OFFICE OF THE PRIME MINISTER'S SCIENCE ADVISORY COMMITTEE

Which science to fund: time to review peer review?

Sir Peter Gluckman KNZM FMedSci FRSNZ FRS
Chief Science Advisor to the Prime Minister

December 2012

Office of the Prime Minister's Science Advisory Committee
PO Box 108-117, Symonds Street, Auckland 1150, New Zealand
Telephone: +64 9 923 1788
Website: www.pmcsa.org.nz
Email: csa@pmcsa.org.nz

1. Preamble

Publically funded science systems around the world are undergoing considerable change. In part this change reflects society's wish, manifest through the political process, to see an ever-greater utilitarian role for science. The current systems largely evolved after the Second World War when, while the utilitarian purpose of science was understood, it was not necessarily seen as the dominant basis of funding allocations. More recently, however, the arguments that science can further economic development, assist social and environmental enhancement and contribute to policy formation have been advanced eloquently and effectively by the science community and accepted by the policy community. As the scientific enterprise has expanded and the demand on the taxpayer's dollar also expanded, it is perhaps inevitable that the utilitarian purpose of public science is now expected to be transparently clear. The difficulty is however that the ways in which science impacts on society and the economy are not always direct, but those that are less direct may be no less important for a society. Elsewhere I have written about the multiple purposes of research[1] – the challenge is to find ways to measure and explain these broader impacts.

A major factor must be the increased demands on the public purse: in part this reflects the massive increase in the scientific enterprise, partially driven by the massive expansion of tertiary education. Further, the costs of much research have risen rapidly as a result of technological advances. Other factors are also important: there has been an effect on policy settings arising from the changing nature of science from physical to biological, from reductionist to holistic, and from simple to ever-more complex. Science is increasingly dealing with issues where science and values can be conflated (post-normal science) and must reflect the changing nature of scientific communication and publication, from authoritarian to interactive.

These issues are leading science policy advisors around the world to reflect on how the science system should evolve. These issues are much more acute in a small country, which by definition is more limited in absolute terms in its research spend; yet in small countries the capacities for change are higher and the consequence of poor processes are greater. Concomitantly, pressures on the funding system are more intense because there remains the desire, if not the need, to have a broad range of research endeavours across most domains of potential intellectual enquiry.

Dame Bridget Ogilvie, former director of the Wellcome Trust, on receiving an honorary doctorate at the University of Auckland in 1998 said in her speech "second rate research is a waste of money"; even in applied research this is a maxim that should not be forgotten. The challenge, particularly in discovery science, is that we need to find ways to encourage rather than discourage intellectually edgy research – this is where the leading edge is, where true innovation flourishes, and where high impacts are obtained over the long term.

An area that must and is receiving increasing attention is the process associated with decisions on which researchers and what research projects should be publically funded. Some countries have embarked on major exercises to decide national science priorities – effectively these are already made at a high level when bulk allocations are made to funding bodies. All countries have implied or explicit priorities and these become more important the smaller the science system, although history shows that the most impactful research discoveries are unexpected. Both New Zealand and Australia are in the early stages of reflecting on priorities for the public science system. This discussion paper does not directly focus on that question, but rather on the even more complex and sensitive issue of how funding should be allocated within a contestable system.

There is a range of processes that can be used to make research funding decisions; equally there is a range of funding scales – from a travel grant of a few thousand dollars to a research platform of tens of millions. It has generally been accepted that peer review is a core element of funding allocation in science and that the so-called "Haldane principle" whereby scientists should assess science excellence must operate[2]. That however does not

---

[1] Gluckman PD. Setting priorities for science, 28 August 2012. http://www.pmcsa.org.nz/wp-content/uploads/12-08-28-Setting-Priorities-for-Science.pdf

[2] For example, see section 4 of: Putting Science and Engineering at the Heart of Government Policy. http://www.publications.parliament.uk/pa/cm200809/cmselect/cmdius/168/16807.htm.

mean that society cannot set the priorities, and indeed Governments do establish science priorities through a number of mechanisms. While scientists pride themselves on objectivity, there is surprisingly little in the way of objective assessment of the nature and quality of peer review processes for grant allocation. This is in contrast to peer review processes for publications, where there is a larger literature and where there is also much instability. And yet what literature there is highlights growing concerns about how grant peer review is undertaken. Biomedical research and funding practices have been most intensively studied, but the findings are likely to be applicable to other domains of science.

The greater the competition, the more these issues come into focus. In New Zealand our three major contestable systems[3] all struggle with low success rates – the current success rate for the Marsden Fund is about 7% and for the HRC about 12%. Because the most innovative research tends to involve intellectual risk and thus can invite criticism, it is generally accepted that the general processes of grant awarding bias decisions towards conservatism and are in contradiction to the need of the nation for science to contribute to addressing cultural (in the academic sense), social, environmental and economic goals.

2. What do we know about peer review?

The limited literature suggests considerable discomfort with the peer review system as it generally operates, both from the point of view of the burden of cost and from the perspective of delivering the best outcomes in terms of successful awards, given that any process must be inherently somewhat subjective. Yet the fundamental argument that science is best evaluated by peer review by scientists remains generally accepted. The issue remains how to achieve the best from the system.

Commentaries and papers have long recognized that the issues relating to peer assessment compound in small countries[4]. This is further aggravated in New Zealand by our historically relatively low investment in public R&D. As the tertiary education system has expanded and become performance focused, the expectation for all academics to be research active[5] has flow-on demand effects.

Just because a country is small does not mean that it is not obligated to a broad range of research activities, and the cost structure of research may in fact be greater because infrastructure cannot be as broadly shared. But beyond the reality of more intense competition arising from these factors and others, the small size of a country's scientific community creates particular challenges for the peer review process.

A recent RAND report[6] summarises many of the criticisms that are globally made about the use of peer review in funding processes although its report focuses largely on much more extensive systems. The RAND report breaks questions about peer review into those around efficiency and those around effectiveness. The first category covers questions of the transaction costs and the burden of the process; the second covers questions of reliability, fairness, accountability, fit with strategic goals, and whether the best research is funded.

3. The high burden of assessment

The burden of the peer review system for funding allocation is rarely measured accurately or reported. A recent estimate from the Australian National Medical and Health Research Council[7] shows the costs are a major burden on the science system. A report prepared by the Royal Society of New Zealand to assist this paper's development[8] presents an estimate for the New Zealand Marsden Fund

---

[3] The Marsden Fund, the Health Research Council, and the contestable fund administered by MBIE.

[4] For example, Pouris A, 1988. Peer review in scientifically small countries. R&D Management 18:333-340. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9310.1988.tb00608.x/abstract.

[5] There are other models – for example, a number of high quality liberal studies universities in the USA do not teach beyond masters level and staff are scholarship informed.

[6] RAND Corporation, 2009. Evaluating grant peer review in the health sciences. A review of the literature. http://www.rand.org/pubs/technical_reports/TR742.html.

[7] Graves A et al., 2011. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. BMJ 343:d4797. http://www.bmj.com/content/343/bmj.d4797.

[8] Weston J, Gush J, 2012. Evidence from ten years of research contract management. Wellington: Royal Society of New Zealand. http://www.royalsociety.org.nz/publications/policy.

(disbursable NZ$55 million per annum): the direct administrative costs for operating the grant process are under 3% of the fund size, but this accounting ignores the vast majority of costs to applicants, referees, and panellists. These can only be estimated, but they are substantial. The best estimate puts the total cost at 20-35% of the fund size, some NZ$10-20 million. The majority of the cost falls onto applicants. Estimates suggest that the time spent writing proposals represents over 80% of the total fund cost, with three-quarters of that spent on first stage proposals. International reviewers and panellists make up 10% of the total cost and this is a significant burden upon the small number of people who are called upon in these roles.

For essentially all funding schemes, the major cost is in proposal writing. For unsuccessful applicants, this may not be time that is entirely wasted – there are clear benefits from researching and clarifying ideas, building networks, and the possibility to use those applications for accessing alternative funding – but nevertheless in a small science system it has a major inhibitory effect on research outputs. This is aggravated in New Zealand by the relatively short-term nature of most funding systems, the tendency to underfund requiring multiple sources of support, and the long cycle of assessment: this means that many research active staff are in a constant cycle of either writing grants or assisting others to write grants.

The volume of grants that senior referees are expected to examine means that they increasingly avoid participating in the process[9]. This is a feature increasingly noted in small countries. It is unsurprising that the system is fragile given the variable requirements of review, the sheer amount of reviewing required, the reality that it has often been expected to be conducted over the holiday season, and the facts that institutionally it is unrecognized and, in New Zealand, unpaid. These issues are not trivial and there are increasing signs of senior scientists boycotting requests to participate.

Ironically it is not the assessment and scoring of a grant that takes the time, it is the justification and handling of good, mediocre or particularly adverse scores that takes enormous effort, especially as the reviewer is addressing both the applicants and a committee that has to integrate such scores and comments across a wider pool. This will be discussed further below in considering the question of whether the system is being compromised by the burden being placed on the external referees and committee reviewers to deal with providing commentary on a grant. Anecdotally, it appears that simple but positive reviews are often discounted as if the reviewer has not been serious in his/her evaluation. Conversely, simple but negative reviews carry extra weight in tight funding systems with low success rates.

4. Process biases – subjectivity and biases of panels and reviewers

A recent paper on the science of peer review that models the outcome of a NMHRC round in Australia[10] highlights the inherent subjectivity of the standard process – it observed that the most widely used peer review process (consisting of an evaluation committee with a spokesperson leading an open discussion, informed by external referee reports, followed by scoring) is a very subjective process which is both “costly and somewhat random”. The study concluded that most funding decisions are a result of random effects dominated by factors such as who was the lead reviewer. In general the referee and panel review process is considered problematic[11]. Few scientists are trained to fulfil such roles and bad peer review must result in unfair outcomes.

It seems self evident that such effects will be even more compromising in small countries where the potential for conscious or unconscious bias, whether positive or negative, by the lead reviewer, other domestic members of the panel or referees is greater[12]. Such issues are potentially exacerbated by the

[9] Schroter S et al., 2010. Surveys of current status in biomedical science grant review: funding organisations' and grant reviewers' perspectives. BMC Med 8:62. http://www.biomedcentral.com/1741-7015/8/62.

[10] Graves A et al., 2011. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. BMJ 343:d4797. http://www.bmj.com/content/343/bmj.d4797.

[11] Schroter S et al., 2010. Surveys of current status in biomedical science grant review: funding organisations' and grant reviewers' perspectives. BMC Med 8:62. http://www.biomedcentral.com/1741-7015/8/62.

[12] Pouris A, 1988. Peer review in scientifically small countries. R&D Management 18:333-340. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9310.1988.tb00608.x/abstract.

excessive workloads of expert senior reviewers and a concomitant drift to less experienced researchers as panellists. The latter tend to focus on detail rather than capability, potential and the strategic fit of the science. Experience and seniority have real value in such evaluations.

5. Panel processes

One of the most problematic issues arises because of the chance effects arising from the general and most common model whereby the awarding committee starts with the nominated in-depth reviewer(s) presenting his/her views and then effectively steering discussion, consciously or not, in a certain direction. Studies have demonstrated that such an approach creates enormous variation in outcome depending on which reviewer is appointed to this role. Studies have shown that greater consistency was reached when 10 reviewers independently scored a grant without any discussion[13]. A recent study from Finland[14] shows that discussion does not improve the consistency of panel scores in any way.

The issue of panellist bias is greatest where grants are highly innovative – scores on such applications are likely to be controversial, as the best think outside the box and challenge the orthodoxy. Where success rates are low, it is most probable that there will be more highly rated grants than funds available and thus the impact of bias, marginal negativity or controversy can be a very strong factor in influencing the success of a proposal. The result is that the more innovative and edgy ideas often get disparate scores and are therefore unlikely to be funded. This would appear more likely when the panel is inexperienced and therefore cautious, while conservative research with less impact becomes the norm – even within the lottery of the process. The psychology of the panel also becomes important – adopting a nit-picking negativity can reduce the burden of ranking on panel members when they know they face almost impossible decisions.

6. Conflicts of interest

The integrity of scientific peer review relies on the avoidance of conflicts of interest. In a small science community, conflicts, whether real or perceived, declared or undeclared, create major problems for the granting system. In any one field there are relatively few experts and they are most likely either working collaboratively or are actually competing for the same pool of funds. Thus to fund individual A makes funding of B, even in the next round, less likely. Beyond that, in a small science community personality and extraneous information can easily influence a reviewer or referee, often unconsciously. Similarly, even where private sector interests are involved, the potential exists for other considerations to come into play. This problem of potential bias has led some small countries such as Ireland and Israel to use exclusively extra-jurisdictional referees and, in the case of Ireland, only international panel members, relying on officials to provide local context where appropriate.

7. Multidisciplinary and interdisciplinary research

A further concern is the problem of how to assess multidisciplinary and interdisciplinary research. As multidisciplinary research is often in areas where innovation is particularly likely to arise, then a system to support such research must be a priority. Indeed the need to promote interdisciplinary research is seen as a priority in every research jurisdiction and was strongly endorsed at the recent Transit of Venus forum[15] in Gisborne in 2012. Joint agency funded research faces similar issues unless the assessment is devolved to only one agency. By definition, if a proposal goes to two panels it faces greater jeopardy, and as it may be less directly central to one or other of the panels it may be scored down. However this may depend on whether the panel has a strategic focus or is simply concerned with excellence. The former is more likely to downgrade a grant that crosses boundaries, whereas

[13] Mayo NE et al, 2006. Peering at peer review revealed high degree of chance associated with funding of grant applications. Clin Epidemiol 59:842-8. http://www.sciencedirect.com/science/article/pii/S0895435606000059.

[14] Fogelholm M et al., 2012. Panel discussion does not improve reliability of peer review for medical research grant proposals. Clin Epidemiol 65:47-52. http://www.sciencedirect.com/science/article/pii/S089543561100148X.

[15] Gluckman PD, 2012. Science and New Zealand's future: reflections from the Transit of Venus forum. http://www.pmcsa.org.nz/wp-content/uploads/Transit-of-Venus-Forum-report.pdf.

evidence from the Marsden Fund suggests that multi-panel proposals have been more likely to be funded than those assessed by a single discipline-based panel.

Concerningly, such research was effectively discouraged by the rules of the former Foundation for Research, Science and Technology, which required that grants only went to one panel. Further, the portfolio approach whereby certain fields of research are only available for contestation every three years coupled with the small size of each contestable 'pot' is also actively inhibitory to such research and is in contradistinction to funding systems in all other comparable jurisdictions.

8. Rebalancing people versus projects, relationships versus transactions

A primary issue in any science funding system is the relative emphasis placed on individuals versus ideas. How this is framed can lead to very different outcomes. There should be clarity with all stakeholders as to the extent to which the key points of assessment are about the individual and his/her track record and potential or about the project idea *per se*. In New Zealand we have tended to focus primarily on the project, particularly in the major contestable pool operated by the Ministry/Foundation, but science systems around the world are beginning to give much greater focus to the quality of the applicant(s) and their teams – recognizing that the most 'intellectually entrepreneurial' scientists need to be fostered, thereby allowing teams to be built around them; after all, science is a creative human endeavour. While such an approach is not egalitarian (and grant funding cannot be), there is evidence that track record put in perspective remains a better predictor of performance than anything else[16]. However this will depend on what is meant by 'track record' in the context of the intent to support innovative and impactful research. This focus has to be accompanied by an overt way of identifying and seeding emerging talent[17].

[16] Abrams PA, 1991. The predictive ability of peer review of grant proposals: the case of ecology and the US National Science Foundation. Soc Stud Sci 21: 111-132. http://sss.sagepub.com/content/21/1/111.short.

[17] Graves A et al., 2011. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. BMJ 343: d4797. http://www.bmj.com/content/343/bmj.d4797.

Such an approach must recognize the increasing degree to which research capability resides in teams, especially with the increasing importance of multidisciplinary work in addressing the multifaceted problems that society faces. It is rare for a single project to lead to an innovative breakthrough. Rather, there is a need for a team to compile and build its research over time and the assessment process therefore needs not to look at the project in isolation but in terms of both its trajectory and the performance of those doing it. Undertaking impactful research is a profession which while generally carried out within an academic environment has attributes that mean that some people will be more productive than others. Careers cannot be sustained if such individuals and their teams cannot have a realistic expectation of grant renewal, provided that the progress and potential or actual impact of the research justifies it.

9. National priorities – what level of granularity?

Further issues emerge when factors such as national priorities and other elements of relevance are inserted, as they increasingly will be, into the science review process. Given that the research is publically funded, the taxpayer has the right through the policy/political process to expect impact, which often implies applied relevance. But depending on the type of research, how that is assessed again requires care.

The essential tension here revolves around the granularity of priority settings – to what extent should it occur at national level with the goal of meeting national priorities, at a sector level to meet the needs of industry, at a research organization level with the goal of meeting the strategic plans of those organizations, or at a team level with the aim of building specific research capabilities? One of the tensions that emerges is that of timeframes – industry tends to look for shorter term returns than much research is designed to deliver and developing a balanced portfolio for both shorter and longer term needs is complex. There is value to be gained by making decisions at each level; equally, there are risks by taking an overly prescriptive stance at any level. There are no clear answers here, but this tension should always be borne in mind when considering how the different levels of priorities should feed into proposal assessment criteria.

10. Quality versus relevance

Problems emerge when assessments of science and impact are combined in a single panel or score. The influence of one on the other means that the criteria used for assessing excellence are often lost or obscured. It is particularly problematic when non-scientists and scientists are on the same panel. Increasingly, agencies in other jurisdictions (for example Ireland) are separating entirely the assessment of these two domains. Depending on the skill sets of those assessing impact, short-term goals will be assessed differently to longer-term research. In general assessment of impact/relevance of the proposal and appraisal of the quality of research/researcher require separate skill sets and perspectives. New metrics and criteria may be needed to assess potential and actual impact, an approach several jurisdictions are exploring.

This raises the essential tension of how should the two criteria of research quality and impact/relevance be combined? Should assessment follow the pattern of first assessing for excellence and then filtering for fit with priorities, or should relevance be the first filter? The former makes it more likely that the best research is funded and more likely that the most innovative research will be considered. Putting the relevance criterion first may well block important and potentially transformational research coming forward and risks the funding of research of lower quality which, while apparently relevant, will likely drive the research endeavour over time towards mediocrity; it is counterproductive over the long-term to invest in such research. If the field is important then the policy response should be to seek ways to enhance the quality of the science coming forward. Using the filter of excellence as the first criterion is more likely to be of benefit to both the funder and society and to generate impactful outcomes – otherwise the danger of Dame Bridget's warning is compounded .

11. Tools of funding

One trend that will have an indirect effect on the issues under discussion is that of the increasing funding of larger and longer research programmes. This is a growing feature of the domestic research system, with funding for coordinated programmes such as the Centres of Research Excellence, the Natural Hazards Research Platform and the Sustainable Land Use Research Initiative. Such larger grants generally have longer funding, are multidisciplinary and are built around teams with track records. They can also be matched to national science priorities. By definition such grants tend to have a strategic intent. International evidence from other small countries (e.g. Denmark) suggests that such programmes contribute disproportionately to impactful research as judged both academically and from the policy/business perspective.

12. Addressing the challenges

So how should a small country address such challenges? Firstly, the policy and scientific communities need to agree on what they want to achieve in the grant allocation process. For example, if the system is really being hurt by the time spent by the research community in the application process, then the decision needs to be made as to whether the grant allocation system is primarily about allocation of funds or the constructive development of scientists. It is argued by some that the grant review and rebuttal process is an important part of science career development, although there is little evidence that the review process is particularly educative; the main role of extensive feedback appears to be to create a kind of transparency. Arguably, the educative role is very much a secondary purpose that could be achieved in other ways.

Secondly, it must be decided where and when in the allocation system are people or projects the core determinants of the outcomes being sought.

Thirdly, we also need to decide at what level of granularity priorities are set, and how? The smaller the pots of money, the greater the problem of granularity and the greater the risk of innovative research not being funded.

13. Reducing overall costs by reducing applications

Many of the suggestions commonly put forward for reducing the demand for funding are irrelevant or ineffective in New Zealand. The use of single annual (or less frequent than annual) deadlines is common practice, but the more infrequent the application round the more likely a timely and innovative idea is to die without ever being considered. Limiting the number of applications from particular institutions is a blunt tool that may have a strong effect upon the number of applications, but only by implicitly requiring research organisations to carry out their own culling of potential applications, forcing funding decisions to occur at an additional level, where other intra-institutional factors are bound to be in play, and doubling up on such decision-making.

The major system-wide cost of the funding system is the time spent by applicants in putting forward proposals. Reducing this would allow more researcher time to be spent in research. Equally, reducing the number of applications to review and the time spent per review would reduce the burden on the senior staff who act as reviewers.

Success rates do not appear to act as a self-limiting disincentive for many potential applicants. For prestigious funds such as the Marsden, as for journals such as *Nature*, low success rates act to increase the prestige of the award. Equally, for many kinds of research, there are no other possible sources of contestable funding, so demand for the fund is inelastic and application numbers will remain high and somewhat proportionate to the size of the academic community.

The up-front effort required to pursue a funding application and the number of applications that a fund receives have a complex relationship. But given the major cost associated with an extended application, there is a growing trend towards two-stage applications. But this too can generate further potential risk depending how each stage is conducted. The caveats around panel processes discussed above apply equally at both stages. Given the relatively superficial information available to panel members at the first stage, the potential for bias regarding project or person or domain to creep in is real and is even more difficult to manage – again, to avoid this problem some small jurisdictions undertake grant triage at this stage using only international assessors.

From a wider perspective, the high number of good applications shows the latent potential within the national innovation system The cost of limiting this to reduce transaction costs will be counted in the lost opportunities for benefit. The loss of these opportunities will be hard to measure. The gains in transaction cost saving will be more concrete, but it does not follow that savings in transaction costs will represent overall savings for the science system or for the nation.

14. Changes to review processes and reviewers

Clearly reducing the effort required to assess and review each application should have a clear and strong role in reducing the burden on senior staff who act as reviewers. Where applications look poor at first impression, then little time should be spent in justifying poor scores. Equally, where success rates allow, applications that score highly from all reviewers should be green-lit at the first possible opportunity

Other forms of peer review-based systems have emerged in part to address these issues. Blinded approaches make meaningless any focus on the individual and are difficult in small countries for obvious reasons. Some countries outsource the decisions entirely to a separate jurisdiction, but that too requires consideration of how to address priority and contextual dimensions.

The evidence suggests that increasing the number of expert panellists is more likely to increase objectivity, but only if the inherent biasing effect of the in-depth reviewer is removed. Inter-disciplinarity is addressed through increasing panel size and is improved by the participation of accomplished panellists.

One approach now being increasingly used by a number of bodies such as the European Molecular Biology Organization and the Human Frontiers Science Program relies on a statistical approach based on asking a broad range of senior and expert reviewers to independently rank all grants within the relevant pool as part of a large panel without having to provide any justification (other than perhaps minimal general feedback), which means their time and expertise is entirely devoted to assessment, not writing essays.

There is little substantive difference between ranking, scoring, and fine scoring systems provided that a full range of scores is used. More critical for obtaining a reliable assessment is the consistency of reviewers which comes from their expertise and from independent assessment, rather than panel-based discussions of the details of scores which inevitably lead to some dominance by some individuals over the scoring. Where the grants get consistently high scores or ranks across the entire panel then they are funded (subject to funding being available) without further discussion once the panel convenes (unless there are relevance matters that come into play). The reviewers thus do not spend time justifying/adjusting scores, their engagement and commitment is therefore much higher, and their time is far better focused and spent on grants at the funding margin. In many ways this is similar to a system developed at the McGill University Health Centre[18].

Within this context, experience suggests that a good reviewer generally knows within minutes what they are dealing with, provided that he or she is genuinely knowledgeable in the field (again, a challenge in a small science system). As an extension to the ranking system discussed above, it allows every panel member in a large panel to rank score all grants. It is only then that the panel enters into discussion about those cases where the scores are very discrepant with both very high and low rankings. The question then is whether or not a contentious application is the result of truly innovative and brave ideas, or whether it is really a flawed or a 'me-too' application. Such discrepancy can emerge because there are different perspectives and knowledge bases in the panel and the panel needs to talk these through to resolution.

15. Building people and teams

Partly because of the evidence that supports the central role of key intellectual innovators, funders such as the Wellcome Trust, NIH, and the Howard Hughes Fund have shifted their focus to identifying and supporting the outstanding individual. Such an approach does not require a step away from national priorities.

Recognising potential and building future research leaders requires funding schemes to focus on individuals rather than projects, and thus requires assessment of a new set of intangibles, of personal qualities as opposed to the potential outcomes of an idea. The role of peer review in this process remains widely accepted, with evidence showing reasonable predictive validity for committee peer review of an individual's track record[19]. What factors should be included in that track record is a matter of active research and debate , with numerous suggestions such as the Hirsch index[20], citation count, grant history, and so on. This is a question deserving careful consideration as all indicators have their problems. Many measures of outputs fail to take into account differences in inputs and these inputs can be quite random – a researcher who has been lucky in gaining research funding may have a higher research output than one who has not, despite the second researcher being more capable[21,22]. Where leadership is involved and long-term teams are being funded, despite the subjectivity, the interview is generally accepted as an essential and central part of the assessment process.

16. Points to consider

Research is a different kind of endeavour to most other businesses – by definition it cannot be purchased off-the-shelf from a supplier. Even assessing what research is needed is a complicated matter. Instead, building up the capability of that supplier to meet research needs is a long-term process where capabilities can be only be gained by support over a long period and can be lost overnight. Building leaders and teams requires building enduring relationships between research organizations and research funders.

[18] Mayo NE et al, 2006. Peering at peer review revealed high degree of chance associated with funding of grant applications. Clin Epidemiol 59:842-8. http://www.sciencedirect.com/science/article/pii/S0895435606000059.

[19] Bornmann L et al., 2005. Selection of research fellowship recipients by committee peer review. Reliability, fairness and predictive validity of Board of Trustees' decisions. Scientometrics 63: 297-320.

[20] Acuna DE et al., 2012. Future impact: predicting scientific success. Nature 489:201-202. http://www.nature.com/nature/journal/v489/n7415/full/489201a.html.

[21] Melin G et al., 2006. The top eight percent: development of approved and rejected applicants for a prestigious grant in Sweden. Sci Public Policy 33:702-712. http://spp.oxfordjournals.org/content/33/10/702.short.

[22] Mayo NE et al, 2006. Peering at peer review revealed high degree of chance associated with funding of grant applications. Clin Epidemiol 59:842-8. http://www.sciencedirect.com/science/article/pii/S0895435606000059.

The issues around the use of peer review are complex and there will be many divergent views. Statements of both of these kinds can be found: either that peer review is *"the most effective and respected way to assess the quality of research outputs"*[23] or that *"peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong."*[24]

There is good evidence to suggest that peer review processes are high in burden and less than ideal in outcome: no perfect process exists. In writing this paper, it is not my intent to suggest any particular solution but I do think that it is timely to have a more objective look at the process of funding decisions as this is the most important element in matching our research community to the changing shape of our innovation system. When all is said and done, funding decisions determine both careers and what science will contribute to our world.

Acknowledgements

This paper is intended to stimulate discussion and the views are personal. However the paper has been extensively assisted by Dr Alan Beedle and Dr Stephen Goldson from my office and by Dr Jez Weston and other staff from the Royal Society of New Zealand. The paper was also informed by discussions with my counterparts from Denmark, Finland, Ireland, Israel and Singapore at the Small Advanced Nations meeting in Auckland in November 2012. An early draft was sent to a range of scientists and research administrators and I am most grateful for their comments. In particular I want to acknowledge Professor Phil Baker (National Research Centre for Growth and Development), Mr Len Cook (Victoria University of Wellington), Dr Ian Ferguson (Plant and Food Research), Professor Mark Ferguson (Science Foundation Ireland), Professor Jane Harding (University of Auckland), Professor Harlene Hayne (University of Otago), Professor Peter Hunter (University of Auckland), Professor John Mattick (Garvan Institute of Medical Research, Sydney), Dr Warren McNabb (AgResearch), Dr Robin Olds (Health Research Council), Professor Hamish Spencer (University of Otago) and Dr Prue Williams (MBIE).

[23] The Royal Society of London, 2007. Response to the RCUK consultation on the efficiency and effectiveness of peer review. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/publications/2007/8090.pdf.

[24] Horton R, quoted in Taylor M et al., 2008. The siege of science. Ethics Sci Environ Polit 8: doi: 10.3354/esep00086. http://www.int-res.com/articles/esep2008/8/e008pp13.pdf.