referees are considered by discipline-based panels. These first panels assess on the basis of 60% research ability, 20% demonstrated leadership skills, and 20% project quality. Twenty successful applicants go through to a second stage of interviews with a selection panel, where panel members are appointed by the Royal Society of New Zealand. Selection panellists repeatedly comment on the value of interviews for determining human potential.

The Rutherford Discovery Fellowships provide an illustrative example of how unintentional or unnoticed biases can creep in and become embedded in funding systems. The scheme initially set an eligibility criterion that applicants must be within a certain number of years after finishing their doctorate. However, this blanket rule ignores the reality that many women take time out from research in this period of their career to start families. Simply changing the criteria from “a certain number of years after doctorate” to “a certain number of years of active research after doctorate” removed this unintended bias.

Total costs and benefits of applications to the Rutherford Discovery Fellowship

The Rutherford Discovery Fellowship disburses nearly \$10 million in funding per annual application round. The Royal Society of New Zealand’s best estimate³ of the cost to the sector is approximately \$1.1 million. Approximately 52% of the cost falls on the applicants as they write proposals, 25% is covered by the Society as contract and process manager, and 22% accounts for the time spent by peers involved in the referee reports, panel mediation and interview selection panels.

Applicants selected for interview spent an average of 44 hours on each application, while unsuccessful applicants spent longer, 68 hours. The degree of variability in these numbers is high but the difference between those selected and those not is significant.

³ These figures are based on data from the 2010 survey of applicants and account for researcher time using the 2012 Universities New Zealand report 'University Staff Academic Salaries and Remuneration'. Such calculations inevitably require numerous assumptions, including that referees and panellists are professors and that each spends two hours in assessing applicants.

Rutherford Discovery Fellowships provide examples of the value of failed applications to unsuccessful applicants

One effect of the competitive peer review process may be to raise the quality of research within that discipline, even for research that is unsuccessful at gaining funding. Failed applications have a known cost in the time taken to prepare them, but all applications have a benefit through forcing applicants to attempt to write high quality proposals. Anecdotal examples of this benefit can be found in the Rutherford Discovery Fellowships. One failed applicant considered that the time spent writing their unsuccessful proposal led to gaining funding for that research from a European source. Another stated that the collaboration and network forming during the application process led to a job opportunity with a Crown Research Institute.

Input measures versus output measures for assessing personal track record

Rutherford Discovery Fellowship decisions are made on the history of the person, as a proxy for the expected outcome of their project and future career. This approach to assessing tangible history faces the difficulties of accounting for non-academic track record and the effect of different personal funding histories.

One concern that has been highlighted by these Fellowships is the feed-forward nature of academic track records.

RDF applicants can have three to eight years of post-doctoral research experience. To some extent, the RDF intends to account for differing histories by attempting to be fair to researchers with three or eight years of experience, by clearly presenting to panellists the relative number of years of research experience during the panel discussions and interviews. However, beyond simply the time spent during research, an individual's track record depends upon both the research output of that person and their history of obtaining grants. In many processes such as RDF and Fast Start, historical grant awards to a person are considered as positive evidence of their inherent quality. Certainly, success in gaining grants increases the output of a researcher as it increases their access to material resources, allows them to attract the attention of other researchers and organisational support, and frees them to focus their time on research, rather than applying for grants. However, that success says nothing about the outputs that would have occurred had those resources gone to a better researcher—essentially, any measure of success in obtaining previous grants is a measure of inputs, not outputs. Stated in a different way, if a researcher at MIT and a researcher at smaller, less well resourced, and connected university produce papers of equal quality, which one is likely to be a better researcher? And which will be marginally more productive if given the next grant?

If intrinsic researcher quality is a key criterion for a selection process, then arguably a researcher's track record of outputs should be compared with their track record of inputs. The best researchers should be able to create more outputs from a given set of inputs, thus higher inputs should be scored negatively.

If we accept that previous funding decisions are perfect, then current funding decisions can piggy-back upon the information that previous decisions have created. However, we know that there is a degree of randomness even in the best decision processes. Given the degree of randomness and bias inherent in any research funding decision, it does not follow that a researcher who has been successful in applying for grants is intrinsically a better researcher than one who historically has not. If grants are awarded to a person on the basis that the person has successfully won grants in the past, then this simply ensures that to those who have, more is given and the potential of equally capable researchers with a worse track record is wasted, both through their reduced access to resources and their discouragement.⁴

⁴ Melin *et al*, "[The top eight percent: development of approved and rejected applicants for a prestigious grant in Sweden](#)", 2006

Thus any move to increase the role of track records of individuals in research funding decisions must take care to assess on the basis of measures of research output, not resource input.

Bias in peer review – The Matthew effect, gender, and nepotism

The Matthew effect concerns the preferential accumulation of resources to those members of a group who have historically accumulated resources (named from the Gospel of Matthew: “For unto every one that hath, shall be given”). This effect is an identified feature associated with reviewer bias where a track record of strong grant performance, and therefore research performance, begets further strong grant performance in evaluation of rounds.⁵

Several studies have shown prevalent gender biases in peer-review with results finding a 7% application success penalty based on gender and significant differences in competence scores assigned to male and female groups.^{6,7}

A further bias is nepotistic selection based on reviewer knowledge of a collaborating or affiliated applicant. This bias has been found to result in an approximate 10% improvement of competence scores under some panel environments.⁷

The blinding of applications to referees is a popular method of investigating reviewer bias although difficult to achieve in a small nation. Despite this, it is being attempted in New Zealand with the Health Research Council’s new Explorer Grants.

⁵ Merton, “[The Matthew Effect in Science, II – Cumulative Advantage and the Symbolism of Intellectual Property](#)”, 1988

⁶ Bornmann *et al*, “[Gender differences in grant peer review: A meta-analysis](#)”, 2007

Moss-Racusin *et al*, “[Science faculty’s subtle gender biases favor male students](#)”, 2012

⁷ Wenneras, “[Nepotism and sexism in peer-review](#)”, 1997

International travel grants – Research organisations outsource emotionally challenging decisions

The Royal Society of New Zealand administers a set of international travel grants for the New Zealand Government. The size of the grants is small, generally \$3000 to \$5000 per grant, with approximately 80 grants given per year. The process used is a low overhead one where three non-expert peer reviewers are asked to review grant applications as they come in. From an administration point of view, a key driver is to keep the evaluation costs of these grants low so that they preserve value. As such, the peer review process used for international travel grants is very different to that which is used within other areas of the Society's fund management. The reviewers are not asked to select on the basis of excellence, merely, to confirm that the research is bona fide. Such small grants need low transaction costs, so minimal applications are used to reduce the cost to the applicant.

However, there is a degree of fixed cost in any application process, so the application effort is inevitably high when compared with the size of the reward and in some cases, the administrative costs to the applicant and their organisation may be equal to or greater than the monetary value of the award. Despite the poor returns, these small funds still receive a substantial number of applications. Research institutions support and often require their staff to apply for such travel grants over an internal method of university funded travel grants. Clearly, both researchers and research managers value the schemes highly with the outcomes of these schemes having been reported as very beneficial by recipients.⁸ The overall return on travel grants is stated to be high, providing monetary returns through gaining overseas funding and providing prestigious international collaborations. However, these returns are a result of overseas travel, which could equally be funded by the research organisation at lower administrative cost than applications to the Society. So why do research organisations require their staff to apply to external sources when the relative cost of application is so high?

Travel grants may be seen by employees as discretionary perks from organisational management and thus could be a source of contention within the organisation. Outsourcing such internally challenging decisions on who is to benefit from travel funding may be creating value to the research organisation by removing the onus of selection from the institutions and placing it with the Society and the Ministry of Business, Innovation and Enterprise. The peer review provided by the Society provides high-quality, independent decisions that are uniformly accepted and that are used by other organisations to validate their own decision-making.

⁸ Gush, "[Performance of the ISAT Linkage Fund Scheme: A survey of 2001 to 2003 recipients](#)", 2005

In conclusion:

Based on the Royal Society of New Zealand's experience of administering a wide range of funds over the past decade and more, it is clear that peer review can be efficient and effective in a small country. It remains the best tool for striking a balance between, for instance, excellence and relevance in research. We recommend that:

1. Peer review processes should be matched to the particular features of funding schemes. There is no singular and perfect model for how peer review fits into funding decisions.
2. Peer review processes should be evaluated to understand their impact and adapted to deliver the best possible outcomes. Biases of gender, reward, and nepotism must be monitored and resisted.
3. Low success rates discourage risky proposals. Specific tools should be used to support the riskiest research.
4. In the overall context of a wider research funding system, a balance must be struck between building people and supporting ideas. For funding schemes that aim to develop science leaders, the potential of researchers should be based on measures of outputs and opportunity, not measures of inputs.
5. Peer review creates value beyond direct funding decisions. If it is carried out credibly, the information created has wide value as external validation, as feedback to researchers and their employers, and as a source of prestige. The role of this alternative value should not be discounted.

Except for figures & the RSNZ logo, Royal Society of New Zealand papers are licensed under a Creative Commons 3.0 New Zealand Licence.

