



**MINISTRY OF BUSINESS,
INNOVATION & EMPLOYMENT**
HĪKINA WHAKATUTUKI



Excellence in Research

Published by Gary Evans, Rob Murdoch and Hamish Spencer
Science Advisors, Ministry of Business, Innovation & Employment

August 2021

Not Government Policy:

*Rather than being a formal synthesis of evidence, this document outlines the professional
opinions and advice of independent Science Advisors*

ISBN (online) 978 1 99 100862 6

Table of Contents

EXECUTIVE SUMMARY	1
INTRODUCTION	3
R&D Definition	4
NSSI Definition of Excellence	4
FUNDER IMPLEMENTATIONS OF EXCELLENCE	5
MBIE's Endeavour Fund Implementation	7
Marsden Fund Implementation	7
Health Research Council Implementation	8
PBRF DEFINITION	9
PBRF Review Panel Report	10
ROYAL SOCIETY TE APĀRANGI DEFINITION	11
CORE DEFINITION	12
A MĀORI PERSPECTIVE	13
INTERNATIONAL PERSPECTIVES	14
BIBLIOMETRICS & EXCELLENCE	15
DISCUSSION	17
ACKNOWLEDGEMENTS	20
APPENDIX 1 : DATA UNDERLYING FIG 1.	21

Executive Summary

This document discusses the concept of research excellence: what it is, how it can be assessed and why it matters. Our context for this discussion is the Research, Science and Innovation (RSI) sector of Aotearoa New Zealand (NZ). Consequently, we start with the approach used by the Ministry of Business, Innovation & Employment (MBIE), which administers Vote Business, Science and Innovation. MBIE uses the definition in the *National Statement of Science Investment: 2015-2025 (NSSI)*. We then discuss definitions and assessment in other New Zealand funding agents and organizations (e.g., the Health Research Council [HRC], Tertiary Education Commission's [TEC] Performance-based Research Fund [PBRF] evaluation), before examining what excellence means in Te Ao Māori as well as the international context. The use of bibliometrics is also examined.

We make 4 critical points:

1. The *NSSI* definition of research excellence used by MBIE differs from international practice (and is different from the TEC's approach) in excluding non-academic impact. Research excellence and impact (seen as the second pillar, alongside excellence, of a healthy RSI sector) are assessed separately. In most instances, this approach is unproblematic and practical, provided that impact is clearly defined in a non-overlapping manner. Nevertheless, excellence and impact are closely linked: in mission-led research, for instance, we want both excellence and impact and these are likely to be correlated.
2. Definitions of research excellence need to be broad, ensuring that all forms of research are treated fairly. For example, different fields of research as well as fundamental and applied (developmental) research must all be evaluated in an equitable manner, even if the specific criteria may be different. This requirement is particularly important for research involving and embedded in Te Ao Māori, given the need to significantly improve the record of the RSI sector in meeting the requirements and aspirations of Māori. We need a view of excellence (and knowledge systems) that recognizes what elements of such research are important to and prioritized by Māori.
3. Peer review remains international best-practice for assessing excellence, at least at the level of individual (people and project) assessments. Nevertheless, peer review is subjective and has potential biases, and peers must be appropriately selected (e.g., in having expertise and a diversity of backgrounds) and trained (e.g., to help minimize bias). The use of bibliometrics, while seemingly objective, is often problematic on a number of grounds.
4. It is essential to be clear about the context in which excellence is being assessed. For example, in making prospective decisions about competitive grant applications, we are looking forward in time, attempting to reward proposals that will produce excellent research, even if, in part, we use researchers' track records and other information from the past. In other cases, such as prizes, promotions and PBRF, the decisions are more clearly retrospective, rewarding research excellence in the past. Granularity matters, too: decisions about individual-level funding (e.g., grant applications) will conceivably use a different approach from those concerning institutional-level funding.

We therefore make the following recommendations:

1. Research funders should examine their assessment criteria for excellence to ensure they value the full diversity of ways of achieving excellence in research. Similarly, funders need to ensure this diversity is recognized in practice, for example, by using assessment panel members with an appropriate variety of research backgrounds.
2. Future definitions of excellence by research funders need to ensure they do not disadvantage or discriminate against Māori, gender, institutions, ethnic groups or non-traditional areas of study. In doing so, funders should draw on international efforts to address this issue.
3. Māori should continue to develop their own definitions and measures for research excellence to guide future assessments of excellence by research funders.
4. Although research excellence and impact are intertwined, the definitions of research excellence and impact should be kept separate. Different funding bodies should balance the weightings they assign to excellence and impact so that they are appropriate to their funds' missions.
5. Consideration should be given to providing a clearer definition of excellence in future drafts of the RS&I Strategy. This view should be high-level and not overly prescriptive, however, so that it can encompass a diversity of approaches to pursuing excellence.
6. Peer review should continue to be regarded as the gold standard for assessing research excellence. Nevertheless, funders should ensure that documented problems (e.g., biases) with peer review and other forms of assessment are acknowledged and addressed (e.g., through training of assessors).
7. Funders should ensure that support for research based on excellence does not disadvantage or reduce the methodologically excellent day-to-day research and infrastructure that underpins the RS&I sector.

Introduction

For over one hundred years, nations have formally recognized the societal value of government investment in knowledge generation through science and, more generally, research and development (R&D).¹ Moreover, because making appropriate decisions about such investments is so critical, choices about what R&D to support should transcend short-term political imperatives. Accordingly, in most parts of the western world, while broad areas of government investment in research and innovation are determined by politicians and civil servants, the details of research-funding decisions are kept out of the political arena and a set of widely supported policies and procedures have been developed to this end. For instance, under the Haldane Principle politicians do not compare the relative merits of individual R&D projects themselves.² Instead, in what is viewed internationally as best practice, independent funders use peer review and assessment committees, made up of researchers, to estimate the merit of the proposed research, and decisions are made on that basis. Critically, excellent research is usually considered to lead to greater societal impact.³ As such the concept of excellence is fundamental to research. Every researcher aspires to carry out excellent research,⁴ every funder aims to fund excellent research and every research institution hires staff who are directed to contribute to excellent research.⁵

Research that is considered excellent is rewarded in numerous ways, from the promotion of individuals within their institutions and the awarding by professional societies of prestigious prizes and medals, to greater institutional funding (from governments and other funders) to those institutions that carry out what is considered to be excellent research. All these rewards enhance the reputation of the researcher and their institution, which in turn can lead to further rewards (e.g., students choosing that institution and philanthropic donations).

But what exactly is excellence in R&D and research in particular? This opinion piece addresses this question, with an emphasis on the Aotearoa New Zealand context. We also make recommendations on what elements of a definition we see as desirable for the changing needs of our society.

A second question also arises: how can excellence be assessed? In some respects, this question asks how the definition of excellence can be operationalized, especially in decision making by funders. We also address this second question, giving examples of the assessment of excellence, again with an emphasis on the national context, including relevant aspects of Te Ao Māori.

¹ Gluckman, P.D. 2015. Why do governments support research? The evolving role of the State. Available at <https://www.pmcasa.org.nz/wp-content/uploads/Why-do-governments-support-research.pdf>

² For example, in the UK, the Haldane Principle means that “decisions on individual research proposals are best taken by researchers themselves through peer review ... [T]he awarding of grants to specific research activities should not be taken by Ministers or central government.” See Higher Education and Research Bill: UKRI Vision Factsheet, Principles & Governance. 2016. Page 5. Available from <http://www.legislation.gov.uk/ukpga/2017/29/section/103/enacted>

³ This assumption is backed up by a number of studies, e.g., Jonkers, K. & Sachwald, F. 2018. The dual impact of ‘excellent’ research on science and innovation: The case of Europe. *Science and Public Policy* 45: 159–174.

⁴ This is not to say that researchers value excellence over “making a difference” (impact).

⁵ Of course, not all research or researchers can be excellent, for that would make the extraordinary ordinary.

R&D Definition

In the first instance though, we must define what we mean by R&D. The Frascati manual⁶ measures the scientific, technological and innovation activities across the Organisation for Economic Co-operation and Development (OECD), of which NZ is a member. In part, the manual aims to harmonize the jargon and technical terms used across the OECD to describe R&D. The Frascati manual defines activities as qualifying to be classed as R&D where they are novel, creative, uncertain, systematic, and transferable/reproducible, and describes three types of R&D activity: basic research (“blue skies” research), applied research, and experimental development. Interestingly, the manual does not define research excellence; in fact, excellence is mentioned only once.

NSSI Definition of Excellence

The *National Statement of Science Investment: 2015-2025*⁷ (NSSI) sets out the current governmental long-term vision for the Research, Science & Innovation (RSI) sector and provides a strategic direction to guide future investment. Although focussing primarily on scientific research, many of its arguments apply more broadly to the RSI sector. In this document, excellence and impact are seen as the two main pillars supporting the RSI sector. Excellence is succinctly defined (page 6) as “well-designed, well-performed, well-reported research, recognised as such, e.g., through peer review.” By contrast, impact is “the direct and indirect ‘influence’ of research or its effect on an individual, a community, or society as a whole, including benefits to our economic, social, human and natural capital.”⁸ Note that, under these definitions, what is often called academic or scholarly impact is part of excellence, not impact.

These two pillars are scored separately in the assessment of many MBIE competitive-grant applications (e.g., Endeavour; see below). Excellence and impact are nevertheless inextricably intertwined, since “the quality of the science system and of the people who work within it is the key determinant of impact.”

The definition of excellence is unpacked on page 13 of the *NSSI*, where we are cautioned that

Excellence in science is not easily identified by one measure. It is a concept as applicable to projects and teams as it is to individuals, and is not static; it grows and changes as science does.

Excellent science (and, presumably, research more generally) involves “the best people, a rigorous approach and optimum results.”

The *NSSI* is due to be refreshed and replaced by a new RSI strategy. The draft of this strategy published by MBIE for the purposes of public consultation in September 2019⁹ described research excellence as “the ongoing pursuit of the best thing possible” (page 25). It also noted that excellence applied across research and innovation, but in different ways. The draft argued, too, that “excellence resists a single definition across the broad range of activity that makes up RSI.”

⁶ <https://www.oecd.org/sti/inno/frascati-manual.htm>

⁷ Available at <https://www.mbie.govt.nz/science-and-technology/science-and-innovation/funding-information-and-opportunities/national-statement-of-science-investment/>

⁸ Note that impact as defined this way is not the same as academic impact, which could be defined analogously as the influence of research on the relevant academic communities. Academic impact is hence part of excellence.

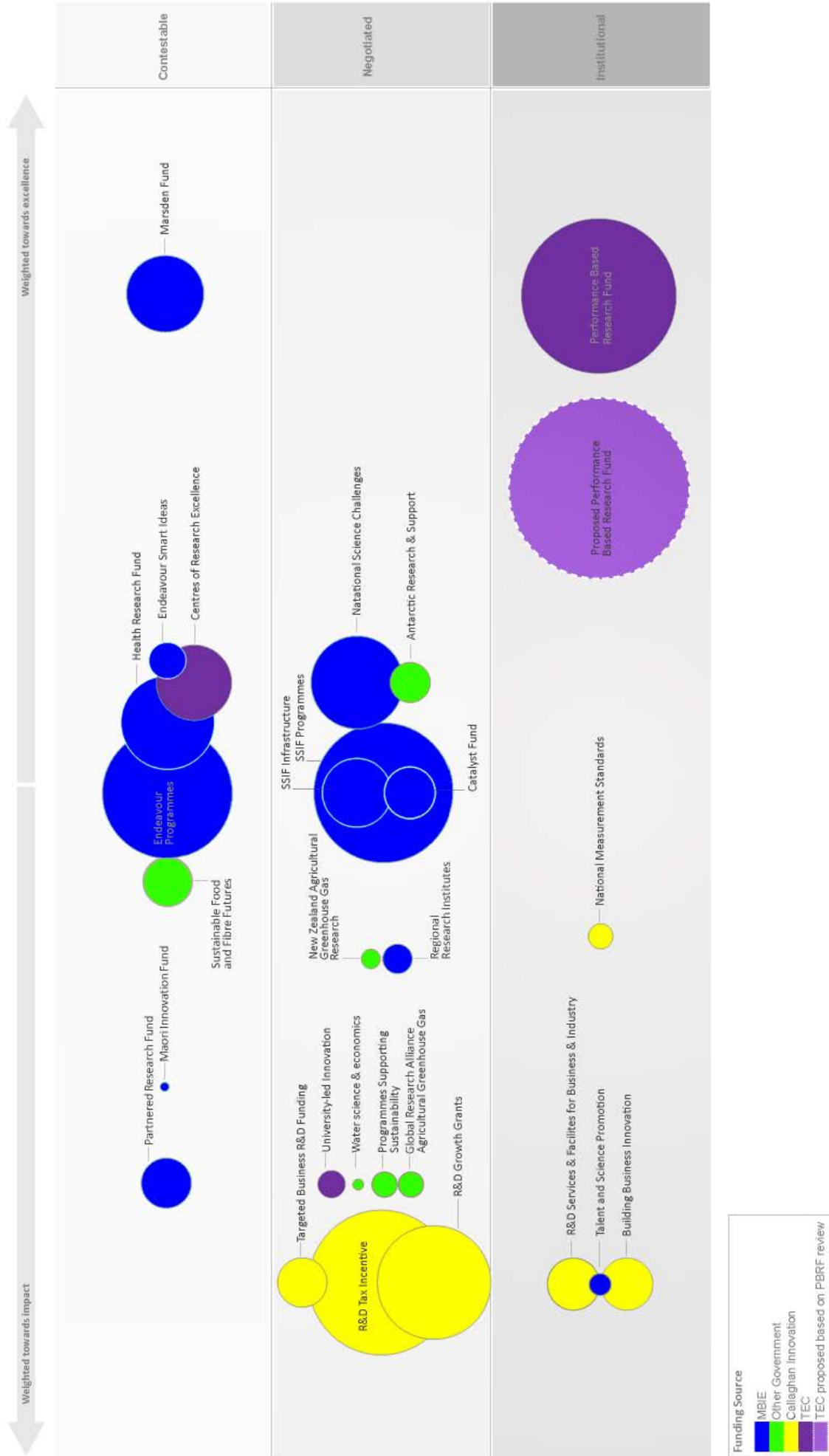
⁹ Available at <https://www.mbie.govt.nz/have-your-say/draft-research-science-and-innovation-strategy/>

Funder Implementations of Excellence

In deciding how to determine the best allocation of resources for research activities, funding organizations often set up criteria against which to score and thereby rank applications. In effect, these organizations operationalize their definitions. Note that these operational measures are prospective or forward-looking: they attempt to discern which proposals will lead to excellent research (among other things). They do so by evaluating the description of the future research, but also by looking backwards, for example, at the track record of the researchers on the application.

Of course, excellence is only one criterion in ranking applications. Many New Zealand funds (see over) also weight impact. Figure 1 (on the following page) shows the relative weighting of excellence and impact across the suite of government funds in the New Zealand RSI system.

Figure 1. The New Zealand RSI system showing the various government-financed research funds. The horizontal axis shows the relative weighting given to impact versus excellence (according to the MBIE definitions) in the assessment step of the application process. These weightings are taken from the various Requests for Proposals. Where other criteria are mentioned, we have ignored them in the weighting or, occasionally, interpreted them as proxies for excellence or impact. For details see Appendix 1. The vertical axis differentiates between contestable, negotiated and institutional funds. The size of the circles is proportional to the size of the funds. The different colours indicate different funders. We also show the proposed larger PBRF, which gives more weight to impact.



MBIE's Endeavour Fund Implementation

The MBIE's Endeavour Fund is New Zealand's largest source of financial support for scientific research awarded through a contestable process.¹⁰ Set up in 2016 as part of the *NSSI* reforms, it amalgamated several smaller, more specialist funds. The fund aims to provide the highest potential impacts across a range of economic, environmental and societal objectives, and give effect to Vision Mātauranga. Both applicants and assessors are given extensive guidelines on how to implement the *NSSI* definition in their applications and scoring. Excellence of the science and the research team are scored separately¹¹, and the criteria are gazetted as follows:¹²

Science. Research should be well-designed, involve risk and/or novelty, and leverage additional value from wider research. Assessment must have particular regard to whether the proposed research, science or technology or related activities:

- progress and disseminate new knowledge;
- have a well-designed research plan and credible approach to risk management;
- are ambitious in terms of scientific risk, technical risk, novelty and/or innovative approaches; and
- are well-positioned in the domestic and international research context.

Team. The proposed team should have the mix of complementary skills, knowledge and resources to deliver the proposed research, science or technology or related activities, and to manage risk.

The key emphases here are on good design, novelty, ambition and connectedness to existing knowledge and activities. The science assessment is more forward-looking; that of the team is predominantly retrospective, largely based on CVs (although some aspects, such as having an appropriate mix of skills, are prospective).

Marsden Fund Implementation

The Marsden Fund is New Zealand's largest fully competitive source of funding for investigator-led, "blue-skies" research. Administered by the Royal Society of New Zealand Te Apārangi using monies from the Business, Science and Innovation appropriation from Government, the Fund's mission statement¹³ is

To drive world-class research in New Zealand by supporting and incentivising excellent researchers to work on their best and boldest ideas and to connect internationally, leading to new knowledge and skills with the potential for significant downstream impact for New Zealand.

Note that the vehicle for achieving research excellence here is support for excellent *researchers*: track record, mana and researcher standing are all important. It appears that "impact" here is predominantly scholarly impact, in spite of the *NSSI* definition: the terms of reference¹⁴ specifically note that supported "research is not subject to government's socio-economic priorities."

¹⁰ <https://www.mbie.govt.nz/science-and-technology/science-and-innovation/funding-information-and-opportunities/investment-funds/endeavour-fund/>

¹¹ This separation has been controversial for several reasons (which are themselves disputed). For example, excellence is sometimes evaluated first and the score received acts as a hurdle that reduces the number of subsequent impact assessments. This timing is seen as privileging excellence over impact, although, in practice, successful applications require high scores for both pillars. More fundamentally, however, if excellence and impact are inextricably intertwined, their separate evaluation requires great care.

¹² <https://gazette.govt.nz/notice/id/2019-go3220>

¹³ See <https://www.royalsociety.org.nz/what-we-do/funds-and-opportunities/marsden/about/marsden-councils-strategic-direction/>

¹⁴ <https://www.royalsociety.org.nz/what-we-do/funds-and-opportunities/marsden/about/tor/>

Nevertheless, impact *sensu NSSI* is not specifically excluded. The Fund's investment plan¹⁵ aligns the assessment criteria used to score applications with the *NSSI*:

- Proposals must have the potential for significant scholarly impact because of the proposal's novelty, originality, insight and ambition
- Proposals must be rigorous, and should have a basis in prior research and use a sound research method
- The research team must have the ability and capacity to deliver
- Proposals should develop research skills in New Zealand, particularly those at the post-doctoral level and emerging researchers

The key concepts underlying research excellence in this view are novelty, insight, ambition, rigor and sound methodology, very like those in the Endeavour Fund's criteria, with the notable addition of scholarly impact. Again, there is a mix of prospective and retrospective evaluation.

Health Research Council Implementation

The *New Zealand Health Research Strategy*¹⁶ underpins the priorities of the Health Research Council (HRC), which funds the majority of health research in New Zealand, again through Vote Business, Science and Innovation. First in its list of four guiding principles is research excellence. The *Strategy* specifically cites the *NSSI* definitions of excellence and impact (pages 41 & 43), but (page 8) argues that adhering to these principles will "increase the impact of government investment in health research." Again, excellence and impact seem inextricably intertwined: research excellence is held to drive greater impact.

In pursuing research excellence, the *Strategy* says (page 10) that the HRC aims to

- Embrace and value a range of research approaches and methodologies that are fit for purpose and rigorous. Investment supports without bias a broad range of paradigms, approaches, methodologies and methods, but scientific rigour and well-designed methodologies are paramount.
- Create the environment to generate innovative ideas: a funding system that supports transformative and innovative ideas from researchers and consumers.
- Conduct ethical research that keeps research participants safe, protects the privacy of individuals, and respects the mana of families and whānau.

Notable here is the emphasis on a diversity of research, but also more traditional mention of rigor, good design and innovation.¹⁷ As befits health research especially, ethical standards are prominent, but it is notable that this requirement is included under excellence.

¹⁵ *Marsden Fund Investment Plan 2018-2020*. Available at <https://www.royalsociety.org.nz/what-we-do/funds-and-opportunities/marsden/about/marsden-councils-strategic-direction/>

¹⁶ Available at <https://hrc.govt.nz/resources/nz-health-research-strategy-2017-2027>

¹⁷ Further details on the HRC's view of the characteristics of excellence research can be found at <https://hrc.govt.nz/what-we-do/investing-excellent-research>. Methodological soundness, scientific robustness, the identification of knowledge gaps and good reporting are all mentioned. Ethical standards are also included here.

PBRF Definition

The Performance-Based Research Fund (PBRF) is a central feature of the New Zealand government's financial support for tertiary education through the Ministry of Education and Vote Education.¹⁸ It is primarily designed to reward and encourage excellent research by awarding more money to institutions in which excellent research is carried out. The way in which the fund is apportioned is determined, in part, by the results of a "Quality Evaluation" of individual researchers across the whole tertiary sector that takes place (currently) every six years. Note that this evaluation examines research across the whole of the education sector, not just science.¹⁹

The most recent Quality Evaluation assessment occurred in 2018, for which the *Guidelines*²⁰ defined excellence in an explicitly multi-dimensional way (page 66):

Excellence, in this respect, is not just about the production of high-quality research articles, books, exhibitions and other forms of research output. It also includes all of the following:

- the production and creation of leading-edge knowledge
- the application of that knowledge
- the dissemination of that knowledge to students and the wider community
- supporting current and potential researchers, such as postgraduate students, in the creation, application and dissemination of knowledge.

Note that some elements of this view of excellence – notably the dissemination of knowledge to the wider community – come very close to the MBIE definition of impact. Indeed, "end-user impact" was explicitly listed as a potential part of the Research Contribution component of each individual researcher's score. Unlike the MBIE and Marsden implementations above, the PBRF excellence evaluation is entirely retrospective, assessing past accomplishments of various types.

The excellence of individual researcher's research portfolios was assessed (in 2018) by one (or occasionally two) subject-specific panel(s) of researchers who acted as peer reviewers. Peer review is often held to be the "gold standard" for assessing excellence.²¹ Nevertheless, peer review has many drawbacks.²² It is expensive²³ (especially compared with bibliometric measures²⁴) and it is subject to a range of biases (e.g., discriminating against ethnic minorities²⁵, as well as women and researchers with unknown names or low-prestige institutional addresses²⁶).

Within the PBRF assessment, there were claims that applied research and some research subjects, especially those with a local focus, were scored more harshly.²⁷ In addition, some Māori researchers argued that the less traditional outcomes and outputs²⁸ of their research were not properly recognized. These criticisms and other issues with the definition of excellence were examined in 2019 by a review panel²⁹ whose report is discussed next.

¹⁸ See <https://www.tec.govt.nz/funding/funding-and-performance/funding/fund-finder/performance-based-research-fund/>

¹⁹ Research was held to be any "original investigation in all domains, including that of a creative, professional or applied nature ... not restricted to theoretical inquiry alone, but occurs across the full spectrum of original investigative activity."

²⁰ *Performance-Based Research Fund: Guidelines for the 2018 Quality Evaluation assessment process*. Tertiary Education Commission/Te Amorangi Mātauranga Matua, Wellington, New Zealand. ISBN 978-0-478-32052-7.

²¹ For example, Rossner (2002) writes, "Peer review remains the backdrop against which all other types of research evaluation appear, and often the standard against which their validity is judged." Rossner, J.D. 2002. Outcome measurement in the USA: State of the art. *Research Evaluation* 11: 85–93.

²² Panel scoring has issues, as well, such as "group-think" (Olbrecht, M. & Bornmann, L. 2010. Panel peer review of grant applications: What do we know from research in social psychology on judgment and decision-making in groups? *Research Evaluation* 19: 293-304.)

²³ The current COVID pandemic, however, has highlighted that using video-conferencing software, etc., can significantly cut costs of panel meetings.

²⁴ Abramo, G. & D'Angelo, C.A. 2011. Evaluating research: From informed peer review to bibliometrics. *Scientometrics* 87: 499-514.

²⁵ Moore, S., Neylon, C., Eve, M.P., O'Donnell, D.P. & Pattinson, D. 2017. "Excellence R Us": University research and the fetishisation of excellence. *Palgrave Communications* 3: 16105.

²⁶ Snodgrass, R. 2006. Single- versus double-blind reviewing: An analysis of the literature. *SIGMOD Record* 35(3): 8-21. Nevertheless, some studies have found that peer review of grant applications is not biased against women. See for example, Marsh, H.W., Jayasinghe, U.W. & Bond, N.W. 2011. Gender differences in peer reviews of grant applications: A substantive-methodological synergy in support of the null hypothesis model. *Journal of Informetrics* 5: 167–180.

²⁷ There was little hard evidence substantiating these two claims but the mere perception of bias is problematic in a process that works better if it is widely supported by the tertiary sector. See also footnote 31.

²⁸ By outputs we mean the products of research (traditionally conference presentations, reports, articles, books, patents, etc.); by outcomes we mean the short- to medium-term effects of the research (e.g., policy changes, new management practices and novel processes in industry).

²⁹ <https://www.education.govt.nz/further-education/policies-and-strategies/review-of-the-performance-based-research-fund/>

PBRF Review Panel Report

The very first recommendation of the panel's report³⁰ was that the assessment framework adopt "a more capacious definition of research excellence" (page 4). The outcome sought by the panel from this change was to ensure that the definition "encompasses the production of research, engagement and impact in relation to that research and support for vibrant, diverse research cultures" (page 8). In addition, the panel wanted "more porous boundaries between research outputs and research contributions" (i.e., "activities that sustain and develop the research" system (page 61)). In short, the assessment should "recognize research excellence in all its diverse forms" (page 62).

This emphasis on the diversity of what constitutes research excellence is, in part, a recognition that some participants in previous PBRF assessments believed, rightly or wrongly, that certain types of research were privileged.³¹ Reiterating (and expanding) the message about research diversity would go some way to alleviating these concerns. In addition, the panel recommended design changes that would better recognize excellence in Mātauranga Māori and associated research practices. Similar concerns also applied to Pasifika research. Again, emphasizing diversity speaks to these issues.

The inclusion of the impact of research in the panel's view of excellence is also worth noting, not the least because the *NSSI* definition separates these two features. The panel was clear that research excellence included "scholarly and non-scholarly (broadly defined) impacts" (page 62). In an illustrative example (page 63) they cite "research leading to change in the way an industry handles a by-product of a construction process," which clearly falls within the *NSSI* definition of impact, not excellence.

³⁰ *Report of the PBRF Review Panel*. 2020. ISBN 978-0-473-51276-7.

³¹ The *Report* describes such views as 'myths.' On page 6, for example, it recommends that the "TEC should work with key stakeholders to develop suitable exemplars that demonstrate how impact can be presented effectively in evidence portfolios, and address other myths about how the assessment system privileges certain kinds of research, research outlets or disciplines."

Royal Society Te Apārangi Definition

The past ten years have seen major changes in Aotearoa New Zealand's elite national academy, the Royal Society Te Apārangi. Its remit has been broadened from scientific and technological disciplines to include the humanities, and there is greater recognition of the need to embrace diversity in its membership and awards, and to honour the Treaty of Waitangi Te Tiriti. In 2019, in response to these changes, the president of the society (and member of the 2019 PBRF Review Panel), Professor Wendy Lerner FRSNZ, gave an address entitled 'Research Excellence in a "Grand Challenge" World,' in which she argued that we need to acknowledge multiple forms of research excellence, including impact and advancement.³² In Lerner's view, the view of research excellence as involving little more than publishing lots of papers in highly ranked journals (and hence scoring highly in various bibliographic metrics) is outdated, over-emphasizing the narrow individualized and competitive aspects.

Drawing on her experience in the British equivalent of the PBRF, the REF, Lerner went on to suggest that excellence should be considered very much more broadly than in either the MBIE or PBRF views outlined above. She saw the deeply engaged researcher, whose research-programme design and dissemination of results has the genuine involvement of stakeholders, as epitomizing a more holistic approach to excellence. Their work may use a multidisciplinary approach and the outputs are likely to be more varied than simply specialist journal articles. Such an approach, Lerner argued, was more likely to recognize excellence more widely, especially in Māori and female researchers, as well as in applied research and creative fields such as literature. In tandem, the Society has been working to enhance the diversity of the membership of its assessment panels and to embed training of these assessors to minimize bias and better recognize the diversity of what excellence can mean.

³² <https://www.royalsociety.org.nz/major-issues-and-projects/presidents-address-2019/>

CoRE Definition

The Centres of Research Excellence (CoREs) Fund was established in 2001 to encourage excellent research in the tertiary sector that was collaborative and strategic, resulting in significant knowledge transfer. The most recent funding round was in 2019-2020 and the applications were scored against four criteria, one of which was “Excellence: excellent research,” which was assessed by considering³³:

- academic strength of the proposed research team;
- academic strength of the proposed research programme;
- commitment to ensuring research is innovative, outcomes-focussed and has impact;
- strength of proposed collaboration and the degree to which partners have contributed to the proposal;
- potential of the CoRE to have national and international influence; and
- commitment to equity and wellbeing outcomes, including encouraging and enabling diversity and inclusion for Māori, Pacific peoples, and other under-represented groups in the proposed research programme and/or in the proposed research team.

This view of research excellence is significantly different from many of the preceding ones, notably in its explicit emphasis on collaboration, at the level of both individual researchers and the partner institutions for which the researchers work. Like the opinions of the PBRF review panel and as Professor Lerner outlined above, though, it points to the importance of outcomes and impact, explicitly including them as part of research excellence. Indeed, excellent research is, in this view, focussed on outcome.³⁴ As with referees for the Endeavour and Marsden Funds, the CoRE evaluators needed to look both at the potential of the proposed research programmes, as well as look back at the track record of the researchers involved.

³³ Appendix II – Selection Criteria. *Centres of Research Excellence (CoREs). Guidelines for the conduct of Expert Selection Panel meetings*. Available at <https://www.royalsociety.org.nz/what-we-do/funds-and-opportunities/cores/cores-selection-process/guidelines-for-expert-selection-panels/>

³⁴ This focus seems at odds with MBIE’s view of the CoREs, which that ministry’s draft *Research, Science and Innovation Strategy* considers to be solely investigator driven, akin to the Marsden Fund.

A Māori Perspective

Almost every recent discussion of research excellence in a New Zealand context has noted that the definitions and assessment of excellence have not served Māori well.³⁵ This failure³⁶ continues despite attempts at ameliorating the issues (e.g., MBIE's inclusion in Endeavour Fund applications of a section outlining how the proposal implements the Vision Mātauranga policy³⁷).

Some of the problems are structural. For example, the *NSSI* definition of research excellence fails to acknowledge or accommodate mātauranga Māori.³⁸ Indeed, Rauika Māngai asserts that “Agencies and institutions should embrace measures of science excellence that include Mātauranga Māori.”³⁹ There is a long-overdue need for criteria that adequately assess Mātauranga-Māori based research that is Māori led or co-developed with Māori, or that addresses the Vision Mātauranga policy.⁴⁰ It is appropriate that these criteria be determined by Māori for inclusion in assessment processes. The development of such criteria for research excellence might be expected to consider the following points.⁴¹

Excellent research would focus on questions, interests and outcomes of importance to Māori, and contribute to Māori-driven aspirations, opportunities and strategic goals. Working with Māori requires partnership and co-design, as well as long-term trusting relationships.⁴² Māori would actively participate in all aspects of the research, from the design of the research questions to the pathways implementing the results. Doing so will ensure that the research is fit for the cultural (and, indeed, national and international) context in which it will be applied. It will also develop Māori capability and capacity across the research workforce, with an intergenerational view of its participants and impact.

Excellent research will often use a trans-disciplinary approach,⁴³ focusing on sharing knowledge and skillsets and integrating knowledge from heterogeneous sources. Such research respects, recognizes and draws on Māori as experts (e.g., as knowledge holders and generators),⁴⁴ and protects their intellectual property rights. Excellent research will recognize the diversity of Māori decision-making contexts; it will follow tikanga, respecting and protecting both whānau and researchers.

Excellence will often be linked to impact and capability development, because of the original focus of the research.⁴⁵ Research outcomes will extend well beyond the life of the programme and reflect overarching Māori-driven needs. This requirement exemplifies the link from the original idea and hypothesis generation to the outcomes and actions.⁴⁶ The effectiveness of the implementation of results will be critical, likely catalysed by knowledge intermediaries who have an understanding of both Māori and science-sector worldviews and who have the trust of both communities.

³⁵ Rauika Māngai. (2020). *A Guide to Vision Mātauranga: Lessons from Māori Voices in the New Zealand Science Sector*. Wellington, NZ: Rauika Māngai. ISBN: 978-0-473-52757-0 (Softcover). ISBN: 978-0-473-52758-7 (PDF). Available from <http://www.maramatanga.co.nz/>

³⁶ It is not clear whether the reasons for this outcome are the consequences of the sorts of questions being asked, the legitimacy or substance of the responses, the quality of the assessors, a bias towards contemporary excellence, the composition and resourcing structure of the research teams, or other reasons.

³⁷ MoRST. 2007. *Vision Mātauranga: Unlocking the Innovation Potential of Māori Knowledge, Resources and People*. Wellington, NZ: MoRST. ISBN: 978-0-478-06219-1

³⁸ Mātauranga Māori is absent from page 13 of the *NSSI*, where excellence is defined.

³⁹ Rauika Māngai. (2020). *A Guide to Vision Mātauranga: Lessons from Māori Voices in the New Zealand Science Sector*. Wellington, NZ: Rauika Māngai. ISBN: 978-0-473-52757-0 (Softcover). ISBN: 978-0-473-52758-7 (PDF). Available from <http://www.maramatanga.co.nz/>

⁴⁰ Indeed, it might be better to go beyond this policy, for example, to embrace knowledge systems and contributors of that knowledge, as well as methods of examination that are useful by not typically included. Similarly, it may be productive to enable Māori knowledge using public funds.

⁴¹ The following points were developed after discussions and consultation with members of the Māori researcher community. They are not intended to be comprehensive: obviously, any fuller view would also need to be led and developed by Māori.

⁴² This means everyone knows each other's priorities, capacity and capability well, and should lead to an alignment of goals. There may be a formal or informal partnership as mutually agreed, enabling joint opportunities to address various research needs overtime, and resulting in a joint track record of delivery (via previous kaupapa).

⁴³ By trans-disciplinary, we mean an approach that integrates the findings from research drawing on several disciplines or knowledge-holders, to address a problem that does not clearly fall within one discipline, as opposed to multi-disciplinary, in which the problem is narrower. Trans-disciplinarity implies researchers from different disciplines, and also other knowledge holders, working collaboratively to create new understandings that integrate and move beyond discipline-specific approaches to address a specific issue or need. Trans-disciplinarity, as opposed to multi-disciplinarity and inter-disciplinarity, requires that researchers co-develop new research at the intersection of their respective fields.

⁴⁴ It is not only mainstream research approaches and philosophies that can be used to deliver research excellence.

⁴⁵ The current interests of Māori around, for example, sustainability, cultural identity and economic development.

⁴⁶ This is not to suggest that what might be seen as “the research of adventure and discovery” with a less obvious link to impact should not be supported.

International Perspectives

Many national funding agencies around the world (e.g., those in UK, Canada, US, Australia, EU) also allocate research funds based on assessments of proposal excellence. Definitions of research excellence, however, vary; measuring research excellence is considered even more difficult⁴⁷. At a broad level, assessment of research excellence has traditionally included consideration of (i) the track record of individual researchers and teams, especially publication records and citations; (ii) the novelty of the proposed research; and, more recently, (iii) the potential of the research to have relevance and impact.

While research excellence is generally accepted around the globe as a fair and efficient assessment criterion for allocating scarce research resources, a range of international studies have identified several factors that need to be considered when basing funding decisions on excellence⁴⁸. These studies highlighted the following potential problems:

- Rather than improve national research capacity, rewarding excellence can result in underfunding the ‘normal research’ that maintains basic research capacity and capability. Similarly, it may distract attention from national priorities to focus on performance measures.
- Rewarding excellence has an impact on research capacity, as it disproportionately allocates resources to those academics and institutions already well rewarded.
- In situations where there is ‘hypercompetition’, emphasis on excellence tends to encourage ‘grantsmanship’ and there is an overemphasis on novel and glamorous research at the expense of ‘normal research’. Publishers can also exacerbate this effect through favouring novel research that exhibits positive findings.
- There is a tendency for researchers to embellish results and impact claims in research proposals and doing so is considered by many researchers and institutions to be a necessity and par for the course in applications for competitive research funds.
- Research that provides replication, negative results or the collection of new types of data, while fundamental to the research process and capacity, generally fail to meet ‘excellence’ criteria.
- Concentrating on ‘excellence’ may promote scientists to work with those similar to themselves (because it is easier to do so), thereby generating a bias within research to the exclusion of the opposite gender, ethnic groups or non-traditional areas of scholarship.

To overcome some of the above limitations and unintended consequences, some countries, such as Canada, have begun to redefine research excellence. Their focus is on ensuring that their future research is open, international and interdisciplinary, as well as equitable, diverse and inclusive. The EU is similarly proposing to place greater emphasis on research soundness, capacity and diversity⁴⁹. These new definitions of research excellence place greater emphasis on thoroughness, completeness, reproducibility, statistical or bibliographic appropriateness, evidence and probity, rather than the existing definitions based on promised outputs and impact. This redefinition of excellence is considered to be essential if the research culture of competition is to be reduced, the distribution of resources expanded to promote diversity and support is provided for the incremental science advances that research relies upon.

⁴⁷ Canadian Science Policy Centre, 2019 Conference – Science & Policy, Panel 138: The future of research excellence: a conversation with Canada’s granting agencies.

⁴⁸ Moore, S., Neylon, C., Eve, M.P., O’Donnell, D.P. & Pattinson, D. 2017. “Excellence R Us”: University research and the fetishisation of excellence. *Palgrave Communications* 3: 16105.

⁴⁹ Eulife, 2016. Key elements for excellence in research: a contribution from EU-LIFE. 11pp. Available at <https://eu-life.eu/newsroom/publications/key-elements-excellence-research-contribution-eu-life>

Bibliometrics & Excellence

The view that excellence cannot be encapsulated by a single measure (or even a suite of numbers) is a recurring theme in discussions of the topic. In part, this caution is a reaction to the spread of bibliographic metrics, such as the Hirsch index (h)⁵⁰, which come up with a single number for a researcher or institution. Similarly, research excellence is surely more than simply publishing in journals with high impact factors⁵¹. Indeed, a number of metrics (including journal impact factor) were developed by publishers as new services rather than for reasons to do with assessment of excellence.

The use of journal-based metrics in evaluating completed research is highly contested.⁵² Indeed, there is evidence that, even within a single field of research, excellence as assessed by peer review and excellence as inferred from bibliometric data (e.g., high citation rates) are different.⁵³ All bibliographic measures have structural issues that cloud their interpretation. For example, although attractive in its simplicity and ease of calculation, the Hirsch index has a number of drawbacks, such as being strongly affected by the field (and subfield) of research and yet unaffected by the number of authors on a paper.⁵⁴ In addition, h is obviously affected by citation practices, which are known to be biased in numerous ways (e.g., against early-career researchers and in favour of review papers rather than original research).⁵⁵

Metrics also drive researchers' behaviour, often in undesirable ways.⁵⁶ For example, the increase in the number of published papers, the trend towards "least publishable units" and the increase in the number of co-authors on papers, all in the last few decades, are thought to be driven in part by gaming of bibliometric scores⁵⁷. More worryingly, PhD projects are reportedly steered towards safer, less risky topics, which boost student numbers and are more likely to lead to publications.⁵⁸

Most fundamentally, as several commentators have pointed out, bibliometrics give the illusion that non-experts can judge the quality of the research, overcoming their informational disadvantage.⁵⁹ Having decisions about research funding, etc., made by those unfamiliar with the research itself violates the Haldane Principle and contravenes international best practice. For all these reasons (and more), there has been widespread condemnation of the overuse of bibliometric data in making decisions about research quality. The *San Francisco Declaration on Research Assessment*,⁶⁰ for instance, recommends they not be used "as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion, or funding decisions."

⁵⁰ The Hirsch index or h index combines one measure of a researcher's productivity (the number of papers produced) and a measure of their academic impact (the number of citations of the researcher's papers) into a single number, h , the number of papers that have been cited at least h times.

⁵¹ The impact factor of a journal is the number of citations to articles in that journal divided by the number of articles published in that journal, taken over a certain period (usually the previous two, or occasionally five, years). This metric is about academic impact, not impact in the sense of the *NSSI*.

⁵² Moreover, many of these metrics can be subjected to gaming, e.g., "coercive citation," in which referees of manuscripts submitted to journals recommending the citation of their own papers. See Wren, J.D., Valencia, A. & Kelso, J. 2019. Reviewer-coerced citation: Case report, update on journal policy and suggestions for future prevention. *Bioinformatics* 35: 3217–3218.

⁵³ Smolinsky, L., Sage, D.S., Lercher, A.J. *et al.* 2021. Citations versus expert opinions: Citation analysis of featured reviews of the American Mathematical Society. *Scientometrics* 126: 3853–3870.

⁵⁴ Costas, R. & Bordons, M. 2007. The h -index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics* 1: 193–203; Schreiber, M. 2008. To share the fame in a fair way, h_m modifies h for multi-authored manuscripts. *New Journal of Physics* 10: 040201.

⁵⁵ Biscaro, C. & Giupponi, C. 2014. Co-authorship and bibliographic coupling network effects on citations. *PLOS ONE* 9: e99502.

⁵⁶ For an amusing take on life in a dystopian metrics-driven university see Helmer, S., Blumenthal, D.B. & Paschen, K. 2020. What is meaningful research and how should we measure it? *Scientometrics* 125: 153–169.

⁵⁷ Reviewed (with further references) in Helmer, S., Blumenthal, D.B. & Paschen, K. 2020. What is meaningful research and how should we measure it? *Scientometrics* 125: 153–169.

⁵⁸ Weingart, P. 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics* 2: 117–131.

⁵⁹ For example, "Metrics provide the necessary simplification, and they promise to be impartial, deterministic, and decision-friendly." Helmer, S., Blumenthal, D.B. & Paschen, K. 2020. What is meaningful research and how should we measure it? *Scientometrics* 125: 153–169. Others go further: "The appeal of numbers is especially compelling to bureaucratic officials ... A decision made by the numbers ... has at least the appearance of being fair and impersonal. ... Quantification is a way of making decisions without seeming to decide." Porter, T.M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press, Princeton, New Jersey.

⁶⁰ <https://sfidora.org/>. This declaration has been signed by >2000 research organizations (e.g., the [Australian Academy of Science](#), the Association of Australian Medical Research Institutes, but none from New Zealand) and >16,000 individuals.

Bibliometrics do have an advantage in that they are objective – the same value will always be calculated from the same bibliometric data – but their interpretation may be less so. Indeed, bibliographic data does have some role to play in assessing excellence and can be very useful if suitably contextualized. For example, in assessing temporal changes at the institutional level, issues such as different citation rates across different disciplines are not relevant because the same disciplines are being compared. Bibliometric assessments also have a distinct advantage in being relatively inexpensive.⁶¹

⁶¹ Abramo, G. & D'Angelo, C.A. 2011. Evaluating research: from informed peer review to bibliometrics. *Scientometrics* 87: 499-514.

Discussion

There is widespread agreement that the definition of research excellence must be broad, even broader than many have previously recognized. A multi-dimensional view of excellence will lead to better research outcomes.⁶² Narrow concepts that suit particular fields of research are not likely to work across a whole system. A wider view is especially important in understanding just what constitutes excellence in different cultural contexts.

Diversity in excellence is not without practical problems, however. In the context of grant applications, for example, assessing the excellence of applications that involve different researchers from different disciplines, with different cultural perspectives and working in different institutions is likely to be challenging, even if this diversity is desirable. Research funders will need to respond effectively and fairly to such situations, not the least because research assessment shapes research culture.

Peer-review is usually held to be the best way to identify excellence, and we concur. Nevertheless, quality peer review requires careful selection and instruction of those peers. In addition to more obvious matters such as ensuring a good match of expertise,⁶³ panels do better when they are more diverse (both demographically and in their research backgrounds).⁶⁴ Moreover, peer review is not without problems, most notably various biases.⁶⁵ We note that a number of New Zealand organizations – for example, the TEC in its most recent PBRF round, the Marsden Fund and HRC panels and various Royal Society Te Apārangi panels – have instituted assessor training to try to mitigate the effects of such biases.

Bibliometric assessments, while problematic on a number of grounds, do have a role to play, especially if comparisons truly are like-with-like. Unfortunately, in situations when unrelated research domains are involved, or when granular distinction between individuals' grant applications is required, bibliometrics have much less to offer.⁶⁶

Part of the reason that research excellence has no widely accepted definition nor an agreed way for its measure is that, according to a number of authorities, it is “essentially contested” *sensu* Gallie.⁶⁷ For example, one recent examination of research-excellence indicators noted that they are called upon to inform policy makers, but, as an essentially contested concept, research excellence “cannot be defined in a single-best, fixed, and objective way from the outset. Hence, they [research-excellence indicators] are likely to produce endless debates on their interpretation and implications.”⁶⁸

There is significant variation among New Zealand's various research funders in their views and uses of excellence (see table below for a summary). The *NSSI* definition of excellence is unusual internationally in explicitly excluding impact⁶⁹; most other definitions and attempts to measure excellence also evaluate impact. This distinction need not be problematic, however, provided there is clarity about the boundaries between excellence and impact as these are assessed, the definition of the latter (to exclude academic impact so as to avoid “double-counting”), as well as sufficient recognition for impactful research.

⁶² As the Global Research Council puts it, “Diverse R&I sectors create high-quality research and impact.” See Fraser, C., Nienaltowski, M.-H., Porter Goff, K., Firth, C., Sharman, B., Bright, M. & Martins Dias, S. 2021. *Responsible Research Assessment – A Virtual Conference from the Global Research Council*.

⁶³ In the case of applied research, for example, it may be appropriate to involve end-users. If impact is included in excellence, the case for end-user peer-reviewers is even stronger; indeed, the reviewers may not be peers.

⁶⁴ The UK's Engineering and Physical Sciences Research Council, for example, argues that inclusive peer review is fairer and supports its equity, diversity and inclusion goals. <https://epsrc.ukri.org/funding/edi-at-epsrc/evolving-and-upholding-fairness-in-peer-review/>

⁶⁵ See footnotes 22, 24-26.

⁶⁶ Citation rate, *h* indices, etc., are all inclined to favour established senior researchers and existing research networks, which might lead to higher scores for less risky, less novel grant applications, and a less diverse cohort of funded researchers, for example.

⁶⁷ Gallie, W.B. 1955. Essentially contested concepts. *Proceedings of the Aristotelian Society* 56: 167–198.

⁶⁸ For example, Ferretti, F., Pereira, Â.G., Vértesy, D. & Hardeman, S. 2018. Research excellence indicators: Time to reimagine the ‘making of’? *Science and Public Policy* 45: 1–11.

⁶⁹ Nevertheless, this distinction is also made in the Irish RSI sector. See *Agenda 2020: Excellence and Impact*. 2012. Science Foundation Ireland, especially page 3.

SOURCE	SOURCE TYPE	STANCE ON EXCELLENCE
NSSI	Strategy	Excellence defined as “well-designed, well-performed, well-reported research, recognised as such, e.g., through peer review.”
Endeavour	Fund	Aligns with NSSI, excellence of science and team scored separately
Marsden	Fund	Aligns with NSSI, set up to incentivize excellent researchers
Health Research Strategy	Strategy	Cites NSSI, but includes ideas of diversity and ethics also
PBRF	Fund	Incorporates more of an impact element than NSSI
PBRF Review	Review Panel	First recommendation suggests a broader definition of research excellence, including more of an appreciation for diverse research forms and impact (which is split from excellence in NSSI)
Royal Society Te Apārangi	Institution	Excellence should be more broadly defined than as by NSSI and PBRF, and that that will lead to more diversity in research
CoREs	Fund	Explicitly includes collaboration in definition of excellence, as well as outcomes and impact

Nevertheless, there is clearly a link between excellence and impact, with a number of those giving us feedback noting this point in different contexts. For example, what counts as excellent research for Māori will likely have significant impact. Others argued that, quite generally, excellence research is likely to have genuine impact, especially with applied research. Separating excellence and impact assessment might also affect how risk is viewed, which would be problematic given that the “Transform” category of Research Programmes in the Endeavour Fund is designed to encourage high-risk proposals.

The separation of excellence and impact in the New Zealand RS&I sector is, however, long-standing, predating the *NSSI* by several decades. The predecessors of MBIE’s Endeavour Fund, for example, were assessed against four criteria: science, team track record, benefit to New Zealand and pathway to implementation. The first two of these remain the two excellence criteria against which applications to the Endeavour Fund are scored; the latter two are effectively the two scorable impact criteria. Previously, however, all four criteria were scored by the same assessors; the Endeavour Fund has separate excellence and impact assessors. Analysis of pre-*NSSI* scores showed strong correlations between the excellence and impact scores: assessors clearly considered proposals they saw as excellent as simultaneously impactful and *vice versa*. Moreover, riskier research, which is likely underpinned by less established science, was inclined to score poorly on impact as well as excellence. With the advent of separate excellence and impact assessment, however, these correlations became less apparent,⁷⁰ which suggests that assessors may have been confounding these two concepts. Consequently, we do not recommend a return to using a single assessor for both excellence and impact, nor do we think that the conceptual separation of these two pillars of the *NSSI* is problematic.

The draft *RSI Strategy* definition of excellence strikes us as circular and impractically vague: What is “the best thing possible”? Indeed, this view appears to echo the “you know it when you see it” type of approach. We suggest that this draft definition be revised.⁷¹

In measuring or assessing excellence, we need to be explicit about why we want to do so and what use we will make of the assessment. Such considerations are particularly important around investment decisions: the intent of the investment needs to be reflected in the assessment criteria. Depending on context, it may be more important for research to advance the particular discipline; in other situations, making the knowledge practical and advancing the use of that knowledge (perhaps in a different discipline from where it was first researched) may be the prime objective. Indeed, the

⁷⁰ This decrease may be partially an artefact, however, as not all proposals are currently assessed for impact. Smart Ideas concept proposals are assessed only for excellence; only those exceeding some hurdle are invited to submit a full proposal, which receive both excellence and impact scores. There may be a good case for assessing more (all?) proposals for impact in order to ensure that excellence is not privileged over impact.

⁷¹ Several of those commenting on the draft of this document explicitly recommended returning to the *NSSI* definition. All the same, “the best thing possible” might be seen as significantly broader, depending on just what is meant by “best”.

issue of whether or not impact is part of excellence is really a matter of what we want to measure. In the context of the PBRF, for instance, including impact is clearly favoured. If impact is assessed separately (as in the Endeavour Fund), then it should not be included in excellence assessment. For these reasons, we do not recommend a single definition of excellence across the RS&I sector.

The granularity at which we measure excellence is also important. Individual research grants (and similar, individual-level assessment for prizes, promotion, etc.) have different requirements from institutional-level assessment (e.g., PBRF). In the latter case, any biases from using proxies for excellence such as bibliometric measures are likely to be less than in the case of individuals.

We also need to be more aware of unintended consequences of over-emphasizing excellence (as outlined in the bulleted list above). Clearly, by definition, not all research can be excellent, but these problems suggest that not all research *should* be excellent (at least without suitable safeguards to ameliorate some of the problems). Governments often have other priorities (political, environmental, economic and social) where impactful, methodologically sound but possibly less novel research might suffice. Similarly, an overemphasis on novelty or risky research may preclude the impactful application of existing knowledge in different contexts. That is not to say, of course, that increasing the proportion of excellent research (by whatever measure) as an aspiration is not a worthy goal, especially if we employ a broader definition of excellence.

New Zealand is not the only place in which the view as to what counts as research excellence has evolved. The view in the European Union has changed profoundly over the last twenty years, but, surprisingly, not necessarily in the same direction. According to Sørensen and colleagues⁷² the understanding of excellence evolved from “a fuzzy concept, intrinsically embedded in research and researchers and revealed by peer review” to something “more sharply defined,” quantifiable, “rooted in measures of research outputs and their commercial applications,” “and connected with a particular sort of knowledge that which produces breakthroughs.” Interestingly, some of this quantification relies on bibliometric data, as well as university and research-organization rankings⁷³ and numbers of granted patents.

The way in which past research performance is used in assessing proposals for future research – a kind of “back to the future” approach – also raises questions. Ideally, we would employ system settings that promote excellence rather than outputs that are proxies for excellence.

Based on the above review and discussion, we make the following recommendations:

1. Research funders should examine their assessment criteria for excellence to ensure they value the full diversity of ways of achieving excellence in research. Similarly, funders need to ensure this diversity is recognized in practice, for example, by using assessment panel members with an appropriate variety of research backgrounds.
2. Future definitions of excellence by research funders need to ensure they do not disadvantage or discriminate against Māori, gender, institutions, ethnic groups or non-traditional areas of study. In doing so, funders should draw on international efforts to address this issue.
3. Māori should continue to develop their own definitions and measures for research excellence to guide future assessments of excellence by research funders.
4. Although research excellence and impact are intertwined, the definitions of research excellence and impact should be kept separate. Different funding bodies should balance the weightings they assign to excellence and impact so that they are appropriate to their funds’ missions.

⁷² Sørensen, M.P., Bloch, C. & Young, M. 2016. Excellence in the knowledge-based economy: From scientific to research excellence. *European Journal of Higher Education* 6: 217–236.

⁷³ Such rankings are highly controversial, however, even as they have become increasingly influential with policymakers and government decision-makers. For a recent critique see Hazelkorn, E. 2019. University rankings: There is room for error and “malpractice”. *Elephant in the Lab*. <https://doi.org/10.5281/zenodo.2592196>. Similarly, the Global Research Council has recently noted that such rankings “often do not measure what matters to the R&I system and they do not demonstrate excellence at a useful scale for users of this information.” See Fraser, C., Nienaltowski, M.-H., Porter Goff, K., Firth, C., Sharman, B., Bright, M. & Martins Dias, S. 2021. *Responsible Research Assessment – A Virtual Conference from the Global Research Council*.

5. Consideration should be given to providing a clearer definition of excellence in future drafts of the RS&I Strategy. This view should be high-level and not overly prescriptive, however, so that it can encompass a diversity of approaches to pursuing excellence.
6. Peer review should continue to be regarded as the gold standard for assessing research excellence. Nevertheless, funders should ensure that documented problems (e.g., biases) with peer review and other forms of assessment are acknowledged and addressed (e.g., through training of assessors).
7. Funders should ensure that support for research based on excellence does not disadvantage or reduce the methodologically excellent day-to-day research and infrastructure that underpins the RS&I sector.

Acknowledgements

We are grateful to many people for the insights they have shared with us in response to our request for feedback on a previous draft of this document. We thank David Bilkey, Richard Blaikie, Giselle Byrnes, Sunny Collins, Wayne Cutfield, Sally Davenport, Gill Dobbie, Charles Eason, Will Edwards, Vince Galvin, Juliet Gerrard, Peter Gluckman, Margaret Hyland, Ian Lambie, Wendy Lerner, Marc Lubbers, Jim Mann, Sarah McDermott, Jim Metson, Richard Newcomb, Louise Parr-Brownlie, Daniel Patrick, Kay Saville-Smith, George Slim, Richard Smith, Ian Town, Jenny Webster-Brown, Liz Wedderburn, Rob Whitney, Prue Williams and David Wratt. We are especially grateful to Erica Williams and Marino Tahī (NIWA) and Willy-John Martin (MBIE) for their contribution to our discussions about a Māori perspective on excellence. Olivia Spratt (MBIE) discussed the whole document with us and, together with Kerry Duckmanton (MBIE), put together Figure 1. The views expressed above are ours, however, and those people and organizations named do not necessarily agree with us. We have tried to indicate some instances of disparate opinion in the footnotes.

Appendix 1 : Data Underlying Fig 1.

In Figure 1: The New Zealand RSI system (page 6) the various government-financed research funds are shown. The horizontal axis shows the relative weighting given to impact versus excellence (according to the MBIE definitions) in the assessment step of the application process. These weightings are taken from the various Requests for Proposals. Where other criteria are mentioned, we have ignored them in the weighting or, occasionally, interpreted them as proxies for excellence or impact. For details see the table below.

Department -RSI bubble Diagram	Appropriation Name	2020/21	Contestable/ Negotiated/ Institutional	Position	Duration	% Excellence	% Impact	Rationale
Callaghan Innovation	R&D Growth Grants	\$ 172,586	Negotiated	50	Variable	0	100	Granted based on eligibility but overall focus is on economic development/ impact.
Callaghan Innovation	Building Business Innovation	\$ 36,378	Institutional	15	Annual	0	100	Operations, so focussed on impact.
Callaghan Innovation	R&D Services and Facilities for Business and Industry	\$ 36,160	Institutional	25	Annual	0	100	Operations, so focussed on impact.
Callaghan Innovation	Targeted Business R&D Funding	\$ 32,500	Negotiated	75	Variable	0	100	Strong focus on industry, hence assumed to be strongly impact driven.
Callaghan Innovation	National Measurement Standards	\$ 8,118	Institutional	20	4 years	35	65	World-standard work being undertaken, primarily to feed in to other research - hence more impact weighted.
Health Research Council	Health Research Fund	\$ 117,489	Contestable	100	Up to 2 years	58	42	Assessed on five criteria: - Two excellence based (quality of health research (22.5%), research team collaboration and integration (22.5%)) - Two impact based (potential for outcomes (22.5%), Māori health advancement (10%)), and - One that is not relevant to either and has hence not been included in ratio (vision of programme (22.5%)).
Inland Revenue Department	R&D Tax Incentive	\$ 281,000	Negotiated	60	Variable	0	100	Geared towards industry, hence assumed to be strongly impact driven.
Ministry for the Environment	Water science and economics	\$ 1,500	Negotiated	65	?	10	90	?
Ministry for Primary Industries	Global Research Alliance on Agricultural Greenhouse Gases	\$ 8,500	Negotiated	55	?	10	90	?
Ministry for Primary Industries	Programmes Supporting Sustainability	\$ 8,567	Negotiated	60	?	10	90	?
Ministry for Primary Industries	New Zealand Agricultural Greenhouse Gas Research (NZ Ag GG Research)	\$ 4,850	Negotiated	63	?	33	67	Assumed to be mostly impact driven, however one of the three key aims is to build research capacity so one third the the ratio has been allocated to excellence.
Ministry for Primary Industries	Sustainable Food and Fibre Futures	\$ 32,468	Contestable	100	Variable	40	60	Nine assessment criteria, assumed to have equal weighting. Two excellence based: - Innovation, - Beyond business as usual. Three impact based: - Sustainable benefits to New Zealand, - The fit with relevant strategies, - Adoption and extension/path to market. Four that have not been included as they aren't relevant to impact or excellence: Ability to deliver, Governance, Risk identification and mitigation, Budget.
Ministry of Business Innovation and Employment	Maori Innovation Fund	\$ 1,000	Contestable	100	18 months	20	80	He Tupu Ōhanga - Commercial Advisors Scheme assesses applications based on: - Alignment with one or more of five pou – employment, rangatahi, enterprise, regions, education (30%), - Potential of the collective's or coalition's assets for growth (30%), - Capability of collective/ coalition (20%), and - Suitability and potential of proposed project (20%) All assumed to be impact related, except capability of collective/ coalition. There are no set funding assessment guidelines for the Pakihi workshops or the Rangatahi Business Challenges, so it is assumed that the excellence to impact ratio is the same as for He Tupu Ōhanga - Commercial Advisors Scheme.

Ministry of Business Innovation and Employment	National Science Challenges (NSCs)	\$ 114,744	Negotiated	65	10 years	62.5	37.5	Assessed based on three criteria: - The research, science, and technology will be of excellent quality - The proposal is focused on delivering impact - Decision-making and accountability arrangements are sound and enduring Third criteria has not been included in criteria, and excellence : impact ratio has been assumed to be the same as for the CoREs as NSCs similar in nature.
Ministry of Business Innovation and Employment	Partnered Research Fund	\$ 33,390	Contestable	100	2 years	10	90	Pre-seed Accelerator Fund assessed based on: - Return on investment (50%), - Pathway to market (20%), - Ability to deliver (20%), and - Future commercialisation ability (10%). All criteria apply to impact, however we have interpreted "ability to deliver" to be split half and half to both excellence and impact. There are no set funding assessment guidelines for the Commercialisation Partner Network, Environlink or the Christchurch Innovation Hub, so it is assumed that the excellence to impact ratio is the same as for the Pre-seed Accelerator Fund.
Ministry of Business Innovation and Employment	Catalyst Fund	\$ 34,751	Negotiated	55	Typically 3 years	50	50	Set up to drive increasing excellence and the potential for impact for New Zealand via international science and innovation connectivity. We have assumed excellence and impact both equal drivers.
Ministry of Business Innovation and Employment	Endeavour Programmes	\$ 224,712	Contestable	100	3-5 years	50	50	Assessed based on: - Excellence (science 25%, team 25%) - Impact (benefit to NZ 25%, implementation pathway(s) 25%)
Ministry of Business Innovation and Employment	Endeavour Smart Ideas	\$ 18,000	Contestable	100	2-3 years	65	35	Assessed based on: - Excellence (science 50%, team 15%) - Impact (benefit to NZ 25%, implementation pathway(s) 10%). Note: Smart Ideas was cancelled for 2020/21, but \$18 M is average value for past years and is proposed future value also.
Ministry of Business Innovation and Employment	Regional Research Institutes	\$ 11,353	Negotiated	58	3 years	33	67	Assessed based on six performance indicators, including excellence and impact – have assumed 1/3 excellence and 2/3 impact as purpose of fund is to drive regional growth.
Ministry of Business Innovation and Employment	Strategic Science Investment Fund Programmes (SSIF Programmes)	\$ 260,822	Negotiated	60	Variable	50	50	Assessed based on seven performance indicators, including excellence and impact – it is assumed they are evenly weighted.
Ministry of Business Innovation and Employment	Strategic Science Investment Fund Infrastructure (SSIF Infra)	\$ 62,859	Negotiated	65	Variable	50	50	Assessed based on seven performance indicators, including excellence and impact – it is assumed they are evenly weighted.
Ministry of Business Innovation and Employment	Talent and Science Promotion	\$ 6,070	Institutional	20	Annual	0	100	?
Ministry of Foreign Affairs and Trade	Antarctic Research and Support	\$ 21,320	Negotiated	55	Variable	62.5	37.5	No assessment criteria available, but this appropriation has been assumed to have similar weightings to the CoRES as the Antarctic Platform is similar in its nature.
Royal Society	Marsden Fund	\$ 78,545	Contestable	100	3 years	100	0	Assessed based on following criteria, all assumed to relate to excellence: - Scholarly impact, ability and potential, - Development of research skills, and - Vision Mātauranga (where applicable).
Tertiary Education Commission	Centres of Research Excellence (CoREs)	\$ 74,700	Contestable	95	Annual	62.5	37.5	Assessed based on four criteria: - Research excellence (40%), - Tertiary education system (20%) (assuming this contributes 50/50 to impact and excellence), - NZ's future development (20%) (assuming this contributes to impact weighting), and - Governance and management (20%) (we have not included this in our ratio).

Tertiary Education Commission	Performance Based Research Fund (PBRF)	\$ 321,000	Institutional	20	Annual	100	0	Scoring based entirely on excellence; however excellence as defined by PBRF includes an element of "impact". A recent review of the PBRF has called for increased emphasis on impact in coming years.
Tertiary Education Commission	Proposed Performance Based Research Fund (Future PBRF)	\$ 421,000	Institutional	20	Annual	80	20	PBRF review called for \$100 M more per annum and greater focus on impact.
Tertiary Education Commission	University-led Innovation	\$ 10,000	Negotiated	70	Annual	10	90	Looking to drive: - cutting-edge research to grow New Zealand's competitive advantage - university-led innovation and entrepreneurship in existing and emerging industries - commercially-relevant research that fosters industry collaboration and strengthens economic growth, and - stronger connections with overseas universities, research institutes and businesses. Assumed that most of these criteria related to impact, although there's an element of excellence in nature of research being "cutting-edge".