



Australian Government

Australian Research Council

**Research Quality
Framework**

**RESPONSE TO THE
ISSUES PAPER**

May 2005

RESEARCH QUALITY FRAMEWORK (RQF)

Responses to the Issues Paper

Overall comments

The following comments form part of the Australian Research Council (ARC) response to the Issues Paper, and supplement those included under the thirteen Issues on which comment was requested by the Department.

In this response, the term 'research' is taken to include research undertaken within the universities and publicly funded research agencies. For the purpose of this response, research and development activities undertaken largely by business enterprises are excluded.

The Australian Research Council's (ARC) mission is to "advance Australia's research excellence to be globally competitive and deliver benefits to the community".

The ARC believes that the benefits that stem from ARC-funded research are realised through six distinct benefit channels, namely:

- benefits from building the basic knowledge stock;
- benefits from improving the skills base of Australians;
- benefits from diffusion of skills, knowledge and other intellectual property that lead to enhanced commercial outcomes;
- benefits from improving access to and ability to capitalise on international research outcomes and international networks;
- benefits from better informed policy making; and
- enhanced health, environmental and cultural benefits.

The ARC is charged with the support of excellence in research and is committed to facilitating the uptake of the outcomes of such research, both within the research sector and more widely to secure broader societal impact. Through its National Competitive Grants Program, the ARC identifies excellence in proposals for research funding and research training through an expert-based competitive assessment process, which takes into account the research records of the investigators, the approach and method proposed, the significance of the project and its potential national benefit. The resultant funding is awarded to eligible applicants for research proposals judged to be excellent and of potential benefit wherever and in whatever disciplines (other than clinical medicine and dentistry) they arise.

The ARC therefore supports in principle the establishment of a Research Quality Framework to evaluate the quality and wider impact of Australia's publicly funded research effort more generally. This support is, of course, conditional on the requirement that the scope and conduct of the associated evaluation process should be broadly consistent with and complementary to the annual assessment procedures conducted by organisations such as the ARC and NH&MRC through their major competitive, expert-reviewed schemes. Such a framework has the potential to:

- provide an information base on Australia's publicly funded research mix and focus, which either does not exist at present or exists only in partial forms;
- focus on the outputs of, rather than the input to, Australia's publicly funded research activity;
- provide information on the benefits achieved from Australia's investment in research;

- identify the institutions, research groupings and disciplines in which Australia's research effort excels (including pockets of excellence within otherwise less research-intensive institutions or disciplines);
- provide benchmarks for international comparisons; and
- inform the allocation of a significant part of that government research funding which is not currently allocated via competitive mechanisms.

The ARC believes that the scope and operation of any such framework should be closely aligned to its intended uses. In particular,

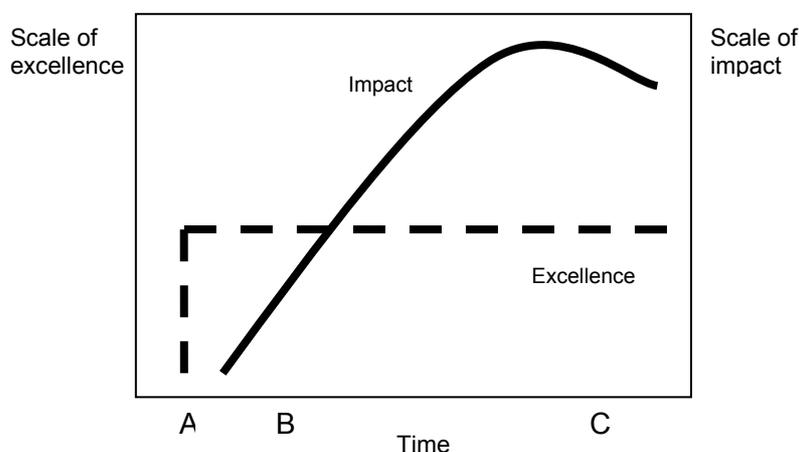
- if the framework is intended to guide funding decisions, the nature and extent of the funding at stake should be clear, as this will influence the design and operation of the framework, including appropriate levels of aggregation of information gathering and reporting;
- if it is believed that funding should follow excellence, then the implications of reducing or withdrawing funding for some disciplines in some institutions and locations – particularly when research is linked to undergraduate teaching or to areas of skills shortage – must be understood and addressed (see the ARC's response under Issue 12);
- if it is to provide information on the nature and national benefits of Australia's publicly funded research activity (including the geographic distribution of that activity and the nature and location of 'pockets of excellence'), it should be comprehensive and include all relevant disciplines, institutions and research groupings, and recognise the vital role of research training, rather than focus on a subset of those categories;
- the framework should not introduce, overtly or otherwise, incentives for institutions and other groups to redirect research resources and activity in ways that are likely to be counterproductive in the national context or to reduce the overall benefits of the activity. Similarly, it should not discourage activity generally viewed as beneficial.

Further, the definition of quality needs to be carefully considered. The ARC agrees that, in this context, the notion of quality encompasses both the intrinsic excellence of research and the impact it exerts on the research community, industry and the wider community. However, it firmly believes that these characteristics should be assessed *separately but in parallel* in any framework. This is because they arise in different timescales, may occur in different research environments, require different skills and are influenced by different incentives and constraints. For example,

- The excellence of a piece of research or of a research facility can generally be assessed on the basis of its content and output at or soon after its release. Expert review processes currently in use in Australia at present are generally regarded as generating robust rankings and comparisons of excellence, which are likely to be relatively insensitive to the particular time at which they were made (see diagram below).
- On the other hand, the impact of a piece or body of research is likely to become apparent over a prolonged period of time. If measured shortly after its release, the impact of much research is likely to be limited to the discipline involved (and captured in indicators such as citations and other bibliometrics). This is particularly the case with 'discovery' research, where the involvement of a broader range of partners may initially be very limited. However, as the results of the research are disseminated within and beyond the discipline, as they stimulate further or applied research and begin to influence ideas, processes, products, services and technologies, they may have a more substantial, sustained and diffuse effect, either on particular groups within the country or more broadly. This is illustrated in the diagram below, where, for example, the impact of the discovery of penicillin would have appeared very much larger in 1944-45 when it saved many Allied lives (Year C) and later millions of lives, than it would have if assessed shortly after its reported discovery in 1928 (Year A – discovery - or B – around 5 years after discovery). (Fittingly, Alexander Fleming, Howard Florey and Ernst Chain - the co-worker of Florey who is credited with rediscovering Alexander Fleming's

work in the literature - shared the Nobel prize in 1945, neatly illustrating that both the primary discovery drivers and the impact facilitators are important in the quest to deliver beneficial outcomes.)

DIFFERING TIMESCALES OF RESEARCH EXCELLENCE AND RESEARCH IMPACT



It follows that, for a given piece of ongoing discovery-based research at a given early point in time (Year A or B), any attempt to amalgamate indicators of excellence and impact into a single upper-level quality rating is unlikely to yield the right information to assist in understanding the mix and focus of Australia's publicly funded research. A better solution is likely to involve identifying means of rating research quality against a range of dimensions, including excellence and impact.

The ARC also believes that excellence and impact should be recognised as separate characteristics of research and supported in complementary ways. For example:

- funding arrangements such as the university block grants should reflect quality assessments based largely on the excellence of the research outcomes of the institutions and research groupings involved (in a similar way to the current competitive research grants) ,
- a 'third stream' of funding could be available to university-based research intended to diffuse and apply new knowledge and so deliver benefits to the nation as a whole. Such funding could take into account a wider range of impacts, also possibly assessable under a research quality framework, including the likely nature, extent, beneficiaries and duration of the impact of the research. It could also recognise that the nature and extent of the impact of excellent research will be influenced by factors – including the media and industry coverage the research attracts, the priorities of industry or community groups at particular points in time, the macroeconomic environment, and development-specific facilitators such as venture capital support, among others – that are sometimes beyond the control of the initiating researchers. It could also recognise that activity intended to add value to the original findings and make it accessible to others in the community may be undertaken in different institutions, locations and even disciplines from those in which the key findings originated. It will almost certainly require different skills and processes from those applied in the original research.

It follows that a Research Quality Framework should desirably enable these different aspects of research quality to be acknowledged and valued.

Finally, the ARC believes that the framework should not duplicate existing research assessment processes (including those of the ARC) for identifying research excellence in funding proposals *and indeed must be capable of delivering outcomes which are consistent with and add value to*

those processes if it is to be justified to the research community. Its expected benefits must exceed its implementation and administration costs.

The ARC does not underestimate the practical difficulties of devising and implementing such a system in ways that are transparent and supported by both the research community and those who use and fund that research. Indeed, the experiences of those countries, including the UK and NZ, which have implemented such schemes indicate the nature of the difficulties and their consequences. However, they also provide lessons from which Australia can benefit.

Research assessment is a major part of the core business of the ARC and the NH&MRC. The mechanisms in place in those organisations, which include expert review informed by commonly accepted metrics, are likely to be appropriate to this broader exercise, at least where it applies to university-based research. The ARC submits that it would be desirable for such existing organisations and mechanisms to be involved in the development of any Research Quality Framework, and possibly even to accept responsibility for one or more of the assessment components of the process, given appropriate additional resources.

Part 2: Creating an Australian RQF

Please indicate your response by placing an X in the relevant box.

2.1 Structuring an RQF

Issue 1: How should an RQF be applied to universities and publicly funded research agencies?

(a) An RQF should be applied in the same way to both universities and publicly funded research agencies.

<input type="checkbox"/>	Strongly agree	<input type="checkbox"/>	Strongly disagree	<input type="checkbox"/>	No comment
<input checked="" type="checkbox"/>	Somewhat agree	<input type="checkbox"/>	Somewhat disagree		

(b) Within the university sector, an RQF should be applied differentially to specific types of institutions.

<input type="checkbox"/>	Strongly agree	<input type="checkbox"/>	Strongly disagree	<input type="checkbox"/>	No comment
<input type="checkbox"/>	Somewhat agree	<input checked="" type="checkbox"/>	Somewhat disagree		

Additional comments

(a) The ARC believes that, ideally, all aspects of publicly funded research and experimental development need to be included in the RQF. The ARC is also of the opinion that, everything else being equal, the maximum long-term impact arises from appropriately connected high-quality basic research. To the extent that common elements of quality can be defined independently of the type of institution in which research is conducted, there are therefore advantages in adopting similar measures to track and compare them. However, where there are obvious differences in the nature or objectives of the research (e.g. in mission, or potential for commercialisation), different measures and/or procedures may be required to capture different interpretations of quality, or different weightings may need to be applied to those measures. (This is more likely to be the case in comparing research impacts than in identifying excellence as such.)

The ARC notes that heterogeneity is likely to exist *within* as well as between university-based research and that of the publicly funded research agencies (PFRAs).

Ultimately, of course, the choice of organisations to be included will reflect the types of research funding at stake as a result of the assessment process.

(b) There are advantages in terms of comparability and fairness in applying a broadly similar assessment process to all institutions. Again, however, where different institutions within the university and PFRA sectors have different research missions or focus, it may be desirable to apply different weightings to the quality measures used in each case. Again, too, the funding at stake as a result of the process will be a relevant consideration.

2.2 Defining and measuring research quality and impact

Issue 2: Research quality and impact should be assessed by appropriately constituted panels.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

Additional comments

The experience of the ARC and the NH&MRC suggests that 'appropriately constituted panels' are the best ways of assessing research quality and impact.

The question of how many panels are required to fairly assess research quality across all areas of research activity is hotly debated. In the ARC's opinion and experience, the discussion of the number of panels cannot usefully be separated from a discussion of the size of individual panels and the sub-disciplines represented on such panels. As a large proportion of current research can be classified as interdisciplinary, there is an argument for having a relatively modest number of panels (e.g. 12) with fairly broad representation to capture a good proportion of the interdisciplinary research without the need for extensive panel-mixing. The ARC currently operates a system of six discipline areas (Mathematics, Information and Communication Sciences; Physics, Chemistry and Geoscience; Engineering and Environmental Sciences; Biological Science and Biotechnology; Social, Behavioural and Economic Sciences; and the Humanities and Creative Arts). The ARC's College of Experts comprises 75 individuals, where 12 or 13 individuals are assigned to each of the six discipline areas. Nevertheless, for specific tasks such as the selection of Centres of Excellence, members of each subdiscipline panel and international experts are assembled into a team capable of forming informed judgments about interdisciplinary applications benchmarked appropriately to international standards.

Inevitably, the mode of assessment chosen will involve a trade-off between the availability and workload of panel members and the cost of operating the panels.

The *ex post* assessment of research quality has, by definition, access to actual research outputs and (early) impacts, some of which can be quantified and described. If the RQF is to function as an *ex post* assessment mechanism, there is somewhat more scope than in the *ex ante* competitive grant allocation process to supplement panel deliberations with other means of assessment, including quantitative and qualitative indicators.

2.3 Measuring research quality and impact

Issue 3: Assessment panel members should include the following (the categories are not necessarily mutually exclusive):

(a) Experts reviewers able to assess impact in a discipline area/academic field.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

(b) Expert reviewers able to assess impact more widely.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

(c) International expert reviewers.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

Additional comments

The inclusion of panel members able to assess research impact will be important. Peer reviewers are typically better equipped than other experts to assess intrinsic research excellence. However, experts drawn from industry and other sectors may be better able to judge the potential impact of the research (except in the case of impact on the discipline involved).

The ARC favours the use of a wide definition of research impact, including social and economic benefits. This reflects the mission of the ARC and the expectations of those who invest in Australia's publicly funded research effort.

It follows that it is desirable to include expert reviewers able to assess impact very broadly (including over long time periods and across diverse industry and community sectors) and international reviewers.

The ARC is of the view that competitive approaches to allocation of funding, guided by expert panels, are likely to deliver superior results. In an increasingly globalised world where trans-national access to research results is facilitated, an international perspective is needed to establish not only the relative quality of Australian research but also the likely benefits to Australia of funding a specific piece of research.

Issue 4: Assessment panels should be informed by metrics whose nature and relative influence may vary across different disciplines.

Strongly agree Strongly disagree No comment

Somewhat agree Somewhat disagree

Additional comments

Metrics will inevitably be a useful supplement to panel processes. The ARC does not believe, however, that any broadly based RQF could or should be driven predominantly by quantitative indicators, and supports the primacy of expert panels in the assessment mechanism. This view is entirely consistent with the way the ARC approaches assessment of applications. Indeed, each application is ideally read by several expert referees and metrics could be said to have a strong impact on only a subset of selection criteria, in particular that of investigator track record. Even in this case, a metric approach is likely to be much more difficult in areas such as the humanities and creative arts than in the physical sciences.

Nevertheless, to the greatest extent possible, any metrics drawn on by panels should be independent of the discipline or institution involved. The questions underlying the assessment process relate less to individual disciplines than to the totality of Australia's publicly funded research effort. They should enable comparisons to be made of the international significance of the research (the size of the pond in which the research is conducted) and its impact (the size of the ripples the research makes). As noted earlier in respect of the assessment process itself, it may be desirable to define an overall set of metrics common to all circumstances, but the relative weightings of which may be varied depending on disciplinary or institutional circumstances.

The ARC notes that the metrics canvassed by the Allen Consulting Group (Table 1 of the Issues Paper) provide a useful starting point for any such set. However, they could be further developed in the area of research uptake and dissemination.

It is important to recognise that any metrics incorporated in an RQF or adopted by panels may be expected to trigger behavioural change in research institutions if they differ significantly from those in use (overtly or otherwise) at present. Consequently, it will be important to have a clear view of desired behavioural change in order to judge the appropriateness of any given metric or set of metrics and the benefits of incorporating discipline-specific metrics or variations.

Issue 5 (a): An RQF should recognise research impact through the measurement of different outcomes for different types of research and disciplines.

Strongly agree Strongly disagree No comment

Somewhat agree Somewhat disagree

Issue 5 (b): An RQF should recognise the production and diffusion of technology and knowledge as elements of research impact.

Strongly agree Strongly disagree No comment

Somewhat agree Somewhat disagree

Issue 5 (c): Where appropriate, users and those commissioning research should contribute to the assessment process by providing an external perspective on research under consideration.

Strongly agree Strongly disagree No comment

Somewhat agree Somewhat disagree

Additional comments

It is clear that research and research organisations with different objectives should be expected to produce different outcomes, including different impacts, and that this should be recognised in any assessment process. However, if a sufficiently diverse set of impacts is considered *within a common assessment framework*, these differences will result in different types of research achieving different ratings on different indicators. The ARC believes that, to the greatest extent possible, all research should be assessed against the same standards and categories, but that the results should illuminate the differences among different institutions and disciplines, identify the types of impacts in which particular research types and groups excel and be interpreted in the light of those differences.

As mentioned earlier, the ARC believes that research impact should not be confined to a narrow range of commercialisation criteria but be broadly defined and include indicators of the production and diffusion of technology and knowledge as well as other social and economic impacts. The required indicators should include a range of indicators of research uptake (including a larger number of indicators than are included in Tables 2 and 3 of the Issues Paper.) Users and those commissioning research are likely to be valuable contributors to this process. However, the ARC reiterates the need to recognise the very lengthy timescales over which such impacts may become apparent.

Part 3: Applying an Australian RQF

Please indicate your response by placing an X in the relevant box.

3.1 Level of aggregation for assessment

Issue 6: What is the most appropriate level of aggregation for assessment?

- Subject/discipline area
- Research grouping/research team/s
- Department/schools
- Faculties/Divisions
- Institutional level - university/PFRA

Additional comments

It is important to align the level of aggregation for assessment with the purpose of the exercise. If the framework is to be used to inform decisions about funding at the institutional level, for example, it would be desirable for the assessment aggregation to be done at that level. If the framework is to be used to guide broad policy development (for example, national priorities concerning research needs or skill sets), then aggregation across institutions and/or across discipline/activity groups may also be relevant.

The use of subject/discipline groupings would provide consistency with current competitive research grant assessment arrangements and, in many cases, with research groupings, but may pose difficulties for dealing with multi-disciplinary arrangements and emerging disciplines if the assessment panels themselves are highly disaggregated along disciplinary lines.

The use of subject/discipline areas *within* research groupings and/or department/schools could allow the aggregation of assessment results to whole-of-discipline and whole-of-institution levels. This could produce a picture of overall research excellence and enable fields and locations of particular excellence to be identified. However, in order to achieve aggregate results of this sort, a core of common measures and common standards would need to be applied across all disciplines and institutions.

If the aggregation of scores is at a finer granularity than Institutions, then it becomes important to exercise flexibility in the level of aggregation as definitions of Centres, Departments, Schools, Faculties etc differ from institution to institution.

The ARC notes that the level of aggregation for assessment need not be the same as the level of aggregation for reporting. For example, the input to the assessment process may include the output of individual researchers, but individuals should probably not be identified in reports.

Issue 7: Who should be assessed as part of an RQF?

- Eligible staff nominated by institutions (based on guidelines to be provided)
- All eligible staff

Additional comments

Both in-principle and practical issues are relevant here. If the assessment exercise is intended to be comprehensive, assessing the quality of the totality of funded research activity in the institutions under consideration, then all eligible staff (and definitions will be required) should be included in the assessment. If it is intended to identify 'the best of the best', then institutions and/or disciplines might be invited to submit for assessment only a subset of their research or research activity. The ARC's in-principle preference is for comprehensiveness, at least partly because it is more likely to enable a full picture to be obtained of the geographic and disciplinary distribution of research quality in Australia and its various 'pockets of excellence' as well as opportunity.

A hybrid approach might also be envisaged, in which the process was detailed for a selected subset and less detailed, or with data obtained from other sources, for other staff.

The ARC supports incentives for institutions to attract and retain excellent staff. However, the final choice concerning eligibility should avoid incentives for artificial manipulation by institutions in relation to the selection of subsets of staff or research for assessment. For example, it should not be possible for institutions to 'import' highly reputable researchers for short periods to augment measured research activity or outcomes in ways that do not reflect genuine research capability and do not contribute materially to future research capabilities.

In addition, the cost of the exercise on all participants must be justified by its benefits. A more selective exercise that yields useful information may be preferred if it releases research and administrative resources that would otherwise be diverted to the production of additional information of lower value. Nevertheless, if institutions that are eligible to apply for, say, ARC and NH&MRC funding have an opportunity to nominate the complement of staff they wish to enter in the assessment exercise, it is possible that the subset submitted for assessment will be broadly reflective of those that are successful in the fully competitive granting schemes conducted by the ARC and the NH&MRC. If that is the outcome of the exercise, the balance of costs and benefits in achieving it is likely to be questioned. We reiterate that, if institutions are offered the opportunity to enter only a subset of staff members employed as part- or full-time researchers, the risk of counterproductive manipulation of that subset is likely to be increased.

3.3 Link to training of researchers

Issue 8: The training received by higher degree students in research requires a separate quality audit and/or assessment process.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

Additional comments

Research training and research have always been tightly coupled in universities and, as stated previously, the ARC believes that one of the key outcomes of publicly funded research is the generation of a highly skilled workforce. Furthermore, anecdotally, it is widely believed that 50 per cent or more of the research activity in universities is conducted by graduate students.

The assessment of the quality and impact of research training in Australia could usefully be expanded, whether or not it occurs in the context of an RQF. Inclusion of post-doctoral staff as well as higher degree research students might also be desirable.

However, the practical issues of decoupling research and research training for assessment purposes are likely to make a separate process undesirable and artificial. In addition, many individual researchers and institutions make decisions about resource allocation between research and training activity which will almost certainly be affected if the quality assessment process excludes research training or treats it under a separate process.

More fundamentally, research training can be considered an investment in future research capability that should be included in any measure of a nation's research achievement.

At minimum, a 'context statement' by university schools or departments indicating the proportion of their research staff that consists of higher degree students, post-doctoral staff or early-career researchers would assist in interpreting the research assessments obtained.

It should be noted that several components of the current DEST block funding grant to universities are heavily influenced by quantitative (not necessarily qualitative) performance in research training. While the ARC believes a measurement of the quality of research training is imperative in any assessment of the university sector's performance, the question of whether it should be performed as a separate exercise depends to a large extent on what funds are in play in the future RQF. If aspects of the current block funding components relating to research training are not included in the funds that are available as a result of the RQF, then a separate exercise is perhaps warranted. On the other hand, exclusion of such funds from the exercise without an indication of the size and origin of alternative funding 'in play' could lead to some degree of cynicism about the whole exercise and legitimate questions about the balance of costs and benefits.

On balance, the ARC's preference is for research training to be included in the RQF.

However, because this issue is likely to have major implications for any final RQF, the ARC suggests that the approach of those countries which have implemented research assessment exercises should be considered in detail.

3.4 Focus of assessment

Issue 9: Assessment for an RQF should include a forward-looking strategic element as well as being based on past performance.

- Strongly agree Strongly disagree No comment
 Somewhat agree Somewhat disagree

Additional comments

It seems obvious that the assessment process should provide incentives for research and funding bodies to consider near future as well as past and current research activity. Where increases in current activity may be at the expense of future capability (for example, where equipment is run down or new research staff are not trained), any assessment mechanism focussing solely on current activity is likely to generate potentially adverse incentives and outcomes.

This concern is clearly dependent on the interval between assessments. If the interval is six years as foreshadowed in the UK assessment, then it would appear that a forward-looking element is required. To exemplify this aspect, the ARC notes that the relative performance of institutions in the competitive ARC schemes can evolve significantly over six years. For example, since 1998, the share of total ARC funding awarded to one particular Australian university has more than doubled, to reach around 4 per cent of the total, while another's has more than halved, to less than 1 per cent of the total.

The assessment mechanisms used in the ARC's competitive schemes incorporate forward-looking elements by examining proposed future activity as well as track record. However, it may be harder to introduce forward-looking components into a broader assessment process. Ideally, a variety of activity would be captured, from research training and the mentoring of early-career researchers to infrastructure investment and international collaborations. There may also be merit in examining strategic decisions made in the past and the impact they had on research profile, focus or capability.

If a variety of disparate activities are included, and if their likely impact on future research capability is unclear, it may be difficult to find a way of evaluating and comparing their contribution to overall research performance. A simple listing of the activities may be all that can reasonably be expected.

For this reason, the ARC submits that the final rating should be informed predominantly by current and past performance (including past performance in implementing strategic change), rather than the forward-looking component.

3.5 Reporting arrangements

Issue 10: How should the outcomes of an RQF be reported?

(a) Reporting the outcomes of an RQF should be aligned to:

- Subject/discipline areas
- Research grouping/research team/s
- Department/schools
- Faculties/Divisions
- Institutional level – university/PFRA

(b) Reporting on subject/discipline areas within any level of aggregation for the RQF should be aligned to the ABS RFCD codes or an appropriate subset.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

Additional comments

(a) In the ARC's view, the results of the assessment process should be reported at both the highest levels of aggregation possible (whole-of-institution, whole-of-discipline and, if possible, whole-of-nation) and in disaggregations aligned with the purpose of the exercise (e.g. by institution or research grouping if funding is to be allocated to such bodies on the basis of the assessment outcomes). The relative strengths of particular institutions, of research disciplines and of Australia's research effort overall are likely to be of considerable policy, educational and industry interest, independent of any funding implications for individual organisations.

If it is considered undesirable to publish simple rankings of institutions, it may still be possible to publish or report other information gathered in the course of the assessment in order to inform policy and other deliberations.

An important consideration with respect to reporting is that institutional rankings alone could disguise pockets of real strength within a larger organisation. In the most undesirable manifestation of such a scenario, students with a real interest in say discipline A could be attracted to study at Institution X with a higher overall ranking than Institution Y, notwithstanding that the former scored much lower in discipline A than the latter. Although the primary purpose of the RQF exercise is to allocate institutional funding based on research excellence, such unintended outcomes should be considered carefully.

(b) Alignment of reporting with standard codes has several advantages:

- it separates the reporting framework from the institutional structures within which research is organised,
- it provides continuity and therefore comparability over time, so that trends and progress can be identified,
- to the extent that the codes themselves align with international codes, and to the extent that the institutional base and definitions against which information is collected align with international practice, it enables international comparison of trends and measures.

Current ARC data collection processes use the RFCD. The use of the RFCD in an RQF would therefore enable comparability across the two datasets and their reports.

However, any deficiencies in the standard are likely to have increasingly greater implications the more widely they are used. More frequent review of the codes than is presently undertaken may be desirable.

Consideration of the use of some supplementary categories (e.g. new and emerging research fields, areas of national research priority) may also be warranted in any new framework. This is particularly so in respect of interdisciplinary research. The dynamic nature of this situation can be illustrated with the emergence of biochemistry. Now a recognised discipline in its own right, it would have been considered an interdisciplinary area between biology and chemistry in the past.

Issue 11: What should be the format of the ratings/rankings/benchmarks of an RQF? Please provide examples.

The format chosen will need to reflect the purpose of the exercise, but ideally should provide for some or all of: (a) international comparisons, (b) cross-institutional comparisons, and (c) cross-disciplinary comparisons, including comparisons over time, so that areas of relative strength and emerging trends can be identified.

If the results are to be compared against international benchmarks, the scales used, the type of information collected, the groups and research base from which they are collected and the weightings applied to individual components will themselves need to be internationally comparable.

The ARC recognises the ‘boundary’ problems associated with any single-measure points-based rating system when quality varies on a continuum rather than in quantum intervals. The implications for funding of achieving a rating of, for example, 3.9 rather than 4.1 may be very serious indeed and this should be avoided if at all possible. Institutions will have an incentive to achieve the highest ranking or rating possible and so can be expected to divert resources and activity in ways that will enhance those rankings. The intensity with which they do this will depend on the funding and prestige at stake. These problems can compromise the picture of overall research quality, make the allocation of research funding more difficult and controversial and reduce the utility and acceptance of the overall research assessment process.

Consequently, any scheme chosen will need to balance the benefits and risks of achieving and disseminating useful ranking information. The ARC suggests that publishing ratings against a number of performance indicators, including separate indicators of excellence and impact, rather than a single aggregated measure, is likely to be preferable.

Where ratings are used to evaluate research excellence, the ratings scale should be neither too coarse to enable meaningful differentiation nor too fine to be difficult to comprehend. A relatively continuous scale might be most appropriate to avoid possible contention and unwarranted disturbances at scoring boundaries. For example, if a 1-5 scale is used, it may be better to allow scores such as 3.9 and 4.1 to avoid the more dramatic implications of the perceived difference between 3 and 4 if integral scores only are used. There are probably advantages in having a scale for research excellence overall and another scale for rating individual items within a research excellence profile. A two-pronged approach of this sort would provide an overall rating for the organisation or grouping under assessment while retaining information that would enable pockets or individuals of excellence to be identified within those groupings.

However, ratings for research impact (other than impact on the discipline concerned) are likely to be more difficult to generate and even more difficult to interpret. This is because, as mentioned in the introduction to this response, research impacts typically occur over a long timeframe, involve the interaction of a wide range of individuals and organisations and are influenced by broader economic and social conditions and constraints than research excellence alone.

The ARC notes that some international frameworks, including the UK RAE, appear to give less weight to impact measures – particularly end-user or socio-economic impact – than to intrinsic research excellence. If Australia is seeking a framework that will *inter alia* enable research effort to be directed towards areas of greatest potential benefit, it will need a better-defined means of identifying and tracking impact. Due to its nature, the ARC believes the determination of impact outside relevant disciplines will necessarily have to rely strongly on case-by-case judgements by appropriately constituted panels.

3.6 Links to funding

Issue 12: The resource intensity required for an RQF should be directly related to the level of funding that it informs.

- Strongly agree Strongly disagree No comment
- Somewhat agree Somewhat disagree

Additional comments

While this statement is clearly true, it will be important to ensure that any RQF avoids duplication with existing effective systems and is designed to minimise the burden on participants, consistent with obtaining information suitable to its purpose.

The ARC also notes that an RQF, properly designed and operated, can be expected to fulfil a wider role than simply a means of allocating funding. It could identify new and emerging areas of excellence. It could also serve to identify discipline areas of national importance that are in general decline and, notwithstanding their overall poor ranking, paradoxically in need of strategic funding in the national interest. Finally, it should also generate value as a monitoring and evaluation tool, indicating the fields in which Australian researchers are active and achieving excellence and, to the greatest extent possible, the benefits of that research effort for the Australian economy and community as a whole.

3.7 Administrative benefits

Issue 13: An RQF ought to lead to commensurate reductions in reporting requirements for other Australian Government research accountability mechanisms.

- Strongly agree Strongly disagree No comment
 Somewhat agree Somewhat disagree

Additional comments

Where the RQF would duplicate information requirements currently managed through other mechanisms, then those requirements should not increase. However, if the RQF is to provide information that is currently not available, or not available in comparable, reliable or accessible forms, then it can be expected to add to the quantum of information collected from the research sector. As noted earlier, the cost of any new reporting requirements must be clearly justified by their expected benefit.

Existing grant administration schemes can be expected to continue to require reports from grant recipients in order to satisfy their own accountability requirements. However, if these reports are at the level of individual funded research projects, they will not necessarily duplicate RQF reporting requirements, which are likely to be at the level of larger research groupings or fields.

The ARC is of the view that there is broad agreement that the approaches it takes, along with the NH&MRC, to judge the excellence of individual project applications are well respected, robust and international in nature. Rather than attempting to duplicate these ongoing annual assessments, a future RQF must wherever possible take advantage of the significant information gathered by these expert-based approaches to defining excellence at the project level.

OTHER COMMENTS

If you have additional ideas or comments on areas not addressed in the paper we invite you to provide these in this submission.

Scope and purpose of the RQF

As noted throughout this document, the ARC believes that definitive design and implementation elements of the RQF can only be determined once there is a clear understanding of the purposes the RQF will serve, including the nature and quantum of the funding decisions that it will inform.

The frequency of RQF assessment

The ARC notes that the frequency of the assessment included under the RQF will influence *inter alia* its cost, the information that can and should be collected and the need for a forward-looking element. The ARC suggests that a 6 year interval may provide a good balance of costs and benefits. However, where particular research institutions are evolving rapidly – for example, a new generation university undergoing rapid change – or are affected by special circumstances, it may be possible for a review to foreshadow a further, interim review of a lighter nature, ahead of the next full review, to re-examine the rating given.