Te Tāhuhu o te Mātauranga
Ministry of Education

Report: TREF Phase 2 – Overall design direction

To:	Hon Dr Shane Reti, Minister for Universities Hon Penny Simmonds, Minister for Vocational Education		
Date:	4 December 2025	Deadline:	15 December 2025
Security Level:	In-Confidence	Priority:	Medium
From:	9(2)(a) Policy Director	Phone:	
Drafter:	9(2)(a)	METIS No:	1353106
Title ok for release?	Yes		

Why are we sending this to you?

- This paper provides high level options for Phase 2 of the design of the Tertiary Research Excellence Fund (TREF). Your decisions on the way forward for universities and non-universities will determine the detailed design options for the TREF, which we will provide advice on ahead of final recommendations to Cabinet in April 2026.
- This paper follows advice regarding Phase 1 decisions on the TREF provided in October 2025. That advice covered changes to the External Research Income component [METIS 1353422] and a research component of investment plans [METIS 1352338].

What action do we need, by when?

- We are seeking your direction on the approach to Phase 2 of the design of the TREF. Your decisions will inform further advice to you in January 2026.

Key facts, issues and questions

- The Ministry of Education (the Ministry) and the Tertiary Education Commission (TEC) have conducted extensive engagement to understand the design implications for replacing the Quality Evaluation (QE) with new metrics in the TREF.
- There are options about the metrics component design for universities which are dependent on your overall objectives for the fund. Different metrics, or a different balance of metrics, will meet some objectives better than others.
- Choices are more limited for the metrics for polytechnics, Wānanga and private training enterprises. This is due to their smaller volume of research activity and diversity of research missions.

Proactively Released

Alignment with Government priorities

1. This report aligns with the Government's Going for Growth agenda, which informed the new Tertiary Education Strategy and recent Cabinet decisions in response to the University Advisory Group. This includes ensuring the higher education sector delivers knowledge, skills and research that will drive productivity and economic growth.

Background

2. On 25 August 2025, Cabinet approved Minister Reti's proposals to replace the Performance-Based Research Fund (PBRF) with a new research capability fund called the Tertiary Research Excellence Fund (TREF) and to strengthen the research component of tertiary education organisations' (TEO) investment plans [CAB-25-MIN-0285 refers].
3. Cabinet agreed that the TREF would retain the existing Research Degrees Completion (RDC) component and an External Research Income (ERI) component, with new weightings to incentivise user-funded research. The existing Quality Evaluation (QE) component will be replaced with a new metrics component. Cabinet agreed that the metrics component would be centred on a citation-based metric and a small number of supplementary metrics, and that modifications to this approach would be considered for non-university TEOs.
4. Cabinet invited you to report back in April 2026 to confirm the detailed fund design including the metric component design for universities, how the metric component would be allocated to non-university TEOs, and overall component weightings.
5. Phase 1 of the TREF design is now complete. You have both approved the new ERI component design and the new investment plan requirements on 10 October 2025 [METIS 1353422 and 1352338 refer], and these will now be implemented as changes to the existing PBRF design through the 2027 Funding Determination and through the Plan Guidance for 2027 (which is released next year). We have provided Minister Reti's office with letters to send to participating TEOs outlining these changes.
6. Officials are now focussed on Phase 2 of the TREF design, which includes the design of the citation-based metric and complementary metrics for universities, and the approach for allocating the metric component funding for non-university TEOs.
7. To support this work over the past two months we have engaged with TEOs who participate in the PBRF and received advice from a Technical Advisory Group (TAG) which was convened to provide input into the development of the citation-based metric for universities [METIS 1350441 refers]. A summary of sector feedback is outlined in **Annex 1**. We have also met with officials from the Australian Research Council, bibliometrics experts at the University of Leiden (publishers of the Leiden Ranking) and researchers at the Research on Research Institute, which investigates global research assessment best practice. Their insights have informed the approaches we have developed.

Structure of this paper

8. This paper presents the Ministry's conclusions based on extensive engagement on the design of the metrics component of the TREF and on analysis of available data. It uses those conclusions to present three approaches that could be followed when designing the TREF. Each approach has strengths and trade-offs, and a different approach could be followed for universities and for non-universities.
9. Your feedback is sought on which approach (or approaches) you prefer. Officials will develop more detailed advice in accordance with your preferences for final decisions next year.

Proactively Released

Insights from Phase 2 engagement and analysis

Limitations and implications of a field-weighted citation metric

10. Cabinet agreed, on the advice of the University Advisory Group, to replace the QE component of PBRF with a field-weighted citation metric (FWCM). A FWCM counts numbers of research citations for a given research output type and normalises that count relative to the average for the subject area and time period following publication. FWCMs are generally accepted across the global research community as a proxy for research quality and academic impact, albeit with some important disciplinary and methodological limitations which we discuss below.

11. A FWCM requires a database of research output metadata in order to compare the citations gained by any given output against as many other outputs within the same field of research as possible. While the choice of database would ultimately be an operational decision for the TEC, likely involving a competitive procurement process, it is important that the overall design of the fund is informed by an understanding of the strengths and limitations of the databases available. This has been a focus of the TAG's analysis.

12. Although no database has 100% coverage of all research, the available options each prioritise accuracy, inclusivity, and cost in different ways and the decision about which database to use will introduce constraints on other elements of the FWCM design.

13. Commercial databases prioritise accuracy and reliability over breadth of coverage to different degrees. They come with substantial direct access costs for TEC and providers but generally lower compliance costs and complexity. There is also a risk of becoming locked into a specific commercial arrangement.

14. An open-source database (OpenAlex, which was recommended by the University Advisory Group, is currently the only option) could offer greater coverage and lower direct costs while avoiding vendor lock-in and aligns with the global shift towards open research. Australia is exploring this approach for research evaluation (not for funding allocation). There was strong support for open access to research across our stakeholder engagement. There was limited practical familiarity with OpenAlex but there was enthusiasm for the future potential of this platform. Nonetheless, we are not confident that it would be sufficiently robust to be utilised in 2027. There are no known jurisdictions that currently use OpenAlex data as a method for funding distribution, although we understand Germany may be exploring this.

15. Feedback from the Research on Research Institute and the University of Leiden indicated that OpenAlex has international sector support and is likely to be more viable in the future, but is not currently sufficiently robust to use without substantial work to improve accuracy, and any curation or audit process could cancel out the lower direct costs. However, if the FWCM is to make up a large share of TREF funding then the costs associated with data curation may be warranted. Open-source databases such as OpenAlex are also likely to develop significantly over time and there may be opportunities to move from a commercial platform in future.

16. A bespoke solution (such as the TEC building its own database), as used in Norway, could be tailored to New Zealand's needs but would be highly complex and costly for TEC and providers, with little sector support. It risks higher compliance costs than the other options and much greater set-up costs, though this has not been explored in depth.

17. Engagement with the sector as well as our own analysis has emphasised the likelihood of coverage gaps, regardless of which database is used. As an example, we compared the University of Waikato's record of research outputs with both Scopus (commercial) and OpenAlex (open-source). Both databases captured between 85-95% of journal articles

Proactively Released

produced from Waikato, both only 40% of book chapters, and approximately 25% of conference proceedings. Neither included any creative works.

18. This means that fields of research that feature fewer journal articles and more of the other types of research output will be underrepresented by a FWCM. Using data from the 2018 QE round we assess the greatest impact to be on the following fields:

    - Creative and Performing Arts, where 55% of the research outputs in the QE were creative works which neither database captures.
    - Humanities and Law, where 23% of outputs in the QE were book chapters, and 5% were books.
    - Engineering, Technology and Architecture, where 21% of outputs in the QE were conference proceedings.

Complementary metrics

19. Cabinet's decisions on the TREF design acknowledge the need to supplement a FWCM with complementary metrics that could both address the coverage gaps discussed in paragraph 18 and incentivise research impact and knowledge exchange activity, particularly aligned with economic growth.

20. We have explored a range of possibilities with the TAG and the sector, which are summarised in **Annex 2**. The key takeaway from our engagements and analysis is that complementary metrics can partially reflect existing impact and knowledge exchange activity, and therefore research aligned with such outcomes, but cannot address the coverage gaps in the arts and humanities. Complementary metrics, if used, are likely to only make up a small portion of TREF funding, which means they are unlikely to move the needle significantly for universities. Their key value is in the signal they send by incentivising a particular high priority activity such as commercialisation.

21. Commercial impact metrics considered include industry co-authorship and a range of commercialisation metrics from the SCOPR[1] survey. There is no support for industry co-authorship, and limited support for commercialisation metrics given the dominance of the University of Auckland and the smaller amount of measurable activity elsewhere in the system. However, both the Ministry and MBIE consider there could be value in metrics to incentivise institutions to support commercialisation activity, given the recent changes to intellectual property policy requirements could disincentivise universities from investing. This would need to be balanced elsewhere in the fund to avoid duplicating incentives to prioritise research that will already benefit from an FWCM.

22. Broader impact and knowledge exchange metrics considered include policy citations, a grey literature output measure (i.e. information published outside traditional academic journals), or open access publication rates. There is support across most universities for a policy impact metric, with grey literature considered likely the most accurate. The universities indicated that while this would involve set-up work on their part to ensure this data could be collected, they are willing to invest that effort as they see wider benefits. There is also some support for an open access metric as a way to incentivise open publishing behaviours among academics.

23. Metrics that would address the FWCM coverage gaps and support a broad research base include public engagement data, subject-based output weightings, or an early career researcher metric. We heard that public engagement and subject-based weightings are too open to gaming and too compliance-heavy for universities. There were mixed views about an early career metric, with Auckland University of Technology being extremely supportive,

[1] Survey of Commercialisation Outcomes from Public Research, conducted annually by Knowledge Commercialisation Australia.

Proactively Released

while University of Auckland and Victoria University Wellington were concerned about the compliance costs associated.

24. Alongside any complementary metrics, changes to the subject-based weighting in the Research Degrees Component was proposed and supported across the sector as the most viable way to address the subject coverage gaps in the FWCM, and we recommend exploring this in addition to any of the approaches Ministers indicate support for. It appears to be the only measure which can address these gaps for the arts and humanities in particular.

Metrics for non-university providers

25. Wānanga, polytechnics and PTEs emphasised the critical role research plays in their activity, the opportunities from strengthening their research capacity, and some of the challenges faced to retain this critical capacity in smaller organisations.

26. There was a split between approaches suggested by TEOs. Most recognised that a simpler proxy for research performance with associated low compliance costs would be beneficial, while others advocated for complex metrics or qualitative approaches that would allow them to illustrate research impact or quality even if such metrics would not yield a greater share of funding for them.

27. The latter view appeared to be driven by two factors:

- a concern that there would be a reputational risk to their degree-level and above provision if university research funding was seen to incorporate more robust quality measures than those for non-universities, and
- a desire to see metrics reflect a particular area of strength for their organisation.

28. Historically, non-university TEOs have only received a very small part of PBRF funding, reflecting lower research activity, subject area funding weightings, the design of the PBRF not rewarding their research types, and some opting out due to compliance costs. Any new reporting requirements will impose proportionately higher costs, reinforcing the case for simpler allocation mechanisms.

29. Officials investigated a range of metrics that would reflect research impact, quality[2] or community engagement. The key challenge is that the non-university grouping is diverse in terms of size, focus, priorities and research output. An approach appropriate for some of the larger institutions with dozens or even a hundred research-active staff would not be appropriate for those that have fewer than ten. Approaches that recognise these differences in scale would significantly increase the complexity of administering the TREF, for limited accuracy gain, and would not be appropriate for the quantum of funding being allocated.

30. There was general – though not universal – support for metrics that are proxies for research capacity, such as the number of active researchers. Such metrics would scale fairly with organisation size and incentivise TEOs of any size to maintain their research capacity. The new research component of investment plans agreed in Phase 1 of the TREF changes [METIS 1352338 refers] is a key component that can mitigate the concerns that simpler research capacity metrics do not incorporate research quality or impact. Many dimensions raised by stakeholders, such as research outputs, research mission and impact will form part of plan content requirements, helping address any gaps in metrics.

[2] This does not include the RDC and ERI measures currently in place, which are recognised as measures of research quality.

Proactively Released

Interaction between funding metrics for universities and other TEOs

31. While Ministers have agreed that the TREF should be open to all providers, a key design question we have discussed with providers is the extent to which universities and other tertiary providers should compete within the same funding pools. Having different metrics for universities and non-universities raises the question of how this component should be allocated between the types of TEOs. There are two broad approaches that we have discussed with stakeholders:

9(2)(g)(i)

- A 'dynamic' approach, whereby all providers compete for the same pool of metric component funding, albeit on the basis of different metrics. This would require establishing a mathematical relationship between the different metrics used.
- A 'fixed pools' approach, whereby the metrics component is split into two pools, one for universities and one for non-universities. Providers would only compete for funding within their pool.

The key trade-off between the two approaches is the degree of flexibility versus stability that it would provide. Under a 'dynamic' approach, the relative share of metrics component funding that the different sectors could earn could change based on their relative performance – i.e. non-university providers' total share of this funding would grow or shrink depending on how the sector performs on their metric/s relative to how universities are collectively performing on their metric/s.

33. On the other hand, it would significantly increase the complexity of the design and administration of the fund, and would create a risk of unintended consequences (particularly if we are using metrics based on information we do not currently collect and may not be able to accurately model).

34. Feedback has generally favoured a fixed pools approach to the metrics component, because any choice about how university and non-university metrics would be translated (e.g. how a university FWCM translates to a research capacity-based measure) would be arbitrary and likely lack credibility with the sector. Stakeholders noted that the size of a fixed pool should be adjusted periodically to reflect changes in relative research performance across the system.

9(2)(g)(i)

35. A small number of TEOs objected to any 'ringfencing' between universities and other providers, due primarily to parity of esteem concerns. They argued in favour of universities and other providers all being included within a single metric or group of metrics, which would not be viable if a FWCM is to be used for universities.

36. Some stakeholders suggested that the whole fund should be split (including RDC and ERI) as this would guarantee a particular share of funding without needing to compete against universities. This was not a widespread view and would be inconsistent with Ministers' earlier decisions. Te Wānanga o Aotearoa suggested splitting funding further between Wānanga and ITPs/PTEs as this would recognise their unique statutory role. This would, however, create very small funding pools and result in the Wānanga competing in a zero-sum game against one another.

37. Under either approach, Ministers would need to determine the starting 'share' of universities versus other providers for the metrics component, either by agreeing the method to mathematically translate one set of metrics to the other (dynamic), or by directly setting the size of the two pools (fixed pool). The basis for doing so would necessarily be a judgment, with relative past performance on the Quality Evaluation exercise providing one, albeit imperfect, input into this decision.[3]

[3] Noting that this data is from 2018 and does not reflect changes in either performance or the number of providers who might choose to participate in TREF compared to PBRF.

Proactively Released

Overall design of the TREF

Objectives for the TREF

38. The overall design of the TREF should be guided by clear objectives for what the fund is intended to achieve, including the relative priority of those objectives. In our earlier advice on replacing the PBRF we suggested objectives for the reforms, which we have updated below as suggested objectives for the TREF:

    - Efficiently allocate research capacity funding, minimising compliance costs and complexity for providers and government
    - Incentivise TEOs to support excellence across the breadth of research activity in New Zealand
    - Incentivise research activity which leads to economic and social benefits for New Zealand
    - Supporting TEOs to deliver quality research-led teaching
    - Supporting research capacity and capability including a future-proofed research workforce
    - Supporting the Crown's Te Tiriti obligations, including the active protection of mātauranga Māori
    - Providing robust and useful public information on research performance.

39. Some of these objectives are necessarily in tension with each other. In the high-level approaches outlined below, we explore the implications of prioritising different objectives.

40. We also acknowledge that the Crown has a shared responsibility with iwi in the active protection and transmission of mātauranga Māori. Mātauranga Māori research - research committed to reclaiming, enriching, and enhancing the Māori knowledge ecosystem - is undertaken across the tertiary education system. Wānanga are the legislated kaitiaki of mātauranga Māori in the tertiary system. We consider that any approach should ensure that:

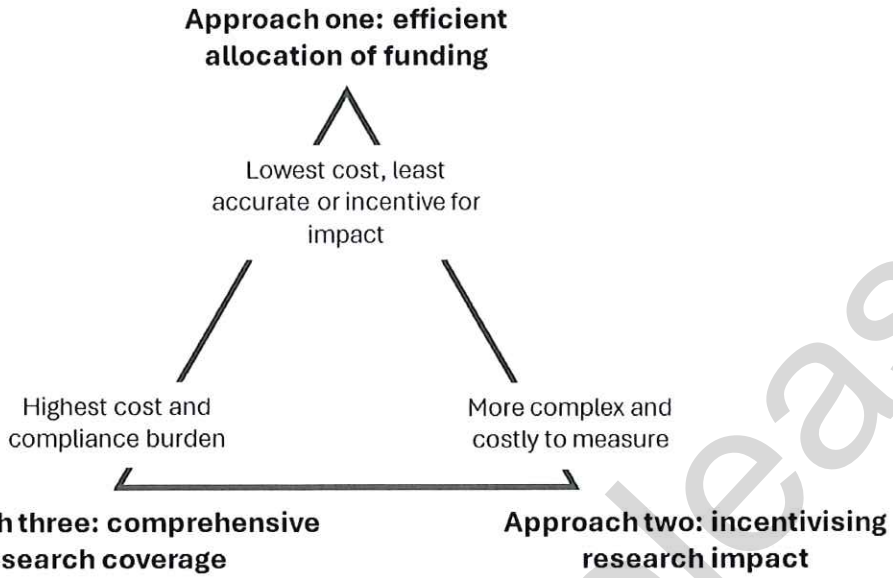
    - staff undertaking mātauranga Māori research across the system are not disadvantaged to do so; and
    - Wānanga are not disadvantaged in the allocation of funding if they focus more on their legislated kaitiaki role than other research opportunities.

The overall design of the TREF will depend on which objectives you want to prioritise

41. Based on the potential objectives and principles we have identified above, as well as what we have heard from the sector, independent experts, and the TAG, we think there are three core approaches to the metric component design for universities. Each approach is based on a different key trade-off:

42. These approaches can be considered as degrees of emphasis rather than discrete packages that completely exclude the other options.

Proactively Released



43. The following section summarises the key design features, system implications and impacts of each approach. **Annex 3** also includes an A3 table summarising the three approaches.

44. It is important to consider how the design of the fund as a whole will best meet your overall objectives, rather than just individual components. Approaches to the metrics component can be offset or balanced by a higher or lower weighting for the ERI or RDC components. For example, a higher ERI component supports an approach that incentivises research impact, while a higher RDC component would support some research areas that are not well represented in a FWCM without additional compliance costs.

45. The design settings under each approach are only intended to be broadly indicative, and any approach would need to be refined following Ministers' feedback. Variations on these approaches could also be considered, particularly if Ministers would like to balance two or more of the objectives. It would also be possible to progress different approaches for universities and other providers.

Approach One – prioritising efficient allocation of funding

	Metric component (universities)	Metric component (non-universities)	External Research Income	Research Degree Completions
Component design	Unmodified FWCM (75% or more of the metrics component) A single complementary metric based on existing data (worth less than 25% of the metrics component).	Single capacity metric e.g. research-active staff FTE or staff contracted research time	As agreed in METIS 1353422.	Possible consideration of changes to subject weightings to reflect gaps in FWCM
How much the components are worth	Significantly lower than QE, recognising lower accuracy and breadth (e.g. 10-30%). University/non-university split discussed below.		30-45%	30-45%

Proactively Released

Design of this approach:

- Prioritises minimising compliance costs over other fund objectives, within the broad framework agreed by Cabinet.
- For universities the metric component would comprise:
  - A simple and unadjusted FWCM based on an existing dataset, worth approximately 75% of the metric component total
  - A single complementary metric for universities, based on data that is already collected, worth no more than 25% of the metric component total. Likely options for this metric are either a combined commercialisation measure based on a number of intellectual property outputs, an open access publication rate, or a policy impact metric.
  - Compliance and implementation costs could be further reduced by relying on a FWCM only (i.e. no complementary metrics), although this would be a change to the broad approach agreed by Cabinet.
- For non-university TEOs the metric component would consist of a single metric of research capacity:
  - Viable options include the number of full-time equivalent active-research staff or the total amount of contracted staff research time.
  - We considered degree level completions as a metric, but do not recommend this approach. While it would cost the least to implement it would shift funding significantly towards providers with significant degree-level delivery in areas not usually research intensive (particularly nursing).
- Under this approach, the metrics component should make up a smaller proportion of total funding compared to the current QE component weighting of 55%, reflecting its lower accuracy and breadth of coverage. ERI and RDC would both be worth more.

9(2)(g)(i)

9(2)(g)(i)

9(2)(g)(i)

System implications:

- Would reduce cost and complexity for participating TEOs and for the Government relative to the PBRF and to other options, but allocate funding on a less comprehensive basis and would create fewer direct incentives for research impact.
- May incentivise universities in particular to deprioritise research areas that do not perform as well under a simple FWCM and/or do not tend to bring in external research income (depending on the additional metric chosen).
- A capacity metric for non-universities would not directly incentivise research impact or broader research outcomes, but would recognise the resource that a provider is dedicating to research and support those providers who are prioritising research-led teaching.
- TREF measures would produce less fine-grained information about the research system, although could be supplemented by the broader research performance monitoring system that Minister Reti has agreed should be developed.
- TEC's investment plan requirements would provide a mechanism for providers to outline their research plans and how they will deliver research impact.

Impacts on funding allocations:

- Universities:

Proactively Released

9(2)(g)(i)

- Initial analysis indicates this approach will likely benefit universities with more research in areas such as physical and biological sciences, medical research and, to a lesser extent, social sciences research.
- Concrete modelling of complementary metric options alongside the FWCM is not possible at this stage, so impacts are difficult to predict for many options.
  - SCOPR data suggest that Auckland would be the overwhelming beneficiary of any commercialisation metric: it reports over 50% of the total volume across 11 SCOPR metrics.
  - Open access rates on the other hand are more evenly distributed across the eight universities.
- These impacts can be mitigated by increases in the ERI and RDC weightings
- In addition, changes to subject area weightings in the RDC component could be considered to address coverage gaps in the FWCM.

- Non-university TEOs
  - TEOs that performed less well under the QE relative to their size – i.e. they had a lower proportion of higher graded evidence portfolios (quality), and in general had fewer evidence portfolios per research FTE (intensity) – would receive more funding under a pure research capacity metric compared to the QE.
  - Overall impacts would be highly dependent on the split of the metrics component between universities and other TEOs, noting that non-university TEOs tended to receive relatively little funding from ERI and RDC.

Alignment with Treaty of Waitangi obligations:

- May underrepresent mātauranga Māori in universities, and this could be considered a breach of the Crown's obligation of active protection.
  - If Ministers chose to progress this approach, we would suggest mitigating this risk by including an explicit expectation around supporting mātauranga Māori in future evolutions of the research component of investment plans.
- Metrics for non-universities based on research capacity – such as number of active researchers – do not create strong incentives in either direction in relation to mātauranga Māori research activity and are aligned with the Crown's responsibilities.

Approach Two – focus on incentivising research impact

	Metric component (universities)	**Metric component (non-universities)**	**External Research Income**	**Research Degree Completions**
Component design	Unmodified FWCM worth a lower proportion of the metrics component – e.g. 50% Two impact and knowledge-exchange metrics with combined value of approx. 50% the metrics component	3-4 impact, knowledge-exchange or community engagement metrics (not recommended)	As agreed in METIS 1353422.	Possible consideration of changes to subject weightings to reflect gaps in FWCM

Proactively Released

How much the components are worth	Lower than QE but higher than Approach One, recognising greater complexity and accuracy of the metric (e.g. 25-40%).	30-45%	20-30%

Design of this approach:

- Prioritises 'system steering' and alignment with the Going for Growth agenda and Government priorities in the Science, Innovation and Technology (SIT) system over other fund objectives.
- For universities the metric would comprise:
    - A simple and unadjusted FWCM based on an existing dataset, worth approximately 50% of the metric component.
    - Two research impact complementary metrics, collectively worth approximately 50% of the metric component.
        - Likely options for research impact metrics are a commercialisation metric based on SCOPR data or similar, and a policy impact metric based on grey literature outputs or research citations in government policy.
        - These options involve trade-offs between compliance costs, degree of robustness, and university preferences. Further analysis and advice would be required.
- For non-university TEOs the metric would comprise a bundle of impact-based metrics, focussed on knowledge exchange and community engagement.
    - Our engagement has not identified any strong contenders for these sorts of metrics that would be sufficiently robust to provide a basis for funding allocation.
    - If Ministers want to prioritise this approach for non-university TEOs we would undertake further work to investigate options suggested by stakeholders (e.g. engagement or impact case study metrics), noting that these would likely be very complex for the amount of funding to be allocated.
- Under this approach the metrics component would have a smaller weighting than the current QE weighting of 55%, but a higher weighting than under Approach One.
    - This would recognise the reduced accuracy relative to the QE but would incentivise impact activity. The ERI component weighting would be increased to further incentivise research commercialisation. RDC would remain unchanged.

System implications:

- Will incentivise research impact and knowledge exchange activity. Universities may be incentivised to shift focus away from subject areas that are less well-aligned to this approach.
    - As such, there is a risk that it could duplicate changes elsewhere in the SIT system, rather than complement them, with implications for breadth of research and research-led teaching in the system.
- Compliance costs for TEOs and TEC cannot be accurately modelled as this depends on metrics selected and whether there is existing data, but they will certainly be greater than Approach One. If the exercise were run less frequently, compliance costs could likely be reduced.
- A commercialisation metric may counterbalance recent changes in intellectual property policy which could otherwise disincentivise universities from investing in this area. MBIE is supportive of this approach for that reason.

Proactively Released

Impacts on funding allocations:

- An unmodified FWCM will likely benefit research in areas that perform well in citations, and modelling suggests the largest benefit would be to the University of Auckland. These effects would be diluted by the lower internal weighting of the FWCM relative to the complementary metrics.
- The effect of impact-based metrics on funding allocations cannot be modelled as we do not hold the relevant data. Initial analysis suggests they are likely to benefit health and biomedical research, as this is where the universities tell us the bulk of their non-academic collaboration and contract research occurs.
- This approach is therefore unlikely to address the gaps in arts and humanities research coverage in the citation metric. As previously noted any commercialisation metric will likely strongly benefit the University of Auckland.

Alignment with Treaty of Waitangi obligations:

- Can be aligned with the Crown's active protection obligations if metrics recognise broader economic and social impacts from research. However, closely targeted measures that reward a specific type of activity could be viewed as a breach of the Crown's obligation if they ultimately divert funding away from mātauranga Māori research.
- How research impact is defined and measured is key determinant of whether this approach aligns with the Crown's obligations.

Approach Three – Comprehensive research coverage

	Metric component (universities)	**Metric component (non-universities)**	**External Research Income**	**Research Degree Completions**
Component design	Modified FCWI including supplementary data with weighting of approx 2/3 3-4 complementary metrics that address gaps in citation metric and research workforce with total weighting of 1/3.	3-4 metrics focussed on research capacity and subsector unique characteristics	As agreed in METIS 1353422.	Possible consideration of changes to subject weightings to reflect gaps in FWCM
How much the components are worth	Closer to QE (e.g. 40-50%)		Approx. 20%	30-40%

Design of this approach:

- Prioritises accuracy and research inclusivity over other objectives, and aims to ensure future research capability pipeline, and a broad research base.
- For universities the metric would comprise:
    - An FWCM based on a broader modified dataset that had been cleansed for accuracy, and potentially supplemented with additional output data provided by the universities (for example, in specific disciplines poorly represented by the core

Proactively Released

dataset). The FWCM would have a larger internal weighting than Approach Two, recognising the increased accuracy, inclusivity, and compliance costs for TEOs.

  - A small number of complementary metrics covering:
    - Arts and humanities research
    - Early career researchers
    - A research integrity or research environment measure such as open access publication rates

- Decisions about which metrics to use involve trade-offs between compliance costs, degree of robustness, and university preferences. Further analysis and advice would be required.
- For non-universities the metric would comprise:
  - A combination of research capacity and performance metrics, such as academic staff count, contracted research time, research output metrics, or potentially a research partnerships metric.
  - As with Approach Two, this could create significant compliance costs for providers and the TEC relative to the value of allocations. The data would be more robust, so this option is viable if Ministers wish to pursue it, but careful design would be needed.
- Under this approach, the metrics component would have a weighting closer to the current QE component, reflecting the greater accuracy and inclusivity and the higher compliance costs relative to the other approaches.
- ERI would remain weighted at 20%, and the RDC measure would increase slightly from 25% to 30-40%. This reflects a focus on disciplinary coverage and research workforce development.

System implications:

- Will incentivise TEOs to maintain broad research capability and to continue to invest in the future research workforce.
- Would complement Government priorities for the SIT system rather than replicating them.
- Would have the highest costs for TEOs and the TEC, and depending on the metrics chosen this could equal or exceed costs associated with the QE. We would report back on the likely costs of a particular approach, if required, as we would not recommend Ministers favour an approach with greater compliance costs than the current QE.

Impacts on funding allocations:

- Initial analysis suggests this approach will likely produce more balanced allocations for universities. We do not have the necessary data for modelling, but an FWCM that is based on more inclusive and accurate output data would likely reflect current QE allocations.
- As with the impact approach, we are not able to model the impacts of all potential complementary metrics. However, Open Access publishing proportions across the eight universities would produce a more evenly distributed outcome than other metrics (for example commercialisation data). Early career researcher proportions (based on 2018 data) indicate a range between 11% of academic staff at Massey University and 24% at University of Canterbury.

Proactively Released

- Impacts for non-universities cannot be modelled in the absence of relevant data, but as with Approach One, it is possible that TEOs who performed less well in the QE could receive more funding under this approach (and vice-versa). Very small TEOs without the research infrastructure to support research capability investment may also be disadvantaged.

Alignment with Treaty of Waitangi obligations:

- This approach has the greatest potential to capture all types of research, including mātauranga Māori, and so is most likely to ensure equitable treatment and active protection of mātauranga Māori research (although this is dependent on how different types of research are weighted).
- This approach would need to balance the impact of higher compliance costs on some smaller institutions, such as Wānanga.

High-level conclusions

There are advantages and risks to all three options for the universities…

46. Each of the three options should be considered as degrees of emphasis rather than discrete packages. They each respond to different prioritisations of the fund objectives, and while each are viable, they have risks which are more pronounced at the more 'extreme' end of the spectrum. The lower-cost option carries the greatest risk of being too blunt and under-representing activity that should be rewarded, while both the research and the comprehensive option could create large compliance costs.
47. Depending on what your priorities are for the fund, as well as the extent to which you would like to see shifts in funding allocations, we can design approaches which look to reduce these risks. This could include the frequency with which funding is allocated to reduce compliance costs.
48. Nonetheless, the Ministry prefers approach one, on balance, given that the investment required from both the TEC and TEOs under approach two and three are not guaranteed to generate a meaning improvement in accuracy, coverage and incentives.. There are other avenues – most notably the research component of investment plans – that can be utilised to incentivise broader system shifts and maintain expectations of research diversity.
49. If Ministers want to prioritise minimising compliance costs, the TEC suggests considering using an unmodified FCWM alone without an additional complementary metric. Their view is that complementary metrics would come at a cost to implement and may not significantly improve accuracy. While we acknowledge these concerns, on balance we still consider that a small number of complementary metrics to be worthwhile to give a better balance of incentives to universities.
50. MBIE would prefer a fund design closer to approach two, with a greater focus on complementary metrics that support research impact, including but not limited to commercialisation.
51. The universities are generally supportive of a low-cost approach, although there are concerns about arts and humanities coverage and mixed views about appropriate additional metrics. The University of Auckland, in particular, is very supportive of an FWCM-only approach, if additional metrics do not change funding allocations meaningfully.

…but the research impact option is not a viable approach for the non-universities

52. As we noted in the analysis above, the variation in provider size, focus, and types of activity and output produces across the participating non-university TEOs creates some challenges in designing appropriate metrics.

Proactively Released

53. We do not consider that metrics based on research impact or knowledge exchange activity, or more qualitative approaches are appropriate for the purposes of allocating TREF funding to non-university TEOs because for many of these providers the volume of activity will be too small to be robust or the compliance costs will be disproportionate to the value of the fund.

54. For this reason, we recommend an approach for the non-universities based on a single research capacity metric (likely staff-based, output based or a combination of both) for the metric component. There are options to work through to mitigate the concerns of some stakeholders and officials can incorporate this into the next piece of advice.

55. If Ministers are interested in pursuing a more comprehensive approach for the non-universities officials can report back with the most favoured three to four metrics for further consideration.

TREF participation settings

56. TEOs must currently meet the following criteria to participate in the PBRF:

- be based in Aotearoa New Zealand
- have degree-granting authority
- have an approved Investment Plan
- receive Delivery on the New Zealand Qualifications and Credentials Framework (DQ) funding from the TEC

57. Eligible TEOs can choose to participate in the PBRF by submitting research in the QE. However, if their submitted research is assessed as of fundable quality, then they must also participate in the RDC and ERI components, even if their funding entitlement in those components is zero. TEOs cannot choose to participate in only the ERI or RDC components.

58. This has meant that in the past, participation in the QE process has been the key threshold, with a number of eligible TEOs determining that the that the effort required to make a submission was not worth the likely returns.

59. The replacement of the QE round with the metrics component would reduce compliance costs and therefore lower the barrier to participation in the TREF, potentially opening it up to a larger number of degree-granting TEOs. While this is broadly a desirable consequence of the reforms, it does risk diluting the overall funding pool at the margins and we would want to ensure that access is still limited to providers delivering quality research and research-led teaching.

60. The TEC's additional expectations on the research component of investment plans for participating TEOs [METIS 1352338 refers] may mitigate this risk. The TEC would need to be satisfied that all participating TEOs have a robust plan for building their research capacity and impact that meets the expectations set by the TEC. We will provide further advice on this issue as part of our next advice.

Next steps

61. Officials will develop more detailed proposals for your consideration at the end of January 2026 based on your preferences in this report, seeking approval to undertake targeted consultation with the sector in February 2026. We plan to seek final decisions in early March 2026 ahead of Cabinet decisions in late April 2026.

Proactively Released

Timeline for 2026 advice	Approximate Date
Detailed proposals	29 January 2026
Stakeholder consultation	9-27 February 2026
Final decisions paper	Early March 2026
Draft Cabinet paper for approval	Early April 2026
Ministerial consultation	Early to mid-April 2026
Social Outcomes Committee	~ 22 April 2026
Cabinet	~ 27 April 2026

Implementation timing

62. Cabinet decisions in April 2026 will inform the funding determination for consultation with the sector ahead of finalisation in June or July 2026. This is earlier than it is normally developed due to the pre-election period. That funding determination to TEC will do two key things:
    - Introduce the new ERI weightings agreed previously in METIS 1353422 and METIS 1354163. These will apply to income earned by TEOs in 2026, reported in 2027 and inform funding paid by TEC in 2028.
    - Set out the key policy and design perameters for the Phase 2 changes, providing the TEC with the basis to progress the operational design of the TREF, including any procurement processes.
63. We anticipate that the operational design would be completed in 2027, reported on in 2028 and inform funding paid by TEC in 2029.

Annexes

Annex 1: Detailed outline of stakeholder feedback

Annex 2: Detailed analysis of metrics options for universities and non-universities

Annex 3: Summary of options A3

Proactively Released

Recommended Actions

The Ministry of Education recommends you:

a. **note** that the Ministry has conducted stakeholder engagement with universities, wānanga, polytechnics and private training establishments to discuss potential approaches for a metrics component.

Hon Dr Shane Reti	Hon Penny Simmonds
Noted	**Noted**

b. **Discuss** your priorities for the design of the TREF and how they align with the three approaches outlined in this paper

Hon Dr Shane Reti	Hon Penny Simmonds
Yes / No	**Yes / No**

c. **Indicate** which approach/es you would like the Ministry to further develop as part the next phase of advice on the TREF (circle more than one to indicate a balance of two approaches):

Universities	**Hon Dr Shane Reti**	
1. Efficient allocation of funding	Preferred / not preferred	
2. Incentivising research impact	Preferred / not preferred	
3. Comprehensive research coverage	Preferred / not preferred	
Non-university TEOs	**Hon Dr Shane Reti**	**Hon Penny Simmonds**
1. Efficient allocation of funding	Preferred / not preferred	Preferred / not preferred
2. Incentivising research impact	Preferred / not preferred	Preferred / not preferred
3. Comprehensive research coverage	Preferred / not preferred	Preferred / not preferred

d. **agree in principle** that the metrics component should be split into separate funding pools for universities and non-universities.

Hon Dr Shane Reti	Hon Penny Simmonds
Agree / disagree	**Agree / disagree**

9(2)(g)(i) [redacted]

Proactively Released

e. **Note** that officials will provide further advice on the detailed design of your preferred approach in January 2026.

Hon Dr Shane Reti	Hon Penny Simmonds
Noted	**Noted**

Proactive Release:

f. **agree** that the Ministry of Education release this paper once the final design of the new Tertiary Research Excellence Fund has been finalised with any information needing to be withheld done so in line with the provisions of the Official Information Act 1982.

Hon Dr Shane Reti	Hon Penny Simmonds
Agree / Disagree	**Agree / Disagree**

9(2)(a)

	Hon Dr Shane Reti	Hon Penny Simmonds
Policy Director, Tertiary Policy **Te Pou Kaupapahere**	**Minister for Universities**	**Minister for Vocational Education**
04/12/2025	__/__/____	8/12/25

9(2)(g)(i)

Proactively Released

Annex 1: Additional information on sector feedback

1. Officials conducted extensive sector engagement to inform the options presented in this paper. The Ministry shared initial thinking with the sector stakeholders and encouraged further feedback via email on the options discussed. The key elements of our engagement were:

    - Technical Advisory Group (TAG): The TAG has met for five two-hour workshops over the last two months to discuss the FWCM and complementary metrics for universities.
    - Universities: Officials met Deputy Vice Chancellors – Research on 12 November 2025 and held two three-hour workshops with representatives present from all eight universities. This was followed by written feedback from most universities.
    - New Zealand Institute of Skills and Technology: Officials met with representatives from seven of the NZIST business units and the Director – Rangahau and Research from NZIST on 4 November 2025 for a two-hour workshop followed by written feedback from four of the business units present.
    - Wānanga: Officials met individually for two-hour workshops with Te Wānanga o Aotearoa and Te Whare Wānanga o Awanuiārangi on 3 and 6 November 2025). A further follow-up meeting with Te Wānanga o Aotearoa on 20 November 2025 based on their written feedback.
    - Private training establishments: Officials met with representatives from seven PTEs on 10 November 2025 and an individual meeting (due to scheduling) with Carey Baptist College on 12 November 2025. Five organisations provided further written feedback on the options discussed.

Initial advice of the Technical Advisory Group

2. The Technical Advisory Group (TAG) was convened on 15 September 2025 to provide the Ministry with technical advice on the design of the metrics component as it relates to universities. Chaired by an independent Chair Professor Kath McPherson, the nine members include Australian Research Council officials as well as New Zealand academics and research management professionals. Members were appointed based on their expertise and not as representatives of their organisations.
3. The group's terms of reference focused on providing technical advice on the implications of different design choices, not policy advice on the overall fund design. The group was established primarily to help us understand the design implications of a field-weighted citation index metric (applicable only to universities), so membership was limited to the university sector rather than other TEOs.
4. While the group is still in the process of formulating its final advice, which will inform our advice to you in the new year on concrete options for the metrics component design, its feedback to date has included:

    - Ensuring non-citation research is adequately recognised is a critical criterion for the metric component. The group is particularly concerned that mātauranga Māori, Pacific research, and creative arts research risk being disadvantaged by a metrics-based approach. Although the group recognises that simplicity is important for ensuring a lower-cost exercise, it notes that research which is not visible in the metrics could be at risk of losing institutional support.
    - A bespoke approach to the citation metric will be too compliance-heavy. The group has more mixed views about the relative merits of a commercial provider of publication and citation data versus OpenAlex. While acknowledging a commercial approach will likely be the simplest, the group has concerns about commercial capture, cost, a lack of

Proactively Released

transparency and flexibility, and that this would run counter to global best practice and the accelerating shift toward open research.

- The potential complementary metrics summarised in Appendix 2 and considers that all have limitations as proxies for research quality and performance. The TAG endorses a lower weighting for the metric component in recognition of those limitations. However, the group endorses the use of some grey literature metrics as a proxy for research impact. Some members of the group also support Open Access and early career researcher measures.

University feedback

5. Universities expressed broad support for a metric component design that reduced compliance burdens but raised concerns about the accuracy and fairness of a metrics-based approach, particularly for the arts and humanities. They emphasised the importance of PBRF/TREF funding for sustaining research underpinning teaching, especially given shifts toward mission-led funding in the Science, Innovation and Technology system.
6. We heard some support for a significantly reduced overall metric component weighting to address the limitations of a metrics-based approach. We also heard that introducing subject area weightings through the RDC component would be the most equitable way of addressing the specific limitations for arts and humanities disciplines.
7. The universities are supportive of incentivising research impact but are keen that it is understood as having broad social and economic dimensions and is not narrowly focussed on commercialisation. We heard some concerns about the limitations of existing commercialisation data. Universities were more supportive of metrics that recognised policy impact, as well as Open Access publication. Appendix [x] provides a summary analysis of potential complementary metrics for the universities against a range of criteria, including the sector's views.

Te Wānanga o Aotearoa

8. Overall, Te Wānanga o Aotearoa (TWOA) is supportive of the changes proposed and welcome the change of focus from the individual staff member to the institution. A link between the strategic policy and to the investment plan of the organisation makes sense.
9. Research excellence is a key aspirational endeavour for TWOA. TWOA explained they position themselves within the indigenous research paradigm and consider they have previously been disadvantaged by not adhering to Western education and academic standards.
10. They consider the current PBRF model is very time-intensive for individual academics and defined success on an individual basis. This does not fit the tikanga of TWOA who consider research to be a collective endeavour. The new TREF approach will go some way to prevent this from happening and instead bring a focus to the quality of the institution.
11. TWOA's world view is different to other institutions and their research often contains alternative sources of knowledge such as pūrākau1, which can be difficult to cite and acknowledge the knowledge source. They suggested a metric should be a measure of the community and cultural impact of the research being undertaken but also acknowledged that is difficult to measure.
12. TWOA advised that applied research with a significant practical component and tangible impact on the community needs to be an important metric. An example of this is how the practical component of a degree can be used to start a business or responding to the needs of a community.

Proactively Released

13. They advised that there needs to be a clear definition of "research-led" versus "research-informed" as these are two different approaches depending on the research content. They suggested combining them to create a metric of "research-led and informed" metric.

14. TWOA consider that wānanga should not be included with polytechnics or other PTEs, and that they should be able to have wānanga-based metrics and associated separate funding stream. They view themselves as not fitting within either universities or with polytechnics and PTEs, and this needs to be reflected in the TREF design structure that they occupy a unique position both in terms of research focus and statutory role.

15. When officials put to them that this would create a zero-sum game between wānanga for funding they agreed this would be a concern but thought this could be alleviated with a dynamic funding pool as discussed in the paper.

New Zealand Institute of Skills and Technology

16. The New Zealand Institute of Skills and Technology (NZIST) raised concerns the proposed metrics could disadvantage polytechnics if they are not able to be reviewed each round. Queried if there is potential to put in a metric that acknowledges the unique position of polytechnics and acknowledging the need to review this in future rounds of TREF.

17. NZIST suggested that ring fencing of funding was an attractive option for some, providing this was able to be reviewed against the metrics and did not disadvantage non-universities and have a negative impact on any parity of esteem between the non-universities and universities.

18. They also suggested that an initial allocation for non-universities could be 3.5 - 4% of the total TREF fund, with the ability to review, as there is still a need to retain incentives and still be a good return on the investment.

19. NZIST are of the view that if different metrics for funding allocation are being used for non-universities versus universities, then two different pools of funding made sense as to not disadvantage non-universities. Otherwise, if there is only one pool of funding, everyone is competing for the same amount and the universities will have a larger portion almost by default.

9(2)(g)(i)

20. They suggested a focus on how the new metrics are an evolution, not a complete re-development. Non-universities therefore consider it an opportunity to give a higher weighting to research that involves community development, gives a greater consideration to Te Ao Māori, and has a higher focus on capability and growth.

21. They also suggested metrics that are focussed on incentivising community engagement and impact, as well as working with industry-leaders. This includes strongly defining what an "early career researcher" is, as this means different things depending on if a learner is in a polytechnic or a university environment.

22. NZSIT explained that non-universities are significant employers of early career researchers, and one of their key goals is to develop these researchers, but this does mean the output activity from them is lower. As a result, any volume or quality-based metrics are not appropriate. NZIST suggested that for any metric to be appropriate to their goals, these factors (including community impact) need to be accounted for.

23. NZIST advised that having "Quality" as a metric is a very difficult thing to measure. This would result in ranking and comparing between institutions and not work. Connectivity with industry, productivity and regional impact are more appropriate measures.

24. They advised that generally, polytechnics and non-universities do not have the resource or capacity to provide extensive commercialisation opportunities for their research, unlike universities. If commercialisation is a required metric, there needs to be a focus on any partnership programmes or entrepreneurship opportunities.

Proactively Released

25. Along with other non-universities we engaged with, NZIST consider mātauranga Māori, Kaupapa Māori, and Pacific research needs to be reflected in the metrics to ensure their work in these areas has the support it requires.

26. Finally, they advised that a 2-year frequency cycle for TREF is appropriate. They feel that an annual approach places too much focus on output, whereas a 3-year cycle is too long for close monitoring. They welcomed the opportunity to incentivise collaboration between institutions, lowering their overhead and financial costs, and creating a cross-country network of providers and research being undertaken.

Te Whare Wānanga o Awanuiārangi

27. Te Whare Wānanga o Awanuiārangi (TWWA) are not supportive of a ring-fencing approach. They are of the view that if funding is ring-fenced, the metrics need to ensure that they are not disadvantaging providers who do not meet the standard criteria.

28. They highlighted that the research challenges they face are because what they consider important and impactful, is not necessarily reflected in funding allocations or considered to have the same value as research conducted by other providers. TWWA are cautious of a system that privileges one way of doing things over another.

29. TWWA advised they do not mind being compared to university sector, however there is an issue about what constitutes research. If their research outputs are to be compared to larger institutions, they need to be seen as equal to universities and other providers with degree recognition. Parity of esteem is very important to them, and they do not want their research degree programmes to be seen as "lesser than" other providers.

30. They reported that they compare well when it comes to research outputs with other institutions of a similar size, however any metrics that consider citation levels and outputs need to consider their unique position as a wānanga, as does the definition of what constitutes research. Much of the research output from them is community-impact focussed, and therefore not quantifiable in the same way as other research is.

31. They discussed how any metric weighting needs to also consider the unique position of TWWA. While their RDC level are high for an institution their size, it is low compared to other institutions and therefore could disadvantage them should funding be reduced because of this.

32. TWWA consider that volume is not always the fairer measure for research funding. The number of citations and peer review opportunities are lower for them as much of their research is published in Te Reo Māori, and therefore publication opportunities are limited.

33. They were very clear that they would also like to see the provision of mātauranga Māori and community impacts as a metric. They advised that while their priorities of community impact and focus on social sciences do not reflect the current government's priority and the Going for Growth agenda, they should not be penalised for this focus.

Private Training Establishments (10 and 12 November)

34. The feedback from the 11 Private Training Establishments (PTEs) that currently participate in the PBRF was mixed. Some were supportive of a review and see the shift of focus from individual academics to institutions as positive. Others saw the shift from the individual academics to an institution focus to be a disadvantage for them.

35. The New Zealand School of Chiropractic advised that due to the funding granted to them by the PBRF, they have been able to expand their research capability to 30 full-time researchers. They are therefore concerned that any change to the metrics will reduce their funding and therefore their research capability.

Proactively Released

36. Other PTEs raised that any metric that relied of citations and publications could disadvantage PTEs. Equity with universities and other provider is a significant issue for them, as they do not have the capacity or financial ability to pay for resources like Open Access journals, and therefore their citations are not counted to the same extent as larger institutions with outputs.
37. Some PTEs raised the possibility of post-course evaluation and feedback surveys being counted in the metrics, with this being a possible way to measure community impact. They advised that often community impacts are not seen until 6-12 months after the research has been completed, and therefore surveys and evaluations would ensure this could be counted in each TREF round.
38. Faith-based PTEs raised that theological research does not align with the current government priorities for STEM subject and the Going for Growth agenda, as commercialisation of this type of research is not the goal when undertaking it. Therefore, tailored and flexible metrics that reflect the unique position of PTEs (not just faith-based) should be considered, as they should not be compared to universities.
39. Overall, the PTEs considered that community impact and the inclusion of mātauranga Māori and Pasifika research needs to be considered in any metric. Often, the community impact is the most tangible, however can be difficult to measure.

Proactively Released

Annex 2: Potential complementary metrics for Universities

	OBJECTIVES			ASSESSMENT CRITERIA			
Metric	**Proxy for non-citation research**	**Proxy for research impact**	**Supports a sustainable research workforce**	**Robust and future-proofed metric**	**Does not increase compliance costs**	**Sector credibility**	**Does not create perverse incentives**
Research outputs with non-academic co-author	Not intended to proxy but may be negative proxy – will further benefit disciplines where non-academic authorship is common, and these will already benefit from the FWCI	Will proxy – indicates external focus and collaboration, though not necessarily benefits to stakeholders/realised impact	Not intended to proxy but may be negative proxy – may favour established academics with strong external networks	Robust – established body of data	Low compliance – data can likely be collected through the same source as the citation metric.	Sector does not support – will favour areas that are already well-represented by bibliometrics (esp. biomedical sciences) and potential for gaming	May incentivise non-genuine non-academic co-authorship.
Research outputs with international co-authors	Not intended to proxy but may be negative proxy – will further benefit disciplines where co authorship is common, and these will already benefit from the FWCI	Will proxy academic collaboration but not research impact	N/A – not intended to proxy	Robust – established body of data	Low compliance – data can be collected through the same source as the citation metric.	Sector does not support – will favour areas that are well-represented by bibliometrics	May incentivise non-genuine international co-authorship.
Combined commercialisation data from SCOPR – IP, spin-outs, start-ups, TTO resource	Will proxy - however limited to a small number of mainly scientific and engineering fields.	Will proxy commercial impact – evidence of commercialisation, and may indicate ongoing development if maintenance measured.	N/A – not intended to proxy	Non-robust – currently very low numbers. Would disproportionately favour Auckland	Low compliance – can be collected through the SCOPR annual survey which all 8 universities participate in.	Sector apart from Auckland does not support due to small numbers and UniServices' dominance	None identified so long as metric includes maintenance to avoid incentive to register IP which is not used or to start-up non-viable companies.
Grey literature publications (e.g. reports, guidelines, policy documents, legal opinions, Waitangi Tribunal reports)	Will proxy – will capture research across most disciplines, but less so for creative arts and some humanities, so will not necessarily counterbalance all citation metric effects.	Will proxy non-commercial impact – input into govt, organisational, industry or sector policy etc is a strong indicator of social and economic impact	N/A – not intended to proxy	Further investigation needed. Discussions with sector indicate that data cleansing would be required and clear definitions would also be essential but sector supports doing this work.	Existing grey literature databases include Overton (commercial) and GreyMatters (free). Institutional repositories could also be used. Sector supports this option and is willing to invest in additional compliance in this area	Sector is broadly very supportive of this metric.	None identified
Citations in policy documents, guidelines, legislation etc	Will proxy – will capture research across most disciplines, but less so for creative arts and some humanities, so will not necessarily counterbalance all citation metric effects.	Will proxy non-commercial impact – is evidence that external stakeholders have drawn on research to develop policies.	N/A – not intended to proxy	Further investigation needed but citation practices across organisations are uneven so obtaining accurate data may be challenging.	Further investigation required - several commercial products available including PlumX (Elsevier), Policy Citation Index (Clarivate) and Altmetric (Digital Science), but unclear how comprehensive	Sector is supportive in principle but concerned about additional compliance to ensure cites are being captured.	None identified.
Media and public engagement metrics (e.g. media mentions, public talks,	Will proxy some areas, particularly arts and humanities research, but is more of an indicator of	Will proxy non-commercial impact – is evidence of an external focus and engagement, but is	N/A – not intended to proxy	Likely non-robust – validation of public engagement activity is difficult	High compliance – some solutions such as Altmetric track media and social media engagement but won't capture all types so would require additional data from	Sector does not support due to compliance and robustness issues	Open to inflation of numbers, particularly self-reported data such as number of public talks etc

Proactively Released

	OBJECTIVES			ASSESSMENT CRITERIA			
Metric	**Proxy for non-citation research**	**Proxy for research impact**	**Supports a sustainable research workforce**	**Robust and future-proofed metric**	**Does not increase compliance costs**	**Sector credibility**	**Does not create perverse incentives**
exhibitions, media engagements)	dissemination activity rather than research output.	dissemination rather than impact			TEOs. Sector indicates they don't routinely capture this data.		
Open access metric (e.g. percentage of outputs published under OA, or OA Green)	Will not directly proxy but if based on institutional repositories will include research outputs that aren't included in databases	May proxy – open access publishing is linked to research impact through making research freely available to stakeholders and end users	N/A – not intended to proxy	Robust	Low additional compliance – universities track this information already, and external sources such as COKI and OpenAlex also track OA rates	Strong sector support	None identified
Early career researcher metric (e.g. proportion of academic staff)	N/A – not intended to proxy	N/A – not intended to proxy	Will directly measure – would incentivise investment in ECR workforce	Can be robust but will require definition of an early career researcher and clear audit processes	Medium compliance – would require TEOs to provide data and TEC to verify or audit.	Mixed views – some universities are strongly in favour, some strongly against on basis of potential compliance.	Staff-based measures historically have driven some undesirable behaviours.
ADDITIONAL MEASURES/APPROACHES							
Subject area weighting (note this is not a metric. It would be applied in the funding formula rather than being a complementary metric)	Will directly measure – would ensure a direct recognition of subject areas with significant amounts of research not captured by bibliometrics. Weighting could be calibrated to size of 'gap' between FWCi database and institutional repositories at national level.	N/A – not intended to proxy	N/A – not intended to proxy	If data source is the FWCI research database, weightings will be robust at national level but could be less robust for individual universities e.g. those with much more or much fewer outputs than the national average	Medium compliance – TEOs would need to validate subject codes that had been applied to outputs in the research database (but note this would need to happen anyway to calculate the FWCI) and to apply subject codes to outputs in institutional repositories	Sector does not support this due to potential for gaming	Could incentivise over-publication in weighted subjects if not calibrated correctly, or attempts to opportunistically reclassify outputs, thus requiring stricter auditing
RDC subject area weighting	Will directly measure – would ensure a direct recognition of subject areas with significant amounts of research not captured by bibliometrics. Weighting could be calibrated to size of 'gap' between FWCi database and institutional repositories at national level.	N/A – not intended to proxy	N/A – not intended to proxy	If data source is RDCs, will be extremely accurate	Low compliance for universities – already declare research degree completions and subject areas to TEC through SDR. TEC would need to carry out some additional work to calculate funding.	Strong sector support	Could incentivise over-recruitment into research degrees with higher subject weightings if not calibrated correctly.
Output type weighting (also not a metric – would apply in funding formula)	Will directly measure – would ensure direct recognition of research output types not captured by bibliometrics e.g. monographs, creative works, performances, software, reports etc	N/A – not intended to proxy	N/A – not intended to proxy	Robust but will require definition of output types and data source	Possibly high compliance– TEOs would need to ensure correct output type. May require bespoke research output types in database if existing solution needed. TEC would need to verify the data.	Sector does not support this due to compliance costs	Could incentivise over-publication in weighted output types if not calibrated correctly, or attempts to opportunistically reclassify outputs, thus requiring stricter auditing

Proactively Released

Annex 2 continued: Potential metrics for non-Universities

	OBJECTIVES				ASSESSMENT CRITERIA			
Metric	**Economic impact**	**Growing research capability /capacity**	**Supports unique characteristics of TEO type**	**Contribution to regional and local economy**	**Robust**	**Does not increase compliance costs**	**Sector credibility**	**Does not create perverse incentives**
Research outputs count or per FTE of research active staff	Indirect proxy for economic impact depending on the output type	Incentivising outputs supports growing research capacity	TEOs will be free to determine outputs to be delivered in accordance with their priorities	May be an input proxy depending on the output type	At more robust end of scale. Requires clear definition of what outputs qualify.	Regularly collected data. Will have compliance costs but at the lower end of the spectrum.	Credible measure of research production but likely some reservations about quality not being assessed.	May encourage higher quantity of lower quality research outputs or activity.
Total research active staff (FTE)	Indirect proxy for economic impact depending on the research focus of the TEO	Direct proxy for research capacity	TEOs will be free to determine research focus of staff	Indirect proxy. Enables more regional and local impact but not directly measured.	Robust with verifiable data.	Regularly collected data. Will have compliance costs but at the lower end of the spectrum.	Credible measure of research capacity, but possible reservations about not measuring quality or output.	None identified.
Research time as proportion of academic staff FTE	Indirect proxy for economic impact depending on the research focus of the TEO	Direct proxy for research capacity. More accurate than as it captures how much time is spent on research.	TEOs will be free to determine research focus of staff	Indirect proxy. Enables more regional and local impact but not directly measured.	Robust but will need clear definitions about how research time is counted.	Regularly collected data. Will have compliance costs but at the lower end of the spectrum.	Credible measure of research capacity, but possible reservations about not measuring quality or output.	None identified.
Number of degree completions	Poor proxy for economic impact.	Supports some research capability, but teaching time can detract from research capacity.	Poor proxy for the research role of non-uni TEOs/	Not intended to proxy.	Initial modelling shows poor alignment with research outputs as measured by the QE.	Numbers already reported as part of SDR. Funding received as part of other TEC funding (E.g. DQ)	Little sector credibility for measuring research output or quality	High. Potential for TEOs to pursue volume over quality of provision.
Number of research collaborations	Indirect proxy for economic impact depending on the type of collaboration and the research focus of the TEO.	Indirect proxy for research capability growth.	Supports the industry and community focus of TEOs and applied research.	Indirect proxy depending on the activity type	Less robust. Will rely on very clear definitions and data collection – and the tighter the definition the greater risk of desirable activity not qualifying.	Data is not regularly collected or verified.	Support likely from sector but viability of data collection untested and this may reduce the credibility. Tight definitions may be considered unfair by some.	High. May encourage engagement but does not measure output or impact.
Grey literature publications (e.g. reports, guidelines, policy documents, legal opinions, Waitangi Tribunal reports)	Less direct alignment. Some measured activity will align but will also capture a variety of other activity.	Supports research capacity but captures a smaller range of outputs.	Less appropriate as a proxy for the role / research focus of TEOs.	Less direct alignment. Some measured activity will align but will also capture a variety of other activity.	Further investigation required and obtaining accurate data may be challenging.	Possibly high cost as it may require a commercial database and subscriptions for all TEOs.	Low sector support.	None identified.

Proactively Released

	OBJECTIVES				ASSESSMENT CRITERIA			
Metric	**Economic impact**	**Growing research capability /capacity**	**Supports unique characteristics of TEO type**	**Contribution to regional and local economy**	**Robust**	**Does not increase compliance costs**	**Sector credibility**	**Does not create perverse incentives**
Research engagement satisfaction	Indirect proxy for economic impact – some captured activity will have an economic focus but not all.	Supports high research capability as this is a contributor to positive engagement scores.	Aligns well with the industry and community focus of TEOs	Indirect proxy for economic impact – some captured activity will have an economic focus but not all.	Relies on clear definitions of qualifying activity and a system to consistently seek engagement feedback.	New measure relying on a new system to capture and verify data.	Unclear. Some support but credibility would depend on how a measure could account for the scale of activity.	Unclear, more investigation would be required.
Media and public engagement metrics (e.g. media mentions, public talks, exhibitions, media engagements)	Not intended to proxy.	Difficult to establish direct link to research activity. Small numbers mean unsuitable as proxy.	Not intended to proxy	Difficult to establish direct link to research activity. Could indicate evidence of local/regional engagement	Likely to be a small number of outputs from many non-unis which would lead to an inconsistent measure. High risk of inaccuracy.	Would rely on a new database – possibly commercial – to track and associated costs.	Some support but viability of data collection untested and inaccuracy would undermine credibility.	High. Potential for TEOs to create quantity over quality.
IP metrics including patents granted/maintained, spin outs, startups	Direct proxy.	Poor proxy for capability as this is not a significant activity for non-uni TEOs.	Poor alignment with non-uni research focus.	Indirect proxy but unlikely to capture much activity.	Likely to be a small number of outputs from many non-unis which would lead to an inconsistent measure.	Non-university TEOs do not have Tech Transfer Offices so likely to create additional costs	Poor fit with research focus of non-uni TEOs.	May incentivise activity that is a poor fit with the role of some providers.
Early career researcher metric (e.g. proportion of academic staff)	Not intended to proxy.	Indirect proxy.	Not intended to proxy.	Not intended to proxy.	May be useful metric for some non-university TEOs growth but likely to be widely inconsistent across the sector.	New measure and not currently tracked. Would require a new database to measure.	Some support from TEOs	N/A

Proactively released

Annex 3: Summary of design approaches

Context

The three core approaches for the overall design of the TREF can be seen as points on a spectrum rather than discrete packages that exclude all other options. Thus, different approaches for universities and non-universities can be implemented if desired.

Current weightings for the QE, ERI and RDC components are 55%, 20% and 25%, respectively.

Comparison of component weightings

Component	Current weighting	Option 1	Option 2	Option 3
RDC	25%	30 –45%	20-30%	30-40%
ERI	20%	30-45%	30-45%	Approx 20%
Metric component	55% (Quality Evaluation)	10-30%	30-45%	40 –50%

1: Efficient allocation of funding

Universities metric component
- Unmodified FWCM
- A simple metric based on existing data and worth less than 25% of the metric component

Wānanga, ITP and PTEs component
- Single metric of research capacity

Favours providers that performed less well in QE and larger providers with greater research capacity

ERI	RDC
Based on phase 1 settings	Possible consideration of changes to subject weightings

Overall implications
Lowest cost; least accurate; least incentive for impact

Metric component weighted lower than QE component to allow for the lower accuracy and breadth in the measure

Pros
Reduced cost
Reduced complexity
Basic research in sciences, medicine, social sciences well-represented

Cons
Less comprehensive basis for distributing funding
Would provide less information about the research system
Arts and humanities, matauranga Māori, applied research not well-represented and may be disincentivised.

2: Research impact

Universities metric component
- Unmodified FWCM with lower internal weighting
- 2 research-impact and KE metrics

Wānanga, ITPs and PTEs – we do not recommend this approach for this subsector

ERI	RDC
Based on phase 1 settings	Possible consideration of changes to subject weightings

Overall implications
Aligned with *Going for Growth*. More complex and costly to measure; most incentive for impact

Metric component weighted lower than the QE but higher than Opt. 1 to recognise greater complexity of the measure

Pros
Applied research and sciences, medicine, health well-represented
Will provide additional information about the research system

Cons
Arts and humanities will not be well-represented
Could duplicate SIT system incentives
Narrower focus may not support breadth of research and research-led teaching

3: Comprehensive research coverage

Universities metric component
- Modified FWCI including supplementary data with internal weighting of ~60-70%
- 3-4 metrics that address gaps in citation metric and research workforce

Wānanga, ITP and PTEs component
- 3-5 metrics focused on research capacity and subsector unique characteristics

Distribution of funding would be highly uncertain

Compliance costs would be significant relative to funding value, and some TEO activity will not have sufficient volume

ERI	RDC
Based on phase 1 settings	Possible consideration of changes to subject weightings

Overall implications
Highest cost and compliance burden; most accurate

Metric component closer to QE component to reflect higher compliance costs

Pros
Comprehensive
Complements priorities for the SIT system
Will provide rich data about the research system

Cons
Highest cost and complexity

ERI External Research Income; **FWCI** Field-Weighted Citation Index; KE knowledge exchange; **QE** Quality Evaluation; **RDC** Research Degree Completions; **SIT** Science, Innovation and Technology; **TEO** tertiary Education Organisation; **TREF** Tertiary Research Excellence Fund; **TTO** Technology Transfer Office

Proactively Released