MINISTRY OF EDUCATION
Te Tāhuhu o te Mātauranga

In pursuit of excellence

Analysing the results of New Zealand's PBRF Quality Evaluations

New Zealand Government

Research and knowledge creation

This report forms part of a series called Research and knowledge creation.

Authors
Warren Smart, Principal Research Analyst
Tertiary Sector Performance Analysis, Ministry of Education

Ralf Engler, Senior Research Analyst
Tertiary Sector Performance Analysis, Ministry of Education

Corresponding author
Warren Smart
Email: warren.smart@minedu.govt.nz
Telephone: 04-463-8035
Fax: 04-463-8713

Acknowledgements
The authors gratefully acknowledge comments provided by reviewers at the Ministry of Education and Tertiary Education Commission. The authors also gratefully acknowledge the assistance provided by Sean Alexander (Tertiary Education Commission) in producing the datasets used in this analysis. The authors also gratefully acknowledge Alison Lipski, who proof-read this report.

All views expressed in this report, and any remaining errors or omissions, remain the responsibility of the authors.

Published by
Tertiary Sector Performance Analysis
Tertiary, International and System Performance
MINISTRY OF EDUCATION

This report is available from the Ministry of Education's Education Counts website: www.educationcounts.govt.nz.

June 2013

ISBN (web) 978-0-478-40657-3

In pursuit of excellence

FIGURES

TABLES

1 INTRODUCTION

Between 2004 and 2016, the Performance-Based Research Fund (PBRF) will allocate around $2.7 billion in funding to participating New Zealand tertiary education organisations (TEOs). This represents a substantial investment by the Government in tertiary education.

A key objective of the PBRF is to raise the 'average quality of research'. To assess the extent to which the average quality of research has improved, this report analyses the results of the three PBRF Quality Evaluations. We also examine the types of research outputs and categories of peer esteem and contribution to the research environment items submitted in evidence portfolios. There is also an analysis of researcher pathways over time.

This report complements a previous report by the Ministry of Education[1] that examined the quantitative evidence on the impact of the PBRF in a variety of areas ranging from bibliometric measures to completion rates of research degrees.

This report has the following structure:

- In Chapter 2, we present a brief history and overview of the PBRF. This includes the current aims of the PBRF and describes how the three component measures of the PBRF work.
- In Chapter 3, we outline the data and method used to analyse the Quality Evaluation results.
- In Chapter 4, we analyse changes in measured quality between 2003 and 2012.
- In Chapter 5, we use statistical modelling to look at the factors associated with higher research quality in the 2012 Quality Evaluation.
- In Chapter 6, we look at the pathways of staff who participated in the 2012 Quality Evaluation.
- In Chapter 7, we take a closer look at staff who received an 'A' in the Quality Evaluations of 2006 and 2012.
- In Chapter 8, we examine what happened to new staff with an emerging research profile in the 2006 Quality Evaluation.
- In Chapter 9, we look at the types of research output, categories of peer esteem and categories of contribution to the research environment submitted in the Quality Evaluations in the 2012 Quality Evaluation.
- In Chapter 10, we investigate the correlation between the research output, peer esteem and contribution to the research environment component scores in the 2012 Quality Evaluation.
- In Chapter 11, we look at the profile of staff who submitted evidence portfolios to the Professional and Applied and the Pacific Research Expert Advisory Groups in the 2012 Quality Evaluation.
- In the Appendix, we present a detailed description of the PBRF Quality Evaluations.

[1] See Smart (2013).

2 THE PERFORMANCE-BASED RESEARCH FUND

Background

The PBRF was established in 2002. It is intended to ensure that excellent research in the tertiary education sector is encouraged and rewarded. This entails assessing the research performance of tertiary education organisations (TEOs) and then funding them on the basis of that performance. Only degree-granting TEOs are eligible to participate in the PBRF. All universities and some polytechnics, wānanga, and private training establishments participate in the PBRF.

Aims of the PBRF

The Government's aims in introducing the PBRF were to:

- increase the average quality of research
- ensure that research continues to support degree and postgraduate teaching
- ensure that funding is available for postgraduate students and new researchers
- improve the quality of public information on research outputs
- prevent undue concentration of funding that would undermine research support for all degrees or prevent access to the system by new researchers
- underpin the research strength in the tertiary education sector.

The focus in this report is assessing if the first objective – increase the average quality of research – has been met.

The three components of the PBRF

The PBRF funding formula is based on three indicators that together assess both quantity and quality of research:

a. **Quality Evaluation**: the assessment of the research quality of TEO staff, based largely on peer review of a researcher's evidence portfolio (EP) of research outputs, accounting for 60 percent of the fund

b. **Research degree completions**: the number of postgraduate research-based degrees completed in the TEO, accounting for 25 percent of the fund

c. **External research income**: the amount of income for research purposes received by the TEO from external sources, accounting for 15 percent of the fund.

The first Quality Evaluation took place in 2003, followed by a partial round in 2006.[2] The third Quality Evaluation was undertaken in 2012.

More detailed information on how the Quality Evaluation operates can be found in the Appendix.

[2] The 2006 Quality Evaluation is called a 'partial' round because those who participated in the 2003 Quality Evaluation and were PBRF eligible in 2006 were given the option of not submitting an evidence portfolio in 2006 – instead, carrying their 2003 result forward.

3 METHOD

Introduction

This section outlines the methodological approach that we use in this report to measure changes in research quality. Specifically we:

- describe how the PBRF Quality Evaluations have changed over time and how these changes impact on our analysis
- present how we have calculated a suite of average quality scores
- discuss the statistical modelling approach we have used to assess if changes in measured research quality were statistically significant.

Differences between PBRF Quality Evaluations

Between 2003 and 2012, there have been several changes to the way the PBRF Quality Evaluations have been run. The main differences are outlined in Table 1 and include: a new way of assessing new and emerging researchers from 2006, not collecting data on all PBRF-eligible staff in 2012, and a changing number of TEOs participating in the Quality Evaluations. In addition, the 2006 Quality Evaluation was a partial round, where not all eligible researchers were required to submit a new evidence portfolio, but, instead, they carried over their quality categories and component scores obtained in 2003.

Table 1
Summary of differences between PBRF Quality Evaluations

Difference	2003	2006	2012
New and emerging category included	No	Yes	Yes
Partial round	No	Yes	No
All PBRF-eligible identified in staff census	Yes	Yes	No
Number of participating TEOs	22	31	27

In assessing if the quality of research has increased over time, we need to take these differences in the three Quality Evaluations into account. In particular, we do not know who all of the PBRF-eligible staff were in the 2012 Quality Evaluation and we require a consistent method of selecting the staff to include in the analysis.[3] We take the following approaches to ensure a consistent study population:

- **Analysis A: Progress between 2003 and 2012:** We assess changes in research quality between 2003 and 2012 for those staff who submitted evidence portfolios and received a minimum of a 'C' quality category. Staff who received an 'R', 'C(NE)' or 'R(NE)' quality category are excluded from the analysis.

  For the analysis of trends between 2003 and 2012, we include only those TEOs that participated in all three Quality Evaluations.

- **Analysis B: Progress between 2006 and 2012:** We are also interested in how a wider group of researchers has performed under the current rules (where the provisions for new and emerging researchers are in place), so we also examine trends in research

[3] See TEC (2012) for the reasons why data on all PBRF-eligible staff was not collected for the 2012 Quality Evaluation.

quality between 2006 and 2012 for staff who received at least a 'C(NE)' quality category. This approach examines changes in the quality of measured research under the existing system of rules (i.e. with the provision of new and emerging staff in place).

For the analysis of trends between 2006 and 2012, we include those TEOs that participated in those two Quality Evaluations.

The list of TEOs used in each of the analyses is presented in Table 2.

In any analysis including the 2006 partial Quality Evaluation, we transferred the 2003 quality categories and component scores of staff that had evidence portfolios over to the 2006 Quality Evaluation.

Table 2
List of TEOs in the trend analysis

	University	Polytechnic	Wānanga	PTE
2003-2012	Auckland University of Technology	Unitec		AIS St Helens
	Lincoln University	Wintec		Bethlehem Tertiary Institute
	Massey University			Carey Baptist College
	University of Auckland			Laidlaw College
	University of Canterbury			Whitecliffe College of Arts and Design
	University of Otago			
	University of Waikato			
	Victoria University of Wellington			
2006-2012	Auckland University of Technology	CPIT	Te Whare Wānanga o Awanuiārangi	AIS St Helens
	Lincoln University	EIT		Bethlehem Tertiary Institute
	Massey University	MIT		Carey Baptist College
	University of Auckland	Northland Polytechnic		Good Shepherd College
	University of Canterbury	Open Polytechnic		Laidlaw College
	University of Otago	Otago Polytechnic		Whitecliffe College of Arts and Design
	University of Waikato	Unitec		
	Victoria University of Wellington	Wintec		
		Whitireia Community Polytechnic		

Note: Colleges of education are treated as being part of the universities.

Calculation of average quality scores

One of the key measures of trends in research performance as measured in the PBRF is the *average quality score*.

In the earlier Quality Evaluations, there was a simple and common approach to calculation of the average quality score – we used as a denominator the number of PBRF-eligible staff. This

gave us a reading on the quality of research produced on average by the population of PBRF-eligible people. But in 2012, the TEC did not collect data on the number of PBRF-eligible staff. So we cannot use that measure to track trends between 2003 and 2012. Instead, we use four new averages.

The numerator of all four average quality score measures is a full-time equivalent (FTE) weighted score, where the score for an 'A' quality category = 10, 'B' = 6, 'C' = 2 and 'C(NE)' = 2. The numerator is given by this formula:

$$\sum(\textit{Quality score} \times \textit{FTE weighting})$$

To calculate an average quality score we use four different denominators. These are:

- the FTE weighting of the staff whose evidence portfolios were submitted as part of the Quality Evaluation and who received 'C' or above (for the analysis of progress between 2003 and 2012), or 'C(NE)' or above for the analysis of performance between 2006 and 2012
- the number of full-time equivalent academic and research only staff reported annually to the Ministry of Education in the annual statistical collection
- the number of bachelors degree or higher equivalent full-time students
- the New Zealand resident population aged 15 or older.

The *first* denominator gives a sense of how the quality of the evidence portfolios considered by the panels has changed. The *second* gives us a reading of the extent of change in research quality over the population of academic or research only staff in participating TEOs. The *third* takes account of the statutory requirement that degree-level teaching should be taught by staff mainly active in research; in effect, it is a measure of the extent to which degree teaching in the sector is supported by research. The *fourth* takes a national perspective, assessing how the research activity of the TEOs is growing in line with New Zealand's adult population.

The formula for the calculation of the average quality scores is presented in Table 3.

Table 3
Formula for the calculation of average overall quality score

Values	Formula
A = 10, B = 6 C = 2, C(NE) = 2	$\frac{\sum(\textit{Quality score} \times \textit{FTE weighting})}{\textit{Various denominators}}$

Component scores

In order to help with the process of assigning quality categories to evidence portfolios, the PBRF panels arrive at scores for the quality of research outputs, peer esteem and contribution to the research environment. Each of these components receives a score between 0 and 7. Then a formula is applied to arrive at a single score to help with the assignment of quality categories. The weighting is 70 percent for the research output score, 15 percent for the peer esteem score and 15 percent of the contribution to the research environment score.[4]

We calculate an FTE-weighted average score for the research output score (AROS), peer esteem score (APES) and contribution to the research environment score (ACRES) (see Table 4). For

[4] More detail on these scores is presented in the Appendix.

the analysis of component average scores between 2003 and 2012, the denominator in this analysis is the number of FTE-weighted PBRF staff at 'C' or above. For the analysis of average scores between 2006 and 2012, the denominator in this analysis is the number of FTE-weighted PBRF staff at 'C(NE)' or above.

Table 4
Formulas for the calculation of component average quality scores

Measure	Values	Formulas
AROS	0-7	$\frac{\sum(\text{Research output score} \times \text{FTE weighting})}{\sum \text{FTE weighting}}$
APES	0-7	$\frac{\sum(\text{Peer esteem score} \times \text{FTE weighting})}{\sum \text{FTE weighting}}$
ACRES	0-7	$\frac{\sum(\text{Contribution to research environment score} \times \text{FTE weighting})}{\sum \text{FTE weighting}}$

Testing if the change in average quality is statistically significant

Three sets of analysis were used to test the statistical significance of the PBRF results: one to look at the 2003-2012 results; another to look at the 2006-2012 results; and a third to consider the 2012 results alone. In these analyses, we consider individual researchers, not weighted by their full-time equivalent status.

To determine if there was a significant difference in the likelihood of a researcher receiving a particular quality category, or score, the PBRF data was modelled using logistic regression.[5]

Many of the researchers were present in two or more PBRF rounds, so the regression models controlled for this lack of independence.

For the analysis of the 2003-2012 results, the study population consisted of those researchers who provided an evidence portfolio in any of the three PBRF rounds, and who were employed by a tertiary education organisation which participated in the PBRF round in each of these three years. Researchers who received an 'R' or 'C(NE)' quality category were excluded. There were 15,114 people in the 2003-2012 researcher study population.

For the analysis of the 2006 and 2012 rounds, the study population consisted of those researchers who provided an evidence portfolio in these years, and who were employed by a tertiary education organisation that participated in the PBRF round in these two years. Researchers who received an 'R' or 'R(NE)' quality category were also excluded, but those with a 'C(NE)' rating *were* included. There were 12,472 people in the 2006-2012 researcher study population.

For our analysis of the factors associated with PBRF performance, we used the researchers in the 2012 PBRF round. To ensure there were sufficient numbers of researchers in the different combination of factors, the analysis was limited to researchers in the lecturer, senior lecturer, associate professor and professor positions. New and emerging researchers who received a 'C(NE)' or 'R(NE)' were also excluded, as were those with missing information about gender and birth dates. There were 5,655 people in the 2012 researcher study population.

[5] The dependent variable was the likelihood of a researcher receiving a particular quality category or score. The different categories or scores were regarded as nominal (not having any intrinsic order). This was done, rather than considering them as ordinal (where one category is regarded as 'higher' than another), because, with the software we had available to us, the models that controlled for the fact that researchers were present in multiple years could not handle dependent ordinal variables. Regarding the dependent variable as ordinal does not affect the results or conclusions.

For each set of regression analyses, four dependent variables were considered. These were the quality score, the research output score, the peer esteem score and the contribution to the research environment score. For the 2003-2012 and 200-2012 study populations, the quality categories were ‘A’, ‘B’ and ‘C’. For the 2012 study population, this also included the ‘R’ quality category. For the three ‘score’ quality measures, the categories were ‘0-3’, ‘4-5’ and ‘6-7’ in each study population.

Data

The data for this analysis was provided by the Tertiary Education Commission (TEC) under agreed confidentiality protocols. We received the dataset on 18 March 2013. Our data differs very slightly from the final dataset used by the Tertiary Education Commission in reporting results and allocating funding although the differences are minor and do not affect the conclusions drawn in this paper.

There were a number of cases where there was missing data. The number of missing birth dates was 158 in 2003, 231 in 2006 and 177 in 2012. There were 11 researchers in 2012 with missing gender information. We do not think this missing data has any material effect on the analysis, the results, or the conclusions reached.

4 TRENDS IN THE QUALITY OF RESEARCH

KEY POINTS

We analysed the results of the PBRF Quality Evaluations. The results showed that:

- there was a statistically significant increase in the average quality of research between 2003 and 2012
- the rate of increase between 2003 and 2006 was sustained between 2006 and 2012
- when examining the individual component scores assigned to evidence portfolios, we found that the improvements in the component scores between 2003 and 2012 were all statistically significant
- improvements in performance in the peer esteem and contribution to the research environment scores were greater than in the improvement in the research output score. These improvements may be due to both better performance in these components and/or clearer and more effective presentation of material on these components in evidence portfolios
- the overall improvement in quality correlates with improvements in the rates of citation of research by New Zealand tertiary education institutions since the PBRF was introduced, suggesting that the increases in the quality measures over the three PBRF Quality Evaluations reflect a true lift in performance.

Introduction

In this section, we examine changes in the average quality of research between 2003 and 2012.[6] Specifically, we analyse the quality categories assigned to evidence portfolios and also the component scores used to assist in the holistic process of assigning quality categories. We also apply statistical modelling to see if any changes in research quality were statistically significant.

We examine two periods. First, we examine changes in quality between 2003 and 2012. For consistency, we assume the rules as they applied in 2003 (i.e. no new and emerging researchers) had continued in 2006 and 2012. Then we examine changes in average quality between 2006 and 2012 under the existing set of rules (i.e. including new and emerging researchers).

Changes in the quality of research between 2003 and 2012

This section analyses changes in research quality between the first PBRF Quality Evaluation in 2003 and the most recent in 2012. We define the dataset used in this analysis in Table 5. We only include evidence portfolios that received an 'A', 'B' or 'C' quality category (in other words we assume that the rules in place in 2003 had continued in the 2006 and 2012 Quality Evaluations) and limit our analysis to TEOs that participated in all three Quality Evaluations.

Table 5
Data used in the analysis of trends in research quality 2003 and 2012

Quality categories used in this analysis	A, B and C
TEOs used in this analysis	Those that participated in all three Quality Evaluations (see Table 2 for the full list)

[6] Please note that when we discuss changes in research quality, we are referring to changes in 'measured' research quality.

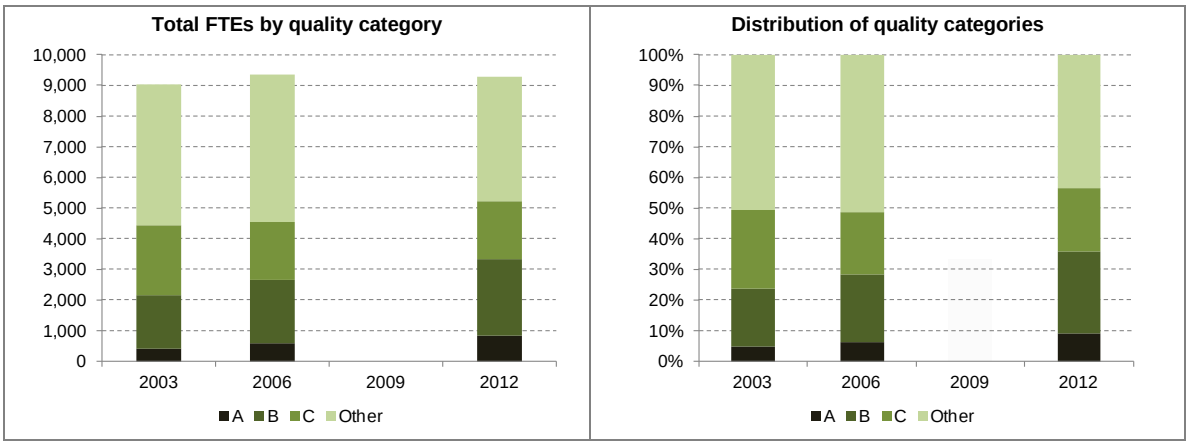
Results

Figure 1 shows the total number of staff FTEs and the proportion of FTEs by quality category, using the annual staffing statistical collection to define the total staff population. The 'Other' category in Figure 1 is calculated by subtracting the 'A', 'B' and 'C' FTEs from the total academic and research only FTEs for the TEOs in this analysis.

The academic and research only FTEs include all academic staff at TEOs and will include staff who teach at below degree level, although this number will be small for the universities. Nevertheless, it provides an element of consistency in the denominator which is crucial to drawing conclusions about any changes in the measured quality of research.

Figure 1

Distribution of quality categories using total academic and research only FTEs to define the staff population 2003-2012



Note: The 'Other' category is calculated by subtracting the total of 'A's, 'B's and 'C's from the total number of academics and research only staff.

The results show that:

- between 2003 and 2006, the proportion of staff who received at least a 'C' quality category remained relatively constant, at around 49 percent, but the composition of quality categories improved. The proportion of 'A's increased from 5 percent to 6 percent and the proportion of 'B's from 19 percent to 22 percent. Conversely, the proportion of 'C's fell from 26 percent to 20 percent

- between 2006 and 2012, the proportion of staff receiving at least a 'C' quality category increased from 49 percent in 2006 to 56 percent in 2012. The proportion of 'A' and 'B' quality categories increased, from 6 percent to 9 percent and 22 percent to 27 percent, respectively. The proportion of 'C' quality categories remained unchanged at 20 percent.

So between 2003 and 2006, there was shift in the mix of 'A', 'B' and 'C' quality categories, but little change in the overall proportion of quality categories of at least a 'C'. But between 2006 and 2012, there was a substantial increase in the proportion of staff assigned at least a C quality category.

The 2006 Quality Evaluation was a partial round and was undertaken only three years after the first Quality Evaluation. There was evidence that the quality of presentation of evidence portfolios improved, resulting in higher grades (TEC 2007, Smart 2008b). So while the proportion of staff receiving at least a 'C' didn't change, there was a shift towards higher quality categories. However, in 2012, there was an increase in the proportion of staff who received at least a 'C' quality category.

This may reflect that it has taken time for the TEOs to make inroads into the proportion of staff who were not receiving at least a 'C' quality category. This could be through either improving the quality of existing staff or employing staff from outside the system with better research performance.

In 2006, 85 percent of staff with a 'C' quality category or higher had participated in the previous Quality Evaluation. In 2012, the figure was 61 percent. Given the greater period between Quality Evaluations, the smaller proportion in the 2012 Evaluation that participated in the previous Quality Evaluation is not unexpected. However, it does suggest that a significant proportion of the increase in a 'C' or higher quality category was staff new to the Quality Evaluation. The pathways of staff in the Quality Evaluations are examined in more depth in Chapter 6.

Trends in average quality score

In Table 6 and Figure 2 we present the average overall score using a variety of denominators: FTE academic and research only staff, EFTS at degree level or higher, and the New Zealand population aged 15 and over. In each case, the average score increased over time.

One difference among the measures is in the rate of improvement. Whereas the measure of average quality using academic and research only FTEs and New Zealand population as denominators shows a relatively constant rate of improvement, the measure using degree-level or higher EFTS as the denominator shows a decline in the *rate* of improvement between 2006 and 2012. This declining rate is partly a result of a higher student per staff ratio in 2012 compared with 2006. In other words, the TEOs were carrying a higher number of students because of higher participation, reflecting the shift in the balance of enrolments towards higher-level qualifications over the period 2005-2012.[7]

These improvements in the average quality of research correlate with improvements in the bibliometric performance of New Zealand tertiary education institutions (TEIs). Smart (2013) found that the rate of citation of research measured in the Thomson Reuters Web of Science by New Zealand TEIs has increased since the PBRF was introduced.

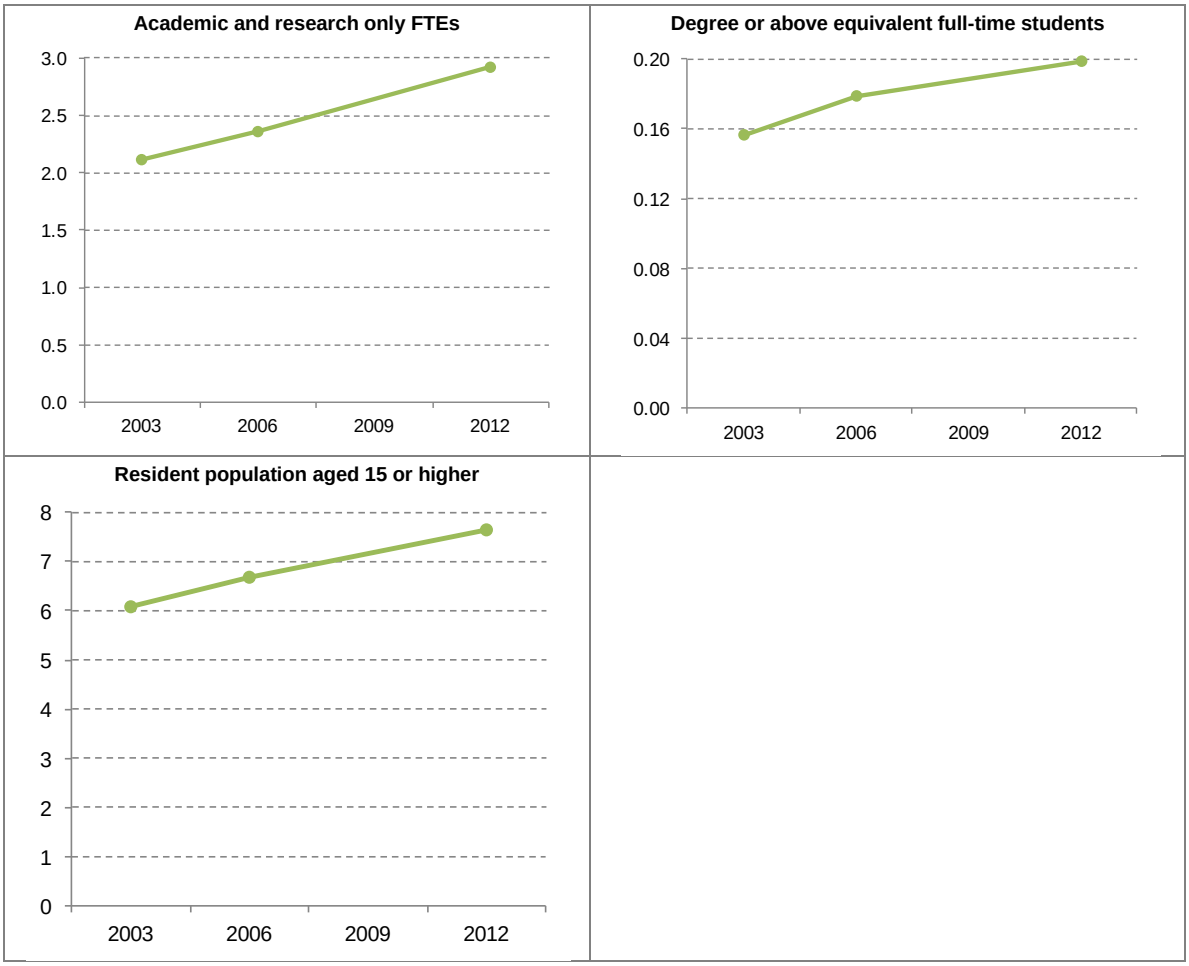
Table 6

Average overall score calculated using various denominators 2003-2012

Denominator	2003	2006	2012
Academic and research only FTEs	2.1	2.4	2.9
EFTS (degree and above)	0.16	0.18	0.20
Population 15 and over (in 000s)	6.1	6.7	7.6

[7] See Wensvoort (2012).

Figure 2
Average quality scores using various denominators 2003-2012



We now limit our study population to staff that received at least a 'C' quality category. Although this means we are analysing a narrower group of staff than in the previous section, this is done to ensure we have a consistent study population over time. In limiting the analysis in this way, we are examining the average quality of research for the group of people who are assessed as having conducted research at a level that attracts funding (under the original 2003 rules). We also extend our focus to include the component scores assigned to evidence portfolios.

Figure 3 and Table 7 show the trends in average research quality between 2003 and 2012 for those that received at least a 'C' quality category. The results show that there has been an increase in the assessed quality of research in this group, although the rate of improvement slowed between 2006 and 2012. Specifically, the data shows that:

- on an FTE-weighted basis, the number of staff receiving at least a 'C' quality category increased by 2 percent between 2003 and 2006 and 15 percent between 2006 and 2012. Overall, this was an increase of 18 percent between 2003 and 2012

- the number of FTE-weighted 'A's and 'B's continued to grow between 2003 and 2006 and 2006 and 2012. In total, the number of 'A's increased by 96 percent between 2003 and 2012, while the number of 'B's increased by 46 percent. However, the number of FTE-weighted 'C's fell by 17 percent between 2003 and 2006 and remained stable between 2006 and 2012

- the proportion of FTE-weighted ‘A’s increased from 10 percent in 2003 to 16 percent in 2012. Similarly, the proportion of ‘B’s increased from 39 percent in 2003 to 49 percent in 2012. The proportion of ‘C’s fell from 52 percent in 2003 to 36 percent in 2012

- although there was further improvement in research quality between 2006 and 2012, the rate of improvement slowed. The average overall score increased from 4.3 in 2003 to 4.9 in 2006, but then increased at a slower rate to reach 5.2 in 2012

- when we examine the underlying component scores, improvements in the peer esteem and contribution to the research environment scores have been important contributors to the overall improvement in quality.

The slow-down in the rate of increase in average overall quality in 2012 was different from what we found using academic and research only FTEs as the denominator (see Figure 2). This is because the reduction in staff receiving less than a ‘C’ quality category is not captured when focusing only on staff who received at least a ‘C’. Hence, the increases in average quality between 2006 and 2012 are likely to be understated by the study population in this case.

The faster rate of increase in the peer esteem scores and contribution to the research environment scores, compared with the research output score, is likely to reflect increased familiarity with the system of evaluation. The TEC (2007) and Smart (2008b) indicated that improvement in the presentation of evidence portfolios was one potential reason for the improvement in scores between 2003 and 2006.

The TEC (2012) noted that there were high correlations in the 2012 Quality Evaluation between peer esteem scores and research output scores and between contribution to the research environment scores and research output scores. They argued that these high correlations plus the lower difference in the scores for the three components suggest that the faster rate of increase in these components reflects a true increase in peer esteem and in the quality of the contribution to the research environment – rather than better presentation of evidence.

However, there is no definitive way of knowing for sure how much of the increase in peer esteem and contribution to the research environment is due to better presentation and how much reflects a true lift in quality. Both are likely to have played a role in the improvement in average research quality.

Table 7

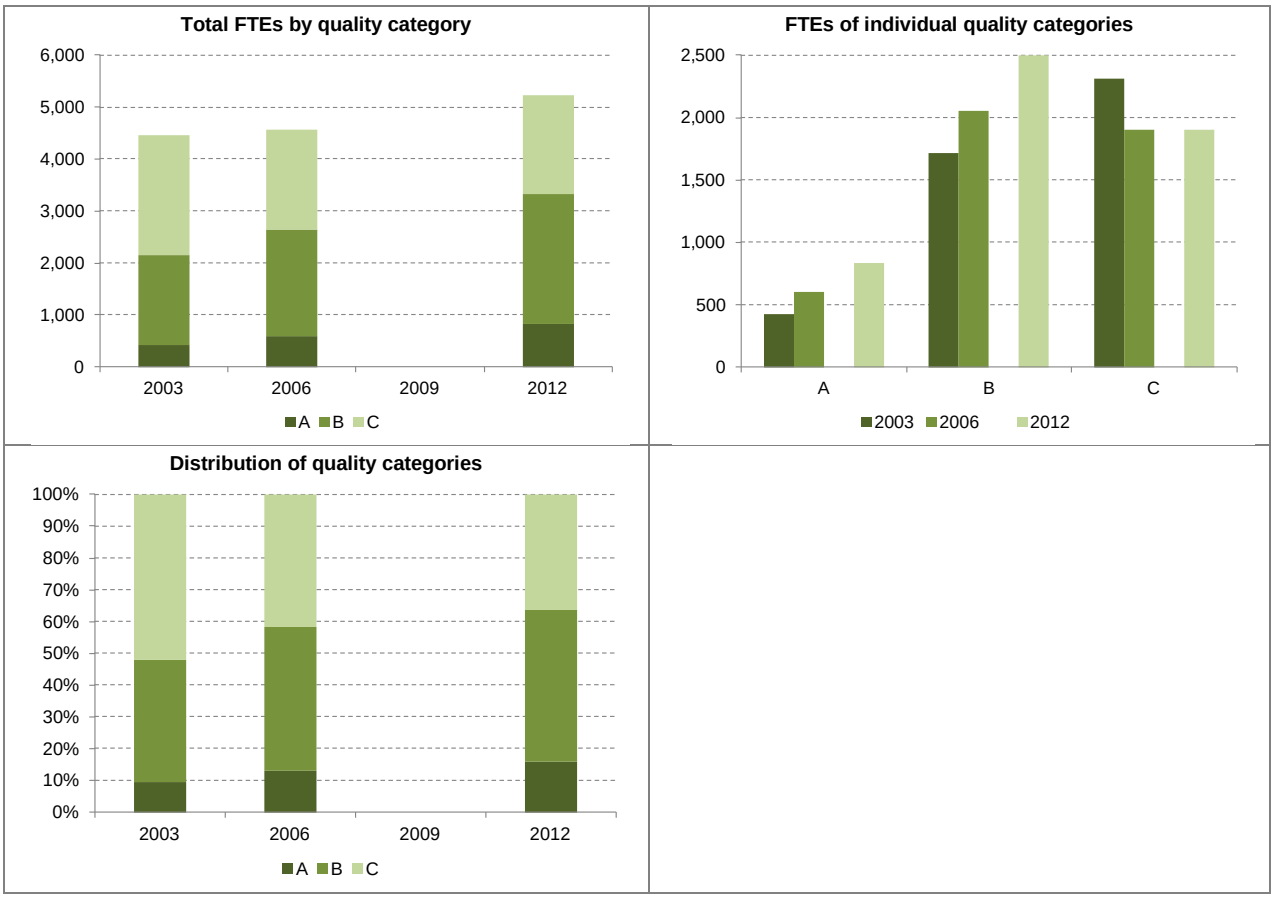
Changes in the assessed quality of research 2003-2012 (FTE weighted)

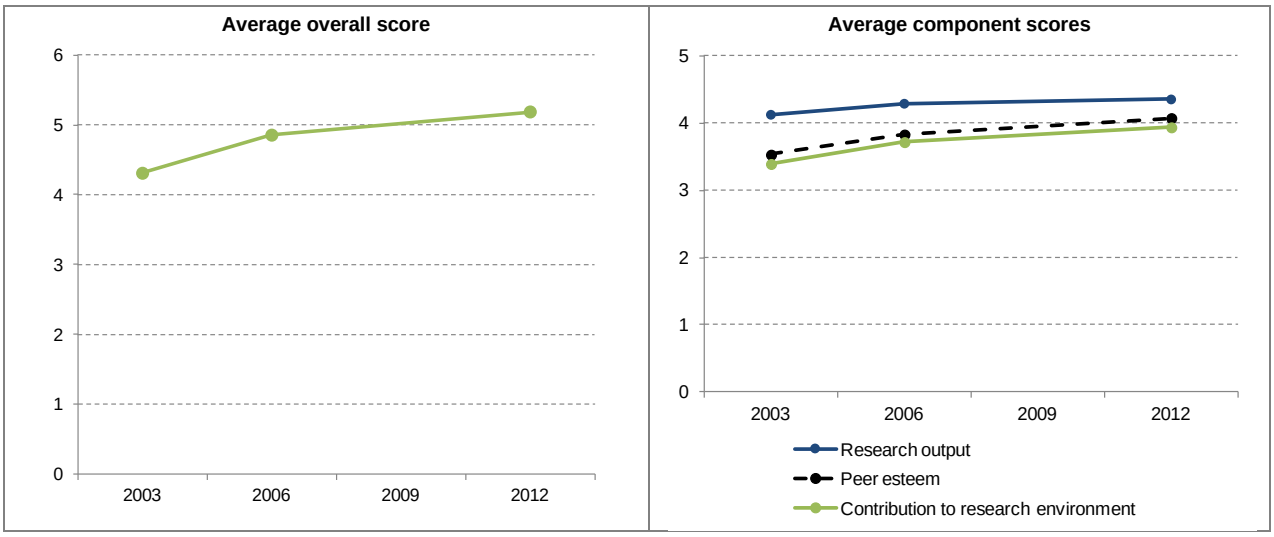
Measure (FTE weighted)	Category	2003	2006	2012
Quality category	A	10%	13%	16%
	B	39%	45%	48%
	C	52%	42%	36%
	Total	**100%**	**100%**	**100%**
Average overall score		4.3	4.9	5.2
Average research output score		4.1	4.3	4.4
Average peer esteem score		3.5	3.8	4.0
Average contribution to the research environment score		3.4	3.7	3.9
Total FTE		**4,452**	**4,554**	**5,232**

Notes: 1. This data is only for TEOs that participated in all three PBRF Quality Evaluations and for those staff who received at least a 'C' quality category. 2. Percentage totals may not add to 100% due to rounding.

Figure 3

Changes in the assessed quality of research 2003-2012 (FTE weighted)





Note: This data is only for TEOs that participated in all three PBRF Quality Evaluations and those who received at least a 'C' quality category.

Statistical modelling results

The statistical modelling in this section is based on a study population where we include staff who received at least a 'C' quality category. Table 8 shows the results of the statistical modelling on the 2003-2012 study population. These same results are shown in Figure 4. The results for the quality category closely mirror those shown in Table 7, which are weighted by FTE. The close similarity of the results indicates that most researchers were working full-time. The average FTE value was 0.94 and the median FTE value was 1.0.

Each measure of research quality shows much the same pattern. The likelihood of a researcher receiving the lowest category or scores showed a significant decline over the years 2003 to 2006 to 2012, while the likelihood of a researcher receiving a higher category or scores significantly improved.

Table 8

Expected probabilities of receiving a particular quality or score category in 2003, 2006 and 2012

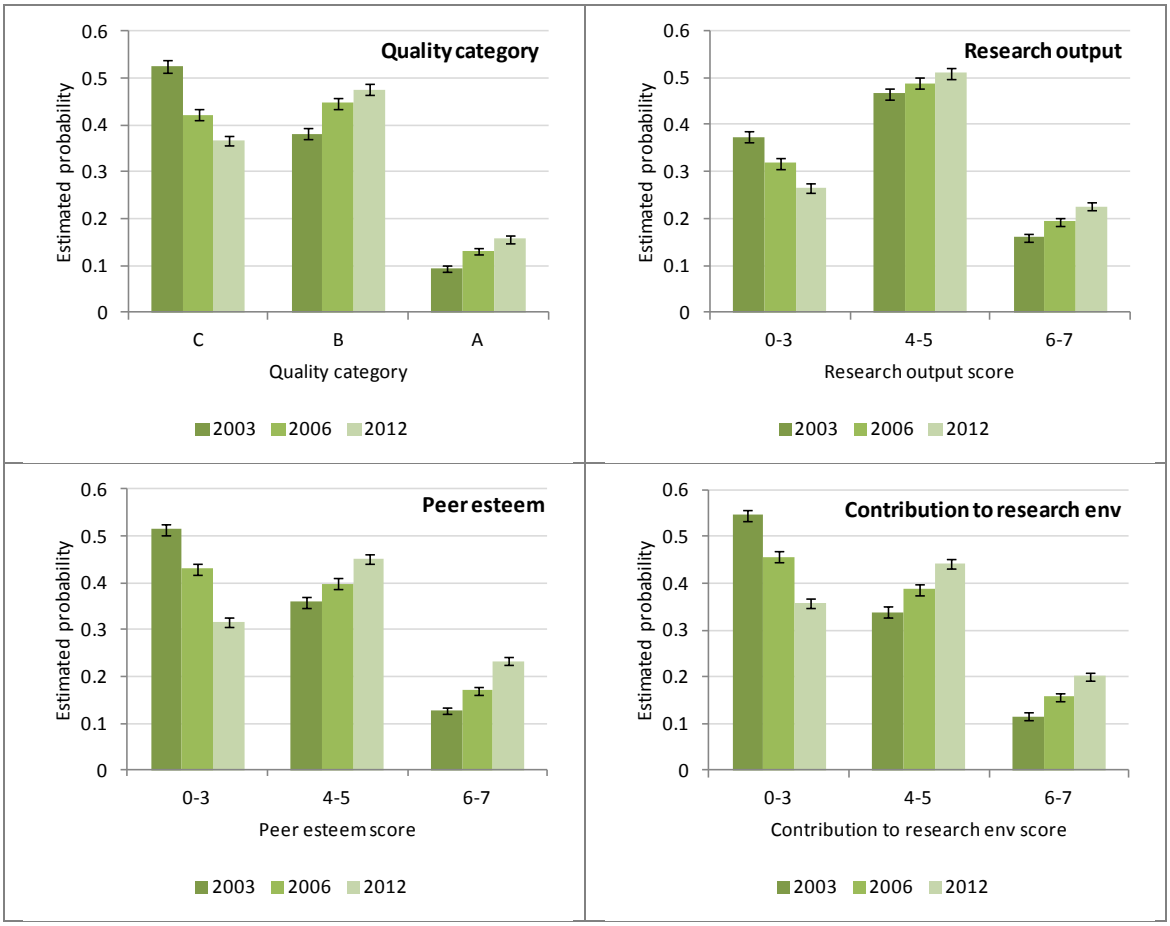
Measure	Category	2003	2006	2012
Quality category	A	0.09 (±0.01)	0.13 (±0.01)	0.16 (±0.01)
	B	0.38 (±0.01)	0.45 (±0.01)	0.48 (±0.01)
	C	0.53 (±0.01)	0.42 (±0.01)	0.37 (±0.01)
Research output	6-7	0.16 (±0.01)	0.19 (±0.01)	0.23 (±0.01)
	4-5	0.47 (±0.01)	0.49 (±0.01)	0.51 (±0.01)
	0-3	0.37 (±0.01)	0.32 (±0.01)	0.26 (±0.01)
Peer esteem	6-7	0.13 (±0.01)	0.17 (±0.01)	0.23 (±0.01)
	4-5	0.36 (±0.01)	0.40 (±0.01)	0.45 (±0.01)
	0-3	0.51 (±0.01)	0.43 (±0.01)	0.32 (±0.01)
Contribution to the	6-7	0.12 (±0.01)	0.16 (±0.01)	0.20 (±0.01)
research environment	4-5	0.34 (±0.01)	0.39 (±0.01)	0.44 (±0.01)
	0-3	0.55 (±0.01)	0.46 (±0.01)	0.36 (±0.01)

Notes: 1. The figures in brackets represent 90 percent confidence intervals. 2. This data is only for TEOs that participated in all three PBRF Quality Evaluations and for researchers who received at least a 'C' quality category.

The only case where there is no evidence of a statistically significant improvement between 2006 and 2012 is for the likelihood of a researcher receiving a score of 4-5 for their research output. This likelihood also showed the smallest change between 2003 and 2012, when compared with changes in all other categories. This result is also seen in Figure 4, which shows the average research output *scores* over time. It is clear there was little change in the research output score between 2006 and 2012, and the least change between 2003 and 2012, when compared with the scores for the other two categories.

Figure 4

Expected probabilities of receiving a particular quality or score category in 2003, 2006 and 2012



Changes in the quality of research between 2006 and 2012

The analysis in the previous section excludes evidence portfolios where staff had received a 'C(NE)'. In this section, we relax that assumption to include evidence portfolios that received at least a 'C(NE)' and widen the number of TEOs in the analysis to those that participated in 2006 and 2012. The data included in this analysis is presented in Table 9.

Table 9

Data used in the analysis of trends in research quality between 2006 and 2012

Quality categories used in this analysis	A, B, C and C(NE)
TEOs used in this analysis	Those that participated in both the 2006 and 2012 Quality Evaluations (see Table 2 for the full list)

Results

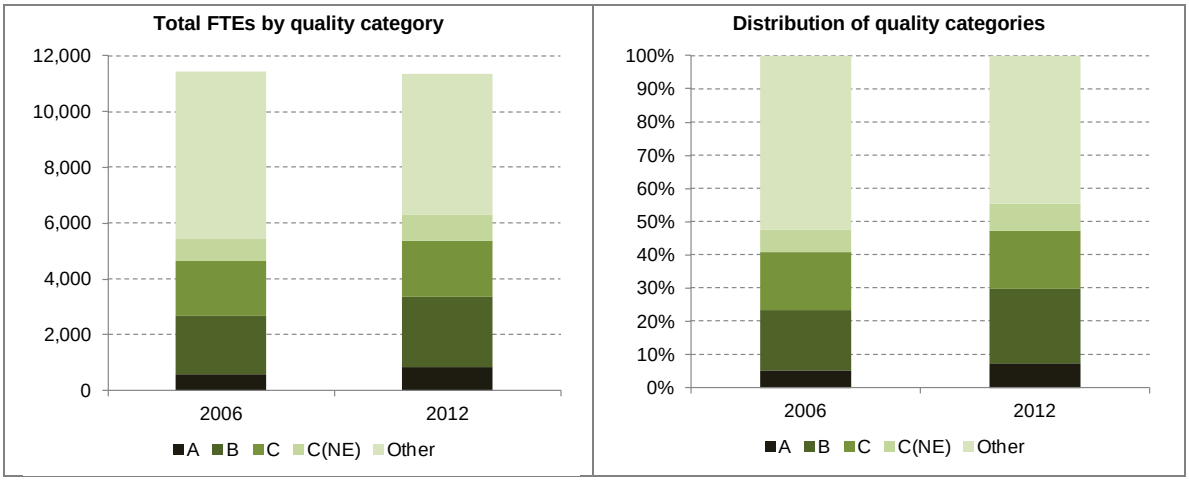
Once again, we used a variety of denominators to assess changes in the average quality of research between 2006 and 2012. Figure 5 shows the distribution of quality categories using **total FTE academic and research only staff to define our study population**. This shows that:

- the proportion of staff receiving at least a 'C(NE)' quality category increased from 48 percent in 2006 to 55 percent in 2012
- the proportion of staff who received an 'A' quality category increased from 5 percent to 7 percent

- the proportion of staff who received a ‘B’ quality category increased from 18 percent to 22 percent
- the proportion of staff who received a ‘C’ quality category increased from 17 percent to 18 percent
- the proportion of staff who received a ‘C(NE)’ quality category increased from 7 percent to 8 percent.

Figure 5

Distribution of quality categories using total academic and research only FTEs to define the staff population 2006-2012



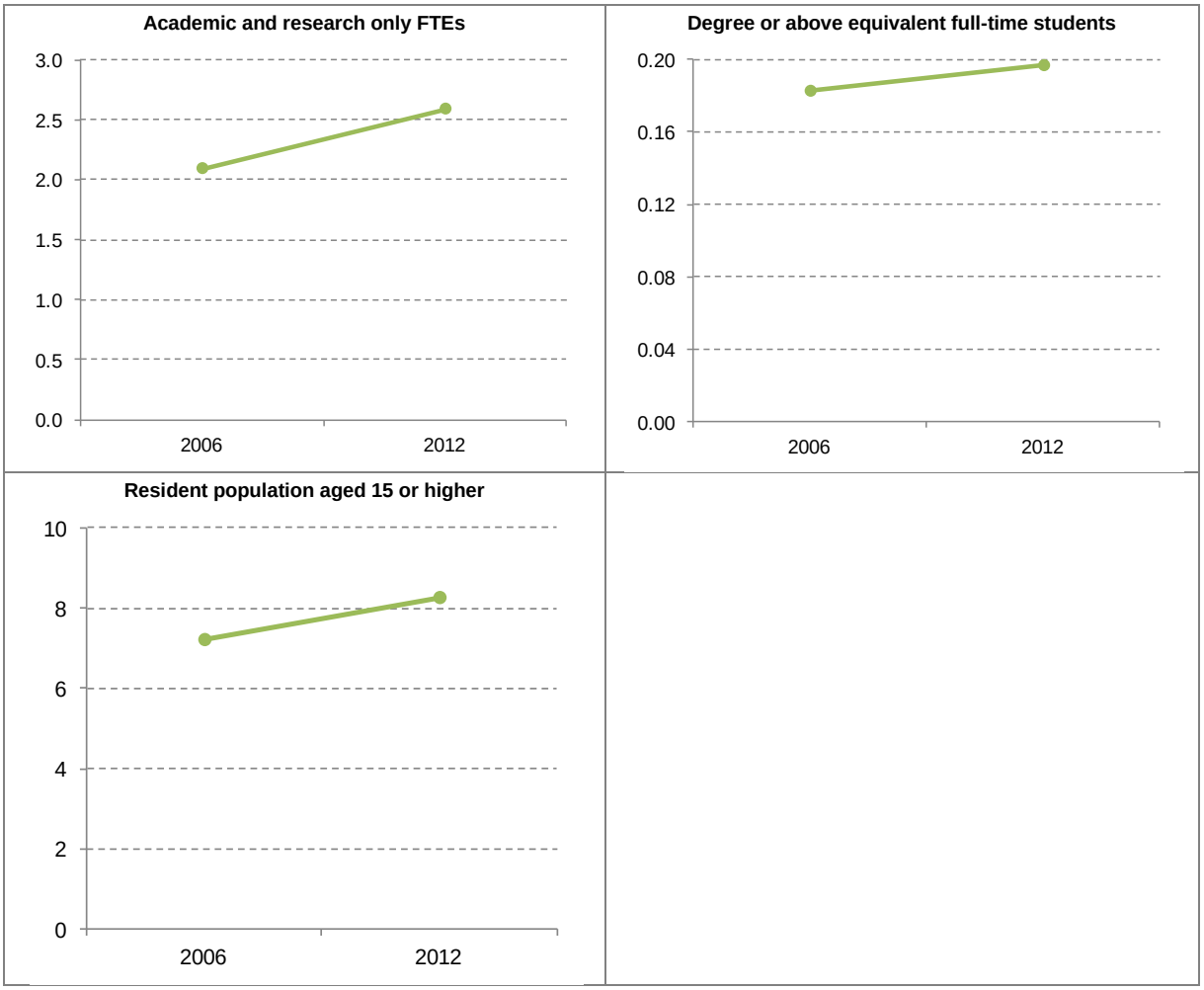
In Table 10 and Figure 6, we present the average overall quality score using the three denominators we used for the 2003-2012 analysis. In each case, the average score has increased between 2006 and 2012. So we find similar results between 2006 and 2012 to those in the previous section analysing changes in the average quality of research between 2003 and 2012. The underlying patterns appear to be the same – the average quality of research increased between 2006 and 2012, and was notable for a reduction in the proportion of the study population achieving less than a ‘C(NE)’ quality category.

Table 10

Average overall score calculated using various denominators 2006-2012

Denominator	2006	2012
Academic and research only FTEs	2.1	2.6
EFTS (degree and above)	0.18	0.20
Population 15 and over (in 000s)	7.2	8.3

Figure 6
Average quality scores using various denominators 2006-2012



We now limit our study population to those staff who received at least a 'C(NE)' quality category. In limiting the analysis in this way, we are examining the average quality of research for the group of people who are assessed as having conducted research at a level that attracts funding (under current rules). Figure 7 and Table 11 show trends in research quality between 2006 and 2012. The results show there has been an increase in the assessed quality of research. Specifically, the data shows that:

- on an FTEs-weighted basis, the number of staff receiving at least a 'C(NE)' quality category increased by 16 percent between 2006 and 2012
- the proportion of FTE-weighted 'A's increased from 11 percent in 2006 to 13 percent in 2012. Similarly, the proportion of 'B's increased from 38 percent in 2006 to 40 percent in 2012 and the proportion of 'C(NE)'s from 14 percent to 15 percent. The proportion of 'C's fell from 37 percent in 2006 to 32 percent in 2012
- when we examine the underlying component scores, improvements in the peer esteem and contribution to the research environment scores were once again important factors in the overall improvement in quality.

Table 11

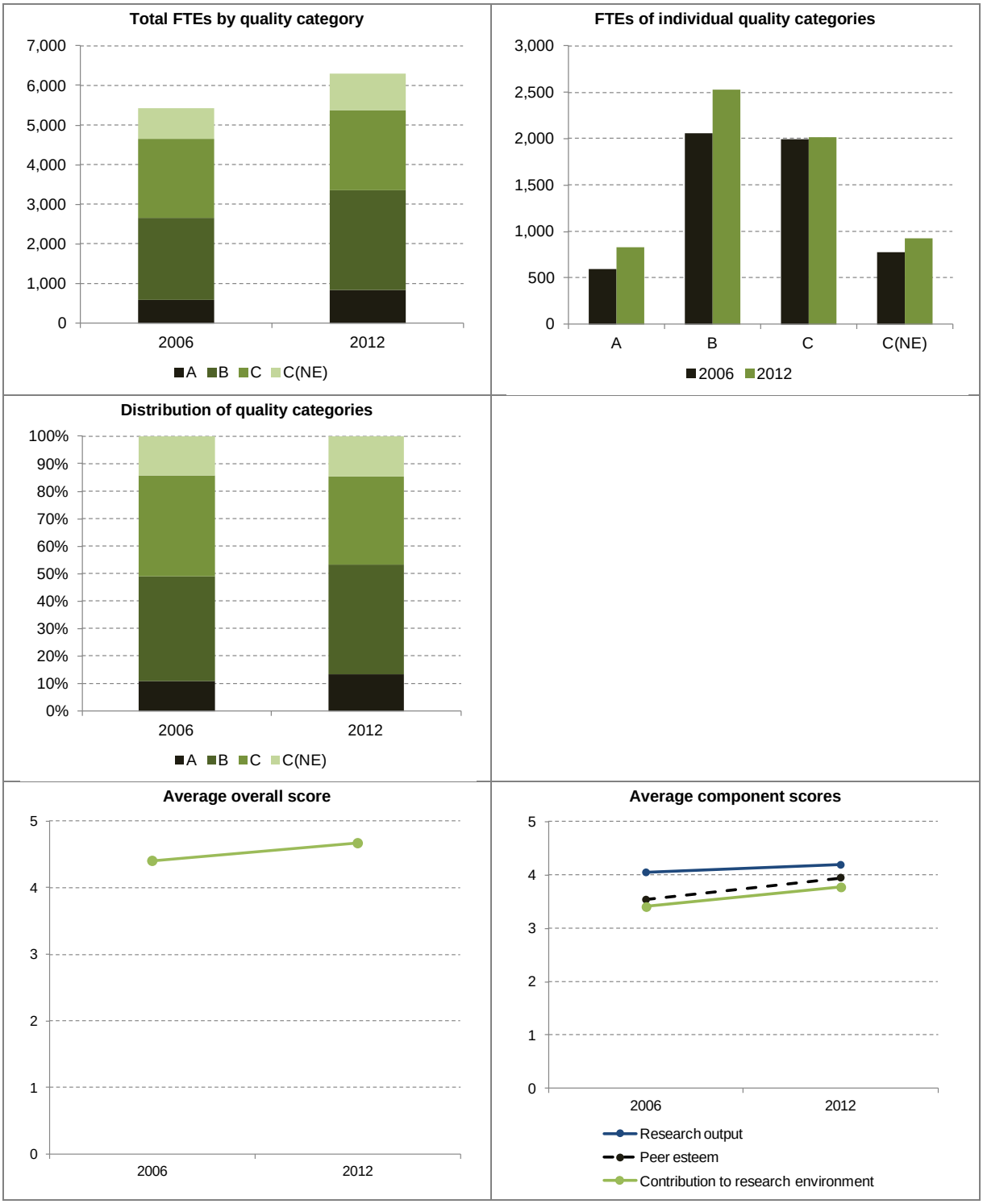
Changes in the assessed quality of research 2006-2012 (FTE weighted)

Measure	Category	2006	2012
Quality category	A	11%	13%
	B	38%	40%
	C	37%	32%
	C(NE)	14%	15%
	Total	**100%**	**100%**
Average overall score		4.4	4.7
Average research output score		4.0	4.2
Average peer esteem score		3.5	3.9
Average contribution to the research environment score		3.4	3.8
Total FTE		**5,433**	**6,294**

Note: This data is only for TEOs that participated in the 2006 and 2012 PBRF Quality Evaluations and for those staff who received at least a 'C(NE)' quality category.

Figure 7

Changes in the assessed quality of research 2006-2012 (FTE weighted)



Statistical modelling results

The statistical modelling in this section is based on a study population where we include staff who received at least a 'C(NE)' quality category. Table 12 shows the results of the statistical modelling on the 2006-2012 study population. These same results are shown in Figure 8. Again, the results for the quality category closely mirror those shown in Table 11, which are weighted by FTE. The close similarity of the results indicates that most researchers were working full-time.

Each measure of research quality shows much the same pattern between 2006 and 2012, with a statistically significant decline in the likelihood of a researcher receiving the lowest category or score, and a statistically significant increase in the likelihood of receiving a middle or higher category or score.

Generally, the likelihoods of receiving a lower score or category were higher in the 2006 and 2012 study population compared with the 2003-2012 population. Similarly, the likelihoods of receiving a middle or higher category or score were lower in the 2006-2012 study population. This is because the 2006-2012 study population included new and emerging researchers who received a 'C' category. This necessarily increases the likelihood of receiving a 'C' category, which affects the likelihood of receiving the other quality categories.

Table 12

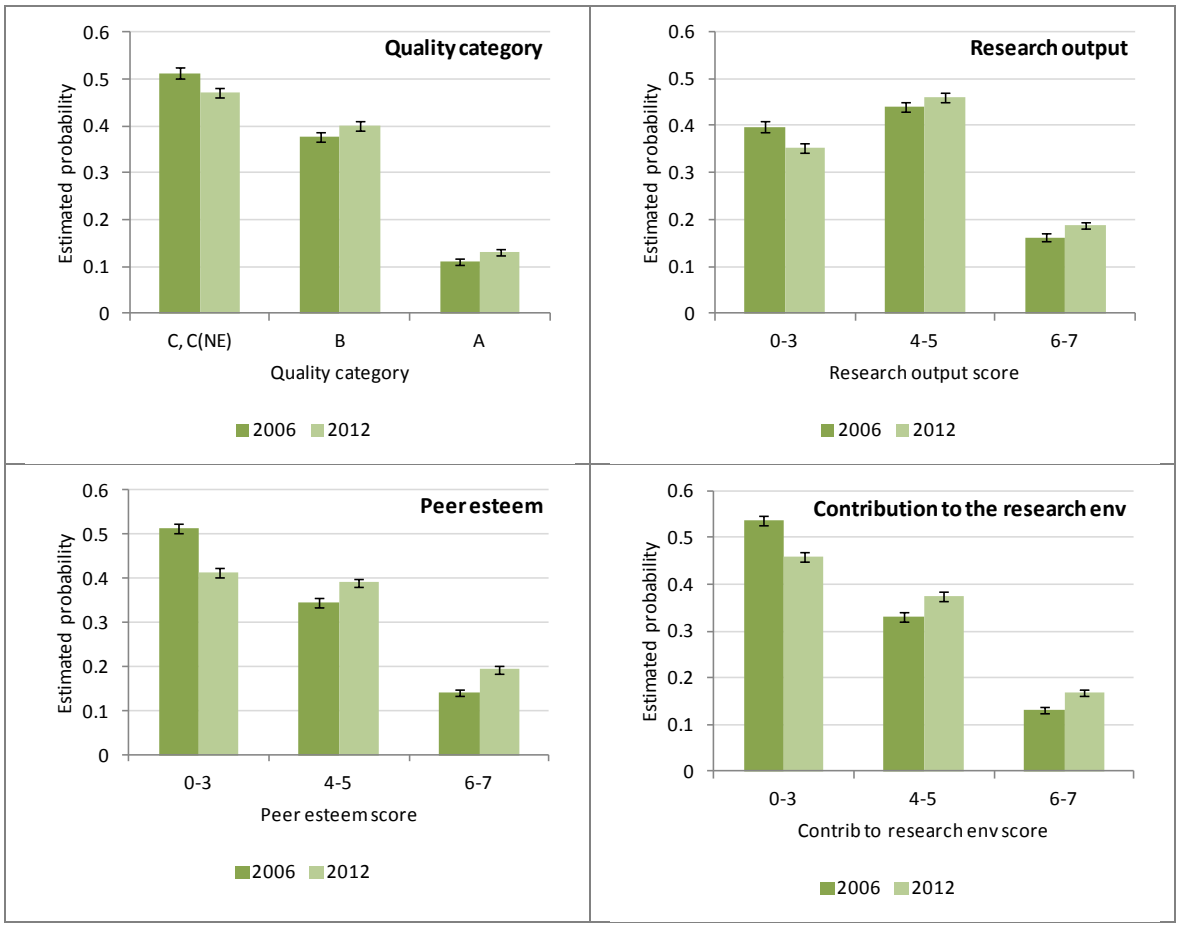
Expected probabilities of receiving a particular quality or score category in 2006 and 2012

Measure	Category	2006	2012
Quality category	A	0.11 (±0.01)	0.13 (±0.01)
	B	0.38 (±0.01)	0.40(±0.01)
	C or C(NE)	0.51 (±0.01)	0.47 (±0.01)
Research output	6-7	0.16 (±0.01)	0.19 (±0.01)
	4-5	0.44 (±0.01)	0.46 (±0.01)
	0-3	0.40 (±0.01)	0.35 (±0.01)
Peer esteem	6-7	0.14 (±0.01)	0.19 (±0.01)
	4-5	0.35 (±0.01)	0.39 (±0.01)
	0-3	0.51 (±0.01)	0.41 (±0.01)
Contribution to the	6-7	0.13 (±0.01)	0.17 (±0.01)
research environment	4-5	0.33 (±0.01)	0.37 (±0.01)
	0-3	0.54 (±0.01)	0.46 (±0.01)

Notes: 1. The figures in brackets represent 90 percent confidence intervals. 2. This data is only for TEOs that participated in the 2006 and 2012 PBRF Quality Evaluations and for researchers who received at least a 'C(NE)' quality category.

Figure 8

Expected probabilities of receiving a particular quality or score category in 2006 and 2012



The other notable result is that the changes in the likelihoods between 2006 and 2012 for the quality category and the research output scores were much lower than those for peer esteem and contributions to the research environment. This suggests that the primary drivers for the increase in research quality between 2006 and 2012 came from increases in researchers' performance on the peer esteem and contributions to the research environment categories.

5 THE FACTORS ASSOCIATED WITH HIGHER RESEARCH QUALITY

KEY POINTS

- Generally, research quality is associated with greater experience, either through being older or by holding a more senior position.
- Because women are less likely than men to hold senior positions, it may *appear* that women have lower research quality.
- However, when considering research quality between males and females *within the same position*, we find there are no differences in the quality categories received by male and female researchers.
- The only exception to this general finding is for a researcher's contribution to the research environment. Here, female professors and associate professors are more likely to receive a higher score than their male counterparts. On the other hand, for lecturers and senior lecturers, there are again no significant differences between males and females in the likelihood of receiving a higher score on this quality measure.
- We found no significant differences in the likelihood of receiving a particular quality category between researchers who submitted their evidence portfolio to the Expert Advisory Groups.
- While there are differences in the likelihoods of receiving a particular quality category between researchers who participated in just one PBRF round versus two rounds, the fact that we do not have information about all PBRF-eligible staff makes it difficult to comment meaningfully about these differences.

For the results in this section, the 2012 researcher study population was used. This study population excluded new and emerging researchers because there were too few of them to model robustly. In addition, researchers were limited to lecturers and senior lecturers, and associate professors and professors, again because there were few researchers in other positions. However, the research quality categories do include the 'R' quality category.

Table 13 gives the proportions of researchers in the 2012 study population against the research quality and other quality variables. The proportions of researchers in the 'A' and 'B' categories, or receiving higher scores, is higher than for the other two study populations because we limited our 2012 study population to particular, mostly senior, positions, and excluded new and emerging researchers who received a 'C' quality category.

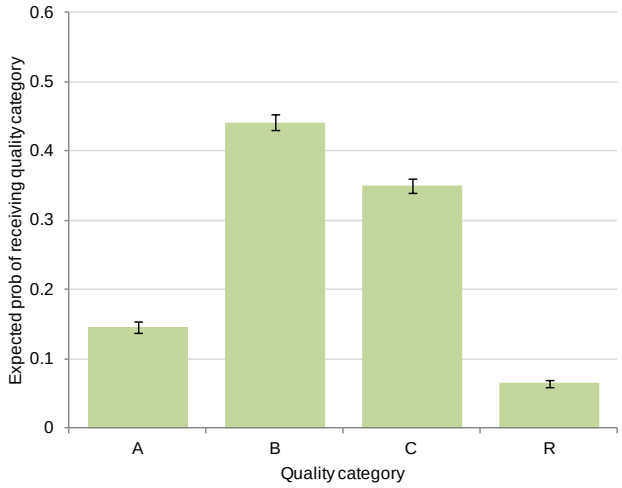
Table 13
Proportions of researchers in each of the quality categories for the 2012 study population

Measure	Categories			
	A	**B**	**C**	**R**
Quality category	0.14	0.44	0.35	0.07
	6-7	**4-5**	**0-3**	
Research output score	0.21	0.48	0.32	
Peer esteem score	0.21	0.42	0.36	
Contribution to the research environment score	0.19	0.41	0.40	

In this section, we mainly provide results for the quality category. In most instances, the results for research output, peer esteem and contributions to the research environment mirrored those for the overall quality category. We do report on the one difference we found for contribution to the research environment score.

The overall expected likelihood of a researcher receiving a particular quality category is shown in Figure 9. The error bars in the figure are 90 percent confidence limits. If the confidence limits don't overlap, then the reader can be quite certain that the categories being compared are different.

Figure 9
Expected probability of receiving a particular quality category in 2012



In this section, we are interested in finding out if this pattern changes when controlling for age, gender, position title and research subject area (using the PBRF assessment panel). We also consider whether there are differences in research quality if the research output was considered by an Expert Advisory Group. Finally, we consider differences between researchers who participated in just one PBRF round versus those who participated in both the 2006 and 2012 rounds.

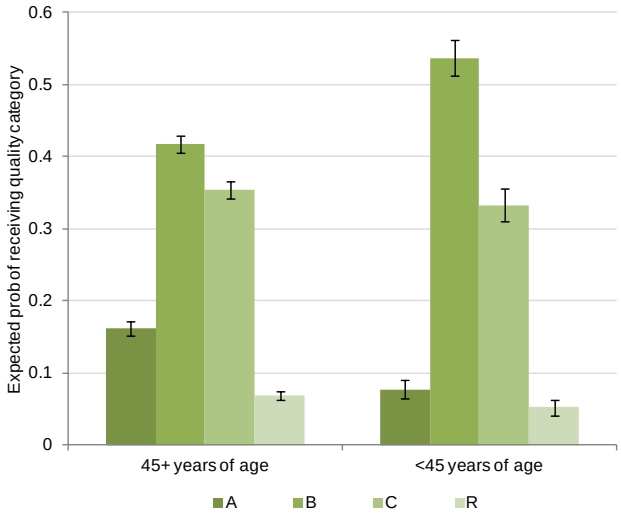
Age of researcher

In the 2012 study population, 80 percent of the researchers were 45 years and older in 2012.

Figure 10 shows the expected likelihoods of receiving a particular quality category by age. It is clear that older researchers were more likely to receive an ‘A’ quality category, although receiving a ‘B’ category was still the most likely. Younger researchers are much more likely to receive a ‘B’ category. This difference in the likelihood of receiving a quality category between these younger and older researchers is statistically significant.

Figure 10

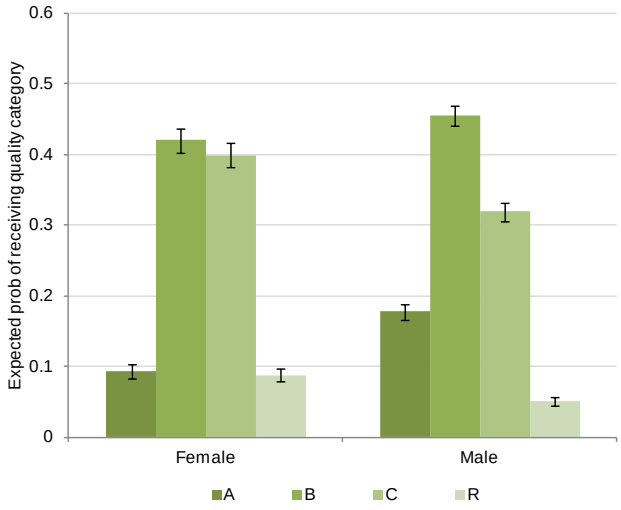
Expected probability of receiving a particular quality category by age of researcher in 2012



Gender of researcher

Almost two-thirds of researchers in the 2012 study population were male (62 percent). Figure 11 shows the expected likelihood of receiving a particular quality category by gender. The differences between gender are statistically significant, with females generally less likely to receive an ‘A’ or ‘B’ quality category, and more likely to receive a ‘C’ or ‘R’ category. However, as we see later in this chapter (Figure 14), this apparent discrepancy mainly reflects the fact that, on average, women were less likely to hold the senior positions associated with higher scores.

Figure 11
Expected probability of receiving a particular quality category by gender in 2012

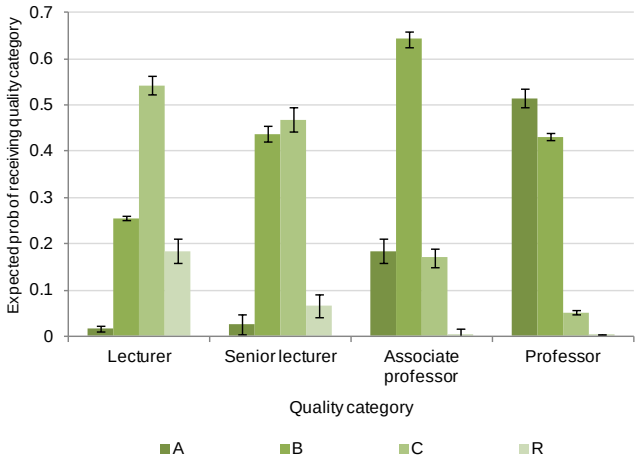


Position title

In the 2012 study population, 18 percent were lecturers, 45 percent senior lecturers, 18 percent associated professors, and 19 percent were professors.

As can be seen in Figure 12, there were clear differences in the likelihood of receiving a particular quality category by position, and these differences were statistically significant. Professors were most likely to receive an ‘A’ category, associate professors were most likely to gain a ‘B’ category, and lecturers were mostly likely to receive a ‘C’ category. Senior lecturers were about as likely to receive a ‘B’ or ‘C’ category.

Figure 12
Expected probability of receiving a particular quality category by position in 2012

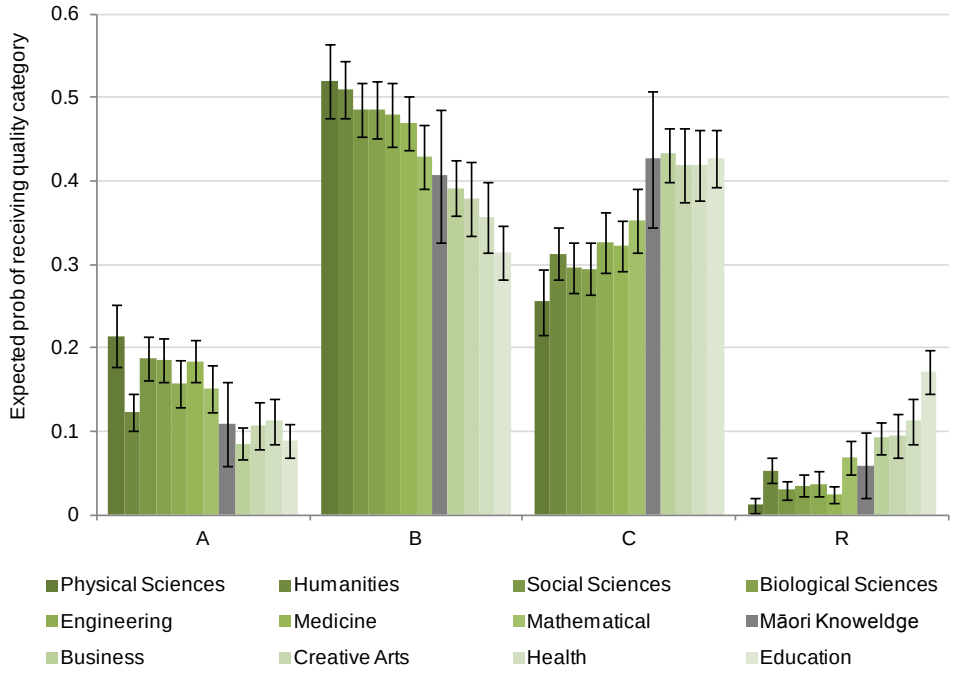


Subject area (panel)

Analysing differences between researchers in the different panels is problematic, because some panels had very few researchers in them. The study population for this analysis did not include all researchers, which exacerbated this problem, but we felt it was better to have a consistent study population for all the results in this section.

Figure 13 shows the results. The wide confidence limits indicate there are few researchers in the different panels. Our focus here was specifically on the Māori Knowledge and Development Panel. It is shown in dark grey in the figure.

Figure 13
Expected probability of receiving a particular quality category by PBRF assessment panel in 2012



Note: The results are presented in rank order of the expected probability of receiving a 'B' quality category.

The small number of researchers in each of the PBRF assessment panels means we are less certain of the results, as demonstrated by the wider confidence limits in the figure. We can't be sure there are real differences in the likelihood of receiving a particular quality category between researchers in many of the assessment panels.

However, it is clear that researchers in the physical sciences, humanities, social sciences, biological sciences, engineering and medicine are most likely to receive a 'B' category, while researchers in education are most likely to receive a 'C' category. For the other assessment panels, we can't be sure there is a difference in the likelihood of receiving a 'B' or 'C' category, as their confidence limits overlap.

In respect of the Māori Knowledge and Development Panel specifically, we cannot be confident that the likelihoods of receiving a particular quality category by researchers in this panel are significantly different from those received by researchers in a range of other assessment panels.

Position title and gender

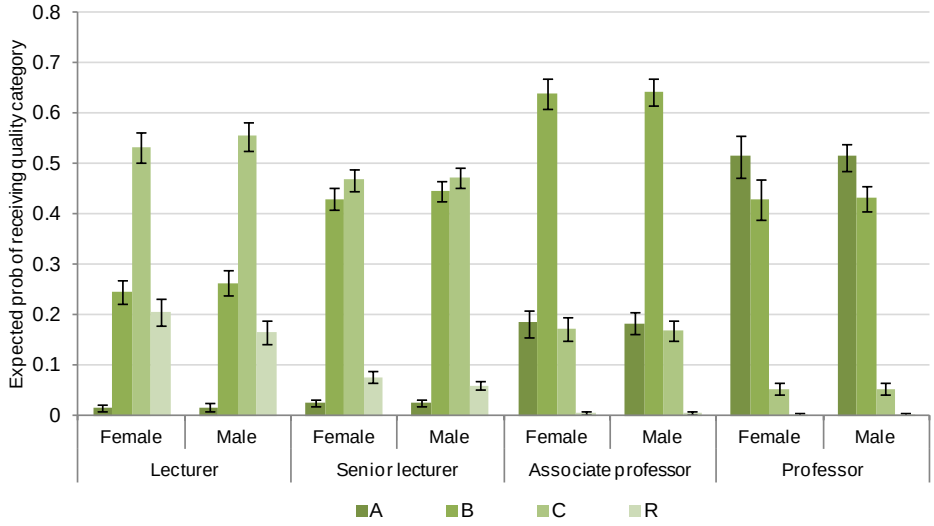
The results up to this point have been for single factors: age, gender, position and panel. In other words, these are the results when not controlling for any other factor.

The single factor results showed that females were less likely to receive an 'A' or 'B' quality category. However, female researchers were more likely to be found in positions that are less likely to receive an 'A' category. For example, while 19 percent of the 2012 study population were professors, 25 percent of males were professors, and 10 percent of females were

professors. Similarly for the other positions: 22 percent of females were lecturers, while for males the figure was 15 percent.

Figure 14 shows the expected likelihood of receiving a particular quality category when controlling for position title and gender. In this case, there is no statistical difference between males and females when controlling for position. This finding is consistent with the analyses by Smart (2005, 2008a).

Figure 14
Expected probability of receiving a particular quality category by position and gender in 2012



When we considered differences between age groups controlling for gender, we found there were no significant interaction between these two factors. That is to say, the differences in the likelihood of receiving a particular quality category were consistent between males and females regardless of age, and vice versa. This is because there is a much more even distribution of genders by age group. In our study population, 22 percent of females were under 45 years of age, compared with 19 percent of males.

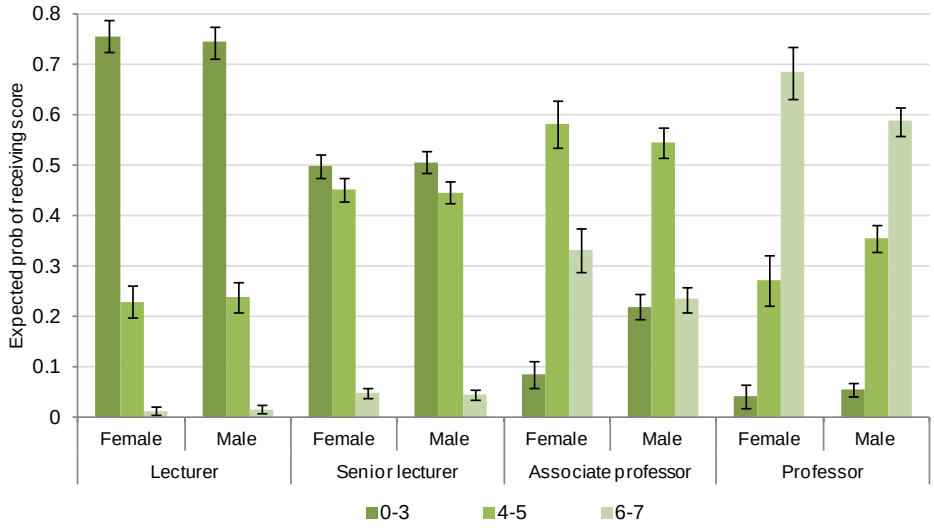
Consistency of results across the quality variables

These results described for age group, gender, position title and panel as applying to the quality category apply also to the research output score categories, as well as to the peer esteem and contributions to the research environment score categories.

There is only one exception. In general, there is no statistically significant difference in the likelihood of males and females receiving a particular quality category or score when controlling for position title. However, this is not the case for the contribution to the research environment score. In this case, there is still no statistically significant differences between genders overall, but females were more likely to receive a higher score than males for professors and associated professors. For lecturers and senior lecturers, there was again no difference between genders. This result is shown in Figure 15.

Figure 15

Expected probability of receiving a particular contribution to the research environment score by position and gender in 2012

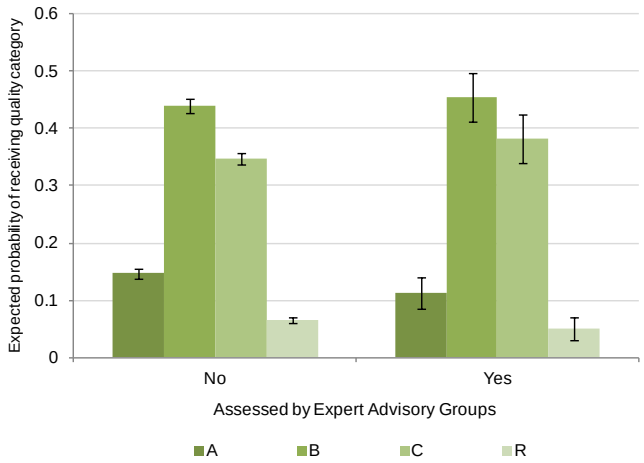


Assessment by Professional and Applied and Pacific Research Expert Advisory Groups

We were interested to see if researchers who had their evidence portfolios assessed by the two Expert Advisory Groups had different likelihoods of receiving particular quality categories. Figure 16 shows the results.

Figure 16

Expected probability of receiving a particular quality category by whether or not the evidence portfolio was assessed by the Expert Advisory Groups in 2012



In our 2012 study population, 356 evidence portfolios were assessed by the Expert Advisory Groups, out of the 5,655 in total. The relatively wide confidence limits for the Expert Advisory Group-assessed portfolios in Figure 16 reflect this low number. It is not surprising therefore to find that there is no statistical difference in the likelihoods of receiving a particular quality category between these two groups, even though the expected likelihood of receiving an ‘A’ category is slightly lower for Expert Advisory Group-assessed researchers. However, the differences in the likelihoods between the two groups are not large in absolute terms, with both groups most likely to receive a ‘B’ category.

We also checked to see if there was any interaction between the Expert Advisory Group-assessed researchers and gender, age, position and panel. For the last two factors, there were too few commercialisation-assessed researchers for the regression models to be stable, so we were unable to test these factors. However, when controlling for gender or age (but not both together), statistically significant differences were found between male and female researchers, and younger and older researchers, in line with the results we have discussed previously. There was also no interaction between these factors and whether the portfolio was assessed by the Expert Advisory Groups. That is to say, there were no differences between the genders, or between age groups, in the likelihood of receiving a particular quality category within the Expert Advisory Group-assessed researcher group or within the group of other researchers. Finally, controlling for gender or age, there was no statistical difference in the likelihood of a researcher receiving a particular quality category between the researchers who had their portfolios assessed by the Expert Advisory Groups or those who did not.

Comparing researchers in single and multiple PBRF rounds

We were interested to see determining whether changes in research quality arise from existing researchers improving their research quality, or if the improvements are driven by an influx of better researchers into the tertiary education researcher workforce, or both.

In this section, we consider the likelihoods of receiving a particular quality category for researchers who participated in the 2006 PBRF round, but not the 2012 round, against those researchers who did not participate in the 2006 round, but were in the 2012 round. We ignored the 2003 round in determining the status of a particular researcher.

We first consider researchers in a single PBRF round, and then we consider those researchers who were in *both* the 2006 and 2012 PBRF rounds.

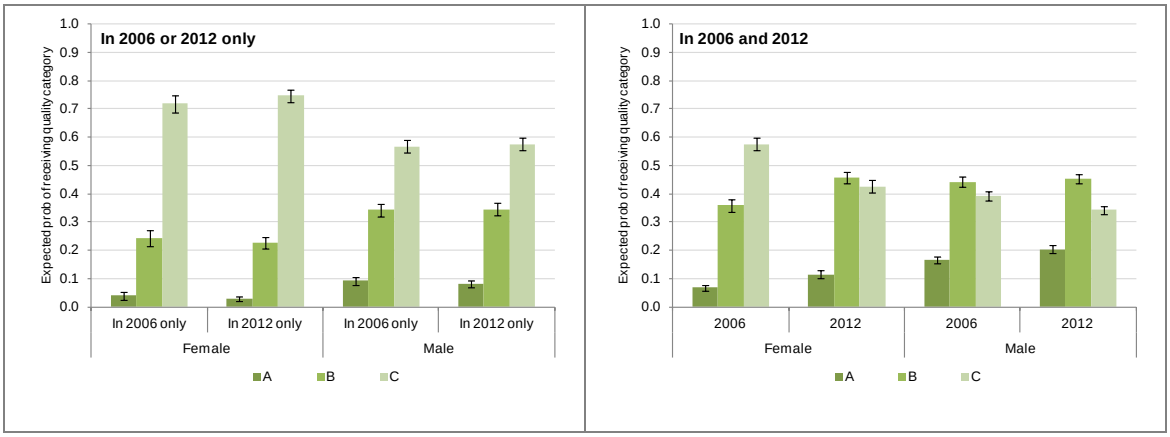
For these analyses, the study population excludes researchers with missing birth dates and gender, and those who received an 'R' quality category. In contrast to the previous sections, we included researchers from all position titles.

In addition to considering whether a researcher participated in a PBRF round or not, we also controlled for gender. As we have seen, there were significant differences between male and female researchers, so controlling for gender provides us with a better understanding of the effect of PBRF round participation.

Figure 17 (left panel) shows the results for researchers who participated in either the 2006 or 2012 PBRF rounds.

Figure 17

Expected probability of receiving a particular quality category by participation status in 2006 and 2012, and by gender



The results show that when controlling for gender there is no difference between those who participated only in 2006 and those who participated only in 2012. While males, as we have seen, were more likely to receive higher quality categories, both male and female researchers who participated in just one PBRF round were most likely to receive a 'C' quality category.

Figure 17 (right panel) shows the results for researchers participating in both the 2006 and 2012 rounds.

The improvement in the likelihood of receiving a higher quality category is clear, especially for female researchers. This reinforces our general finding that higher quality categories are received by more experienced researchers.

When comparing researchers who participated in just one round with those who participated in both the 2006 and 2012 rounds, there are clear differences in the likelihood of receiving a particular quality category.

The greatest difference was that, while researchers in just one round, as a whole, were most likely to receive a 'C' quality category, only females in 2006 who continued on to the 2012 round were most likely to receive this quality category. For females in 2012 and male researchers generally who participated in both rounds, the likelihood of receiving a 'C' category were about the same or less than receiving a 'B' quality category.

Overall, the results show that researchers in 2006 who did not continue on to 2012 were more likely to have received a lower quality category. And the new researchers in 2012 were more likely to receive a lower quality category than those who were also present in the 2006 round.

The average age of researchers who only participated in 2006 was no different from those who were present in both rounds. But the researchers who participated only in the 2012 round were significantly younger than the average researcher by about 11 years. It might be expected that new researchers will be younger and therefore less experienced. This probably explains why these new researchers in 2012 were more likely to receive a lower quality category.

6 RESEARCHER CAREER PATHWAYS

KEY POINTS

We examined the pathways of staff that participated in the 2012 Quality Evaluation and found that:

- the higher the level of quality category awarded in 2012, the more likely the staff were to have participated in the 2006 Quality Evaluation
- there is little evidence that the proportion of people who had not participated in the previous Quality Evaluation has changed over time, once we discount for the different time period between Quality Evaluations.

We also examined what proportion of staff moved TEO between the 2006 and 2012 Quality Evaluations. The data showed that:

- only around 6 percent of staff moved to another TEO, a level that was broadly consistent across levels of quality category. So there is little evidence that researchers who received higher quality categories were more likely to move between TEOs
- there is little sign that the rates of movement have changed over time, once we discount for the different time period between Quality Evaluations.

Introduction

In this section, we examine the pathways of staff who received an 'A', 'B', 'C', or 'R' quality category in the 2012 Quality Evaluation.[8] We look to see if they participated in the 2006 Quality Evaluation and, if they did, what their quality category was. We then compare these proportions with pathways of staff who participated in earlier Quality Evaluations.

We also examine the percentage of staff who moved TEOs between Quality Evaluations. We compare rates of moving between the 2003 and 2006, and 2006 and 2012 Quality Evaluations.

For trend analysis purposes, we restrict our analysis to staff at TEOs that participated in all three Quality Evaluations.

Results

The pathways of staff who received an 'A', 'B', 'C' or 'R' quality category in the 2012 Quality Evaluation are presented in Table 14. The results show that:

- for staff that received an 'A' in 2012, their most likely pathway was to have received an 'A' in the 2006 Quality Evaluation (41 percent). The second most common pathway was to have received a 'B' in the 2006 Quality Evaluation
- for staff that received a 'B' in 2012, their most likely pathway was to have received a 'B' in the 2006 Quality Evaluation (37 percent). The second most likely pathway was not to have featured in the 2006 Quality Evaluation (28 percent)

[8] We omit staff who received a 'C(NE)' or 'R(NE)' quality category.

- for staff that received a ‘C’ in 2012, the most common pathway was to have received a ‘C’ in the 2006 Quality Evaluation (31 percent). The second most common pathway was not to have participated in the 2006 Quality Evaluation (28 percent)

- for staff that received an ‘R’ in 2012, the most common pathway was not to have participated in the 2006 Quality Evaluation (34 percent), followed by receiving an ‘R’ in the 2006 Quality Evaluation (32 percent).

Overall, the data shows that the higher the quality category achieved by staff in 2012, the more likely they were to have participated in the 2006 Quality Evaluation.

Table 14

Pathways of staff who participated in the 2012 Quality Evaluation

Quality category in 2012	Quality category in 2006	N	%
A	A	358	41%
	B	317	36%
	C	30	3%
	C(NE), R, R(NE)	15	2%
	n/a	157	18%
	Total	**877**	**100%**
B	A	124	5%
	B	978	37%
	C	515	19%
	C(NE)	214	8%
	R	45	2%
	R(NE)	36	1%
	n/a	745	28%
	Total	**2,657**	**100%**
C	A, B	244	12%
	C	638	31%
	C(NE)	182	9%
	R	251	12%
	R(NE)	157	8%
	n/a	574	28%
	Total	**2,046**	**100%**
R	A, B, C	49	16%
	C(NE)	12	4%
	R	100	32%
	R(NE)	47	15%
	n/a	105	34%
	Total	**313**	**100%**
Total		***5,893***	

Note: Categories have been aggregated where necessary so as to ensure all entries in this table have five or more researchers.

Because the data used in Table 14 is a snapshot, it is difficult to place the results in context. It is useful to compare the pathways of staff in earlier Quality Evaluations. We can compare the pathways of staff that received an ‘A’, ‘B’ or ‘C’ quality category in the 2012 Quality Evaluation with the pathways of those who received an ‘A’, ‘B’ or ‘C’ in the 2006 Quality

Evaluation (see Tables 15 and 16). We restrict our attention to these three quality categories, due to the change in rules that introduced the new and emerging category, as well as the non-reporting of PBRF-eligible staff in 2012 which causes bias in the lower quality categories.

When we compare the results in Table 15 with Table 16, the data shows that:

- in 2012, in each quality category, there was a higher rate of staff not participating in the previous Quality Evaluation. But this disparity is in line with what we would expect, given the differences in years between Quality Evaluations. For example, the proportion of staff who received an 'A' in 2012 who were not in the previous Quality Evaluation was 18 percent, compared with 9 percent in 2006. Taking into account that the number of years between Quality Evaluations was doubled between 2006 and 2012, this indicates the rate has remained relatively constant over time

- a greater proportion of staff in the 2012 Quality Evaluation experienced a fall in quality category, compared with staff in the 2006 Quality Evaluation. For example, 5 percent of staff that received a 'B' in 2012 had received an A in 2006, compared with no-one in 2006. Given that 2006 was a partial round where staff could carry over quality categories, it is unlikely that staff would have resubmitted if they suspected they would have received a lower quality category. In particular, those who received an 'A' in 2003 had no incentive to participate in 2006. Indeed, there was a strong disincentive for those staff to participate in the 2006 Quality Evaluation.

Table 15
Pathways of staff who participated in the 2012 Quality Evaluation and received an 'A', 'B' or 'C' quality category

Quality category in 2012	Quality category in 2006	N	%
A	A	358	41%
	Other	362	41%
	n/a	157	18%
	Total	**877**	**100%**
B	A	124	5%
	B	978	37%
	C	515	19%
	Other	295	11%
	n/a	745	28%
	Total	**2,657**	**100%**
C	A, B	244	12%
	C	638	31%
	Other	590	29%
	n/a	574	28%
	Total	**2,046**	**100%**

Table 16
Pathways of staff who participated in the 2006 Quality Evaluation and received an 'A', 'B' or 'C' quality category

Quality category in 2006	Quality category in 2003	N	%
A	A	411	65%
	Other	158	25%
	n/a	59	9%
	Total	**628**	**100%**
B	A	0	0%
	B	1,391	65%
	C	430	20%
	Other	14	0.7%
	n/a	311	14%
	Total	**2,153**	**100%**
C	A, B	30	1%
	C	1,367	67%
	Other	276	14%
	n/a	352	17%
	Total	**2,038**	**100%**

Moving between TEOs

In Table 17, we present the proportion of people who participated in both the 2006 and 2012 Quality Evaluations and moved between TEOs. For comparison purposes with earlier years, we limit the set of TEOs to those who participated in all three Quality Evaluations. We also treat staff at colleges of education as being at the universities that absorbed them, so movements as a result of mergers are ignored.

The data shows that the proportion of staff who moved was relatively low, around 6 percent, and that there was little difference in proportion by quality category received in 2006. The one exception was for the 'C(NE)' quality category, with a proportion who moved of around 9 percent. So there was little evidence of a 'market' for staff who received an 'A' in the Quality Evaluations.

Table 17
People who were in both the 2006 and 2012 Quality Evaluations by whether they moved TEO or not

2006 QE result	Outcome in next QE	N	%
A	Moved TEO	30	6%
	Did not move TEO	454	94%
	Total	**484**	**100%**
B	Moved TEO	93	6%
	Did not move TEO	1,456	94%
	Total	**1,549**	**100%**
C	Moved TEO	74	6%
	Did not move TEO	1,166	94%
	Total	**1,240**	**100%**
C(NE)	Moved TEO	37	9%
	Did not move TEO	394	91%
	Total	**431**	**100%**
R	Moved TEO	27	7%
	Did not move TEO	381	93%
	Total	**408**	**100%**
R(NE)	Moved TEO	12	5%
	Did not move TEO	239	95%
	Total	**251**	**100%**
All	Moved TEO	261	6%
	Did not move TEO	4,102	94%
	Total	***4,363***	**100%**

In Table 18, we now compare the proportion of people who moved between the 2003 and 2006 Quality Evaluations, to place the results in Table 17 in context. We focus only on staff who received an ‘A’, ‘B’ or ‘C’, since the non-reporting of PBRF-eligible in 2012 caused bias in those categories that are less than a ‘C’.

The results show that the percentage of people who moved was lower. It was around half the rate of the 2006 data, which was in line with what one would expect given that the length of time between Quality Evaluations was around half. So there was no evidence that the rate of movement between TEOs had changed between those periods.

Table 18
People who were in both the 2003 and 2006 Quality Evaluations by whether they moved TEO or not

2003 QE result	Outcome in next QE	N	%
A	Moved TEO	18	4%
	Did not move TEO	393	96%
	Total	**411**	**100%**
B	Moved TEO	45	3%
	Did not move TEO	1,538	97%
	Total	**1,583**	**100%**
C	Moved TEO	53	3%
	Did not move TEO	1,872	97%
	Total	**1,925**	**100%**

7 A CLOSER LOOK AT THE DISTRIBUTION OF ‘A’S

KEY POINTS

We analysed the evidence portfolios that received an ‘A’ quality category in 2006 and 2012. We found that:

- all 12 panel areas experienced an increase in ‘A’s in terms of FTEs between 2006 and 2012
- seven of the 12 PBRF panels increased their share of ‘A’s between 2006 and 2012. Panels to increase share included Medicine and Public Health and Creative and Performing Arts. Panels to decrease share included Humanities and Law and Engineering, Technology and Architecture
- the share of FTE-weighted ‘A’ quality categories received by women increased between 2006 and 2012.

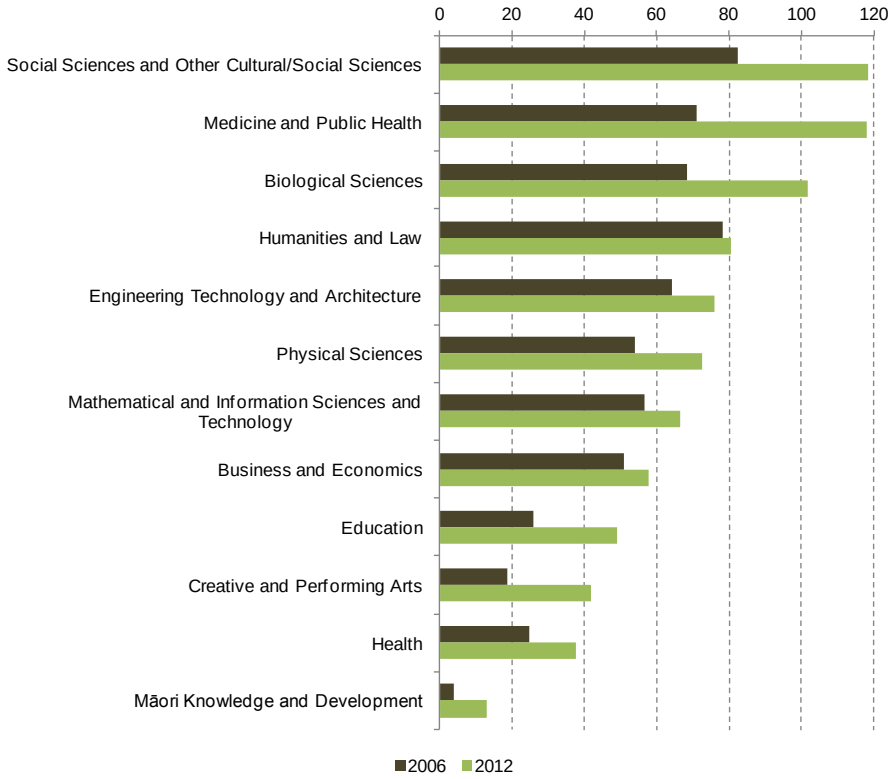
Introduction

In this section, we take a closer look at the distribution of ‘A’ quality categories in the 2006 and 2012 Quality Evaluations. In particular, we examine the share of ‘A’s among PBRF panel areas and by gender. For trend analysis purposes, we restrict our analysis to TEOs that participated in both the 2006 and 2012 Quality Evaluations.

Results

Figure 18 presents the number of full-time equivalent ‘A’s by PBRF panel area. In 2012, the panel areas with the most ‘A’s were Social Sciences and Other/Cultural Sciences and Medicine and Public Health (both with 118 FTEs). All panels exhibited an increase in ‘A’s on an FTE-weighted basis between 2006 and 2012.

Figure 18
The number of ‘A’s by PBRF panel area (FTE weighted)



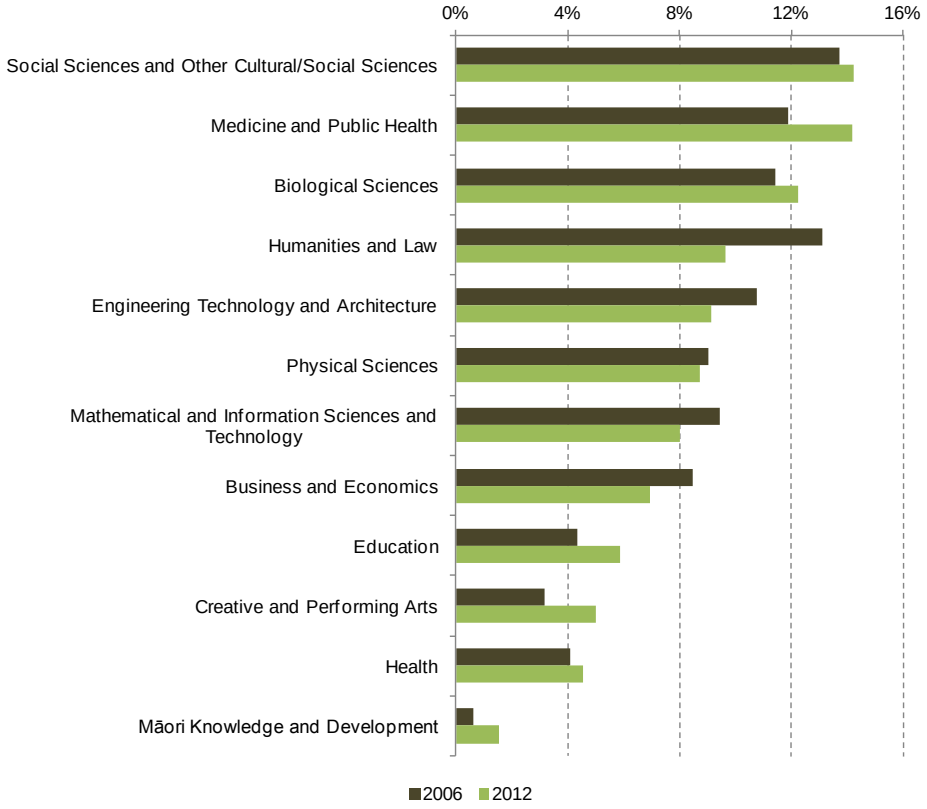
Note: This analysis is restricted to TEOs that participated in both the 2006 and 2012 Quality Evaluations.

In Figure 19 we present the number of ‘A’s as a percentage of total ‘A’s in each year. Seven panels showed an increase in their share of ‘A’s, while five panels showed a decrease in share. Among the panel areas that experienced a reduction in share of total ‘A’s were Humanities and Law (from 13 percent to 10 percent) and Engineering, Technology and Architecture (from 11 percent to 9 percent).

Among the subject areas to exhibit significant increases in share were: Medicine and Public Health (from 12 percent to 14 percent) and Creative and Performing Arts (from 3 percent to 5 percent).

Figure 19

The distribution of ‘A’s by PBRF panel area (FTE weighted)

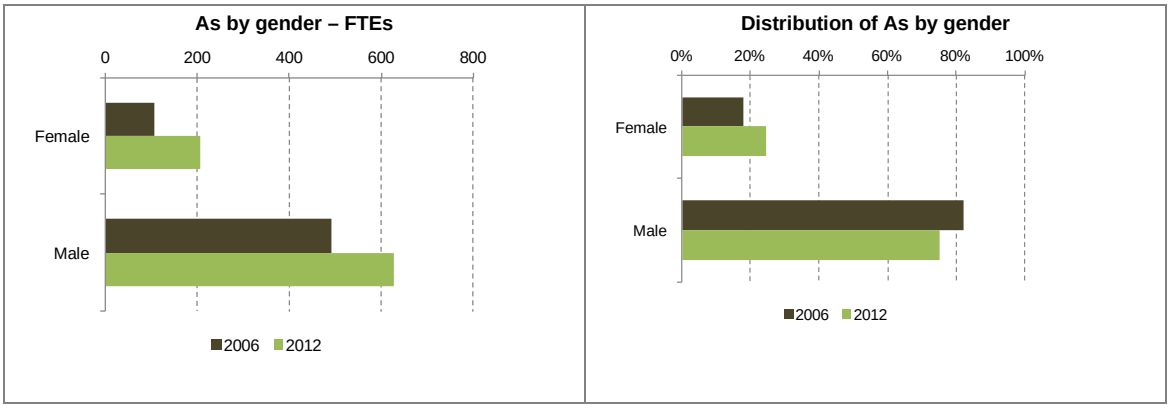


Note: This analysis is restricted to TEOs that participated in both the 2006 and 2012 Quality Evaluations.

We now examine the share of ‘A’s by gender (Figure 20). This shows that both the number (in terms of FTEs) and the proportion of ‘A’s received by women in the Quality Evaluations increased between 2006 and 2012.

Figure 20

The number and distribution of ‘A’s by gender (FTE weighted)



8 A CLOSER LOOK AT NEW AND EMERGING RESEARCHERS

KEY POINTS

We examined what happened to staff who received a 'C(NE)' or 'R(NE)' quality category in the 2006 Quality Evaluation. We found that:

- around a third of this group of researchers received at least a 'C' quality category in the 2012 Quality Evaluation, with the rest either receiving an 'R' or not participating in the evaluation.

We also profiled the distribution of new and emerging researchers across PBRF panel areas. The results showed that:

- the largest number of new and emerging researchers were located in Medicine and Public Health
- the panel with the highest proportion of new and emerging researchers was Health. The panel with the smallest proportion was Education.

Introduction

In this section, we examine the staff who were categorised as new and emerging in the 2006 Quality Evaluation and were assessed as having a 'C(NE)' or 'R(NE)' quality category. We look at whether they participated in the 2012 Quality Evaluation or not, and, if they did, what their level of performance was. We also present a profile of new and emerging researchers in 2012 by panel area.

Results

The results show that around 37 percent of staff who received a 'C(NE)' or 'R(NE)' in 2006 received at least a 'C' in the 2012 Quality Evaluation (Table 19). The remaining 63 percent of staff either received an 'R' or did not submit an evidence portfolio. As might be expected, the performance of staff who received a 'C(NE)' was better than those who received an 'R(NE)'.

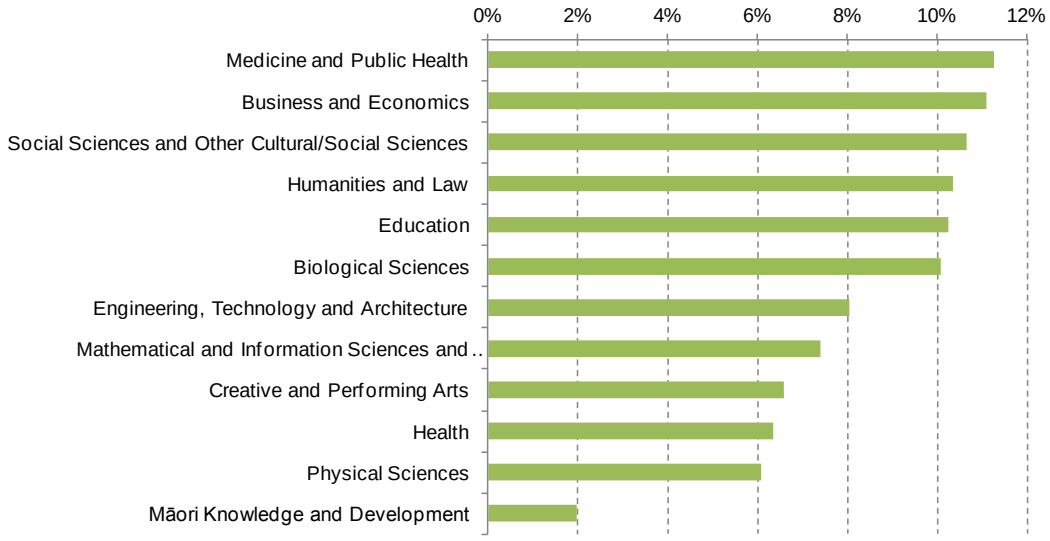
It is difficult to place this finding in context, given there is no baseline data to compare with prior to the introduction of the PBRF; however, a relatively high rate of staff turnover among early career staff is common in many occupations.

Table 19
Destinations of staff in 2012 who were assigned a 'C(NE)' or 'R(NE)' quality category in 2006

Quality category in 2006	Quality category in 2012	N	%
C(NE)	A, B	229	28%
	C	196	24%
	Other	392	48%
	Total	**817**	**100%**
R(NE)	A, B	37	4%
	C	176	19%
	Other	702	77%
	Total	**915**	**100%**
Total	A, B	266	15%
	C	372	21%
	Other	1,094	63%
	Total	**1,732**	**100%**

We now look at the distribution of new and emerging researchers in the 2012 Quality Evaluation. Here we consider all evidence portfolios submitted. Figure 21 presents the distribution of new and emerging staff across the PBRF panel areas. In total, there were 1,342 new and emerging evidence portfolios out of the 7,324 submitted. In terms of volume, the largest number of new and emerging staff were located in Medicine and Public Health and Business and Economics (both around 11 percent).

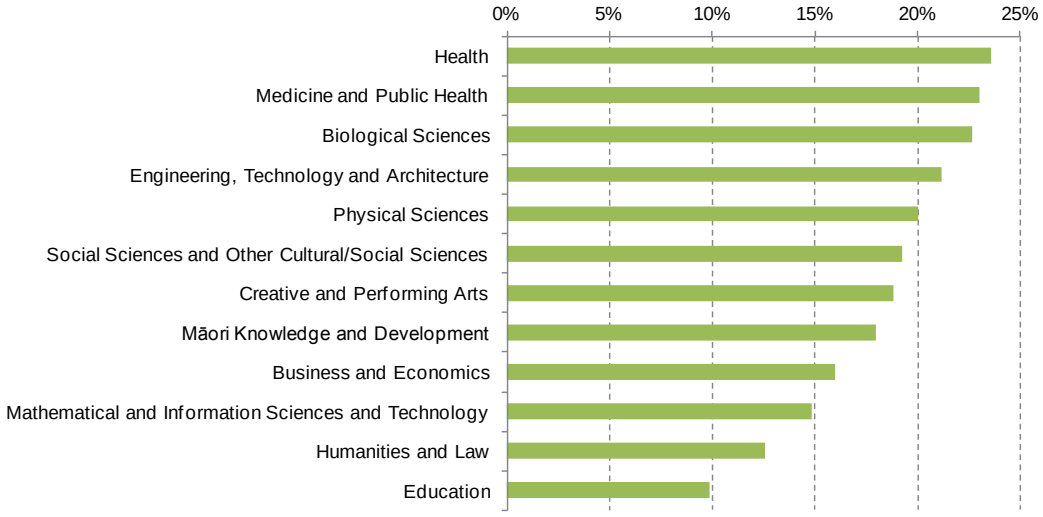
Figure 21
Distribution of new and emerging researchers in 2012



We now look at the proportion of new and emerging researchers within each panel. Figure 22 shows a high proportion of staff in the health areas were identified as new and emerging. The panel with the highest proportion of new and emerging researchers was Health (24 percent), followed by Medicine and Public Health (23 percent).

Areas with a low proportion of new and emerging researchers included Education (10 percent) and Humanities and Law (13 percent).

Figure 22
Proportion of new and emerging researchers within PBRF panels



9 WHAT WAS SUBMITTED IN EVIDENCE PORTFOLIOS?

KEY POINTS

- We found that journal articles were by far the most common type of nominated research output and that this increased between 2006 and 2012.
- The number of Intellectual property nominated research outputs was relatively small and concentrated in the Biological Sciences and Physical Sciences and Engineering, Technology and Architecture panels.
- In 2012, a significant proportion of evidence portfolios had the maximum number of other research outputs submitted.
- In 2012, there was a significant proportion of evidence portfolios that submitted the maximum 30 items for both peer esteem and contribution to the research environment.

Introduction

In this section, we examine the different types of research outputs submitted as nominated research outputs along with the number and categories of peer esteem and contribution to the research environment.

Nominated research outputs

For trend analysis purposes, this analysis of the types of nominated research outputs restricts the evidence portfolios in the dataset to those TEOs that participated in both the 2006 and 2012 Quality Evaluations. We also restrict the analysis to those evidence portfolios that received at least a 'C(NE)' quality category.

Figure 23 shows that the journal article was by far the most common nominated research output type in the 2006 and 2012 Quality Evaluations. Between 2006 and 2012, the proportion of journal articles increased from 64 percent to 72 percent. On the other hand, nominated research outputs such as chapters in books (4.4 percent to 3.5 percent) and books (8.3 percent to 7.4 percent) both decreased.

We also focus on the proportion of nominated research outputs that were in the Intellectual property category. This category decreased from 0.4 percent of nominated research outputs in 2006 to 0.3 percent in 2012.

Figure 23
The distribution of nominated research outputs by type

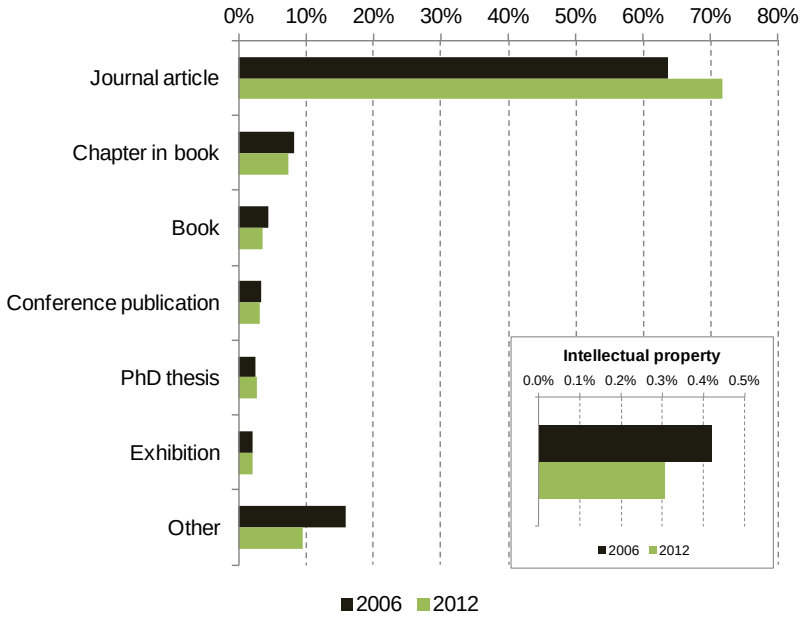
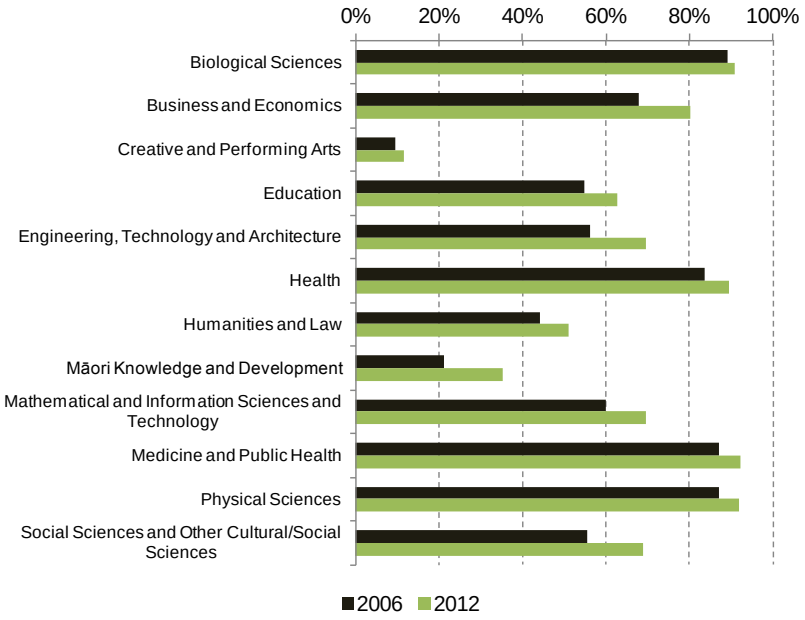


Figure 24 presents the proportion of nominated research outputs that were journal articles by panel in 2006 and 2012. When comparing between panels, the differences in publication convention become apparent. The panels where there were at least 90 percent of nominated research outputs in the journal articles category (Physical Sciences, Medicine and Public Health, Health, and Biological Sciences) is consistent with reviewed articles being the traditional means of publishing research in those fields. Similarly, the low proportion of journal articles in the Creative and Performing Arts Panel (12 percent) reflects other outputs, such as exhibitions, being the main vehicle for disseminating research.

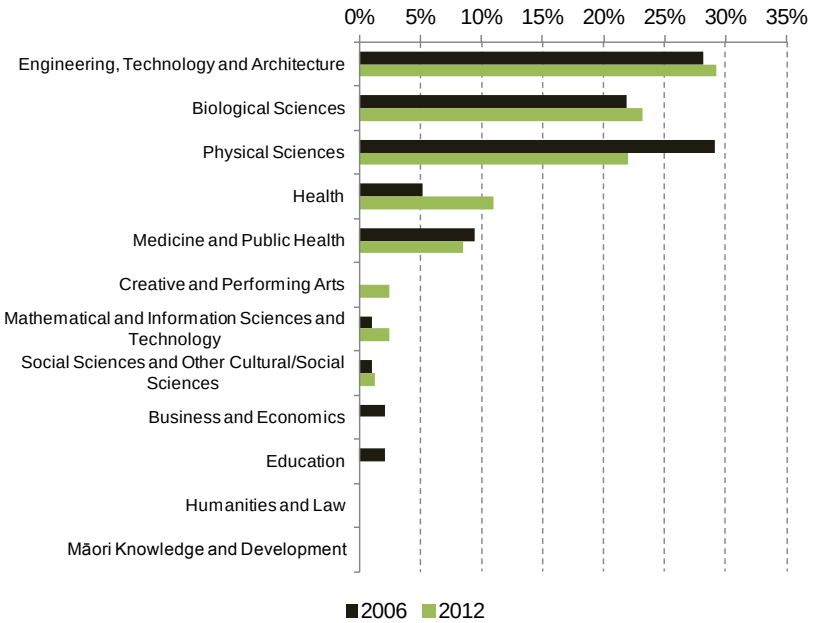
When looking between years, all panels exhibit an increase in the proportion of journal articles submitted as nominated research outputs. The increase in proportions in the Social Sciences and Other Cultural/Social Sciences, Business and Economics and Māori Knowledge and Development panels is significant. This suggests a normalisation in the process of research dissemination towards journal publication.

Figure 24
Change in share of nominated research outputs that are journal articles by panel



We now focus on the Intellectual property category of nominated research outputs. Figure 25 shows the distribution of Intellectual property nominated research outputs by PBRF panel area. This shows that Intellectual property nominated research outputs were mainly clustered around Biological Sciences, Physical Sciences and Engineering, Technology and Architecture. In 2012, 74 percent of Intellectual property nominated research outputs were located in these three areas.

Figure 25
Distribution of Intellectual property nominated research outputs by PBRF panel



Note: The number of nominated research outputs in the Intellectual property category was 96 in 2006 and 82 in 2012.

Other research outputs

In addition to their four nominated research outputs, researchers could list in their evidence portfolios up to 30 other research outputs.

In this section, we examine the number of other research outputs submitted in 2012. The study population is those TEOs that were in both the 2006 and 2012 Quality Evaluations, but we examine the evidence portfolios of all those in that group of TEOs that submitted evidence portfolios. In the 2012 Quality Evaluation, up to 30 other research outputs could be submitted in evidence portfolios.

Table 20 presents the distribution of the number of other research outputs by panel. Panels where there was a large proportion of submissions with high numbers of other research outputs included: Physical Sciences, Engineering, Technology and Architecture, and Medicine and Public Health. Panels with a lower proportion of other research outputs in the 21-30 range included: Māori Knowledge and Development and Creative and Performing Arts.

Overall, around 26 percent of evidence portfolios submitted the maximum of 30 other research outputs.

Table 20
Number of other research outputs submitted by panel in 2012

Panel	0-10	11-20	21-30	N
Biological Sciences	30%	22%	49%	778
Business and Economics	31%	28%	41%	789
Creative and Performing Arts	32%	31%	37%	478
Education	33%	24%	43%	665
Engineering, Technology and Architecture	22%	22%	57%	609
Health	30%	21%	49%	490
Humanities and Law	29%	30%	41%	707
Māori Knowledge and Development	46%	19%	35%	145
Mathematical and Information Sciences and Technology	29%	29%	42%	518
Medicine and Public Health	26%	21%	53%	874
Physical Sciences	22%	20%	58%	454
Social Sciences and Other Cultural/Social Sciences	22%	28%	49%	786
All	**28%**	**25%**	**47%**	**7,293**

Peer esteem

In this section, we examine the categories of peer esteem submitted by researchers who were in TEOs that participated in both the 2006 and 2012 Quality Evaluations. We include all staff at those TEOs that submitted evidence portfolios in the analysis.

Table 21 shows the number of categories submitted in individual evidence portfolios. Staff are allowed to submit up to 30 individual items of peer esteem in their evidence portfolio. In total, 39 percent of staff submitted between 21 and 30 items, with 17 percent of staff submitting the maximum of 30.

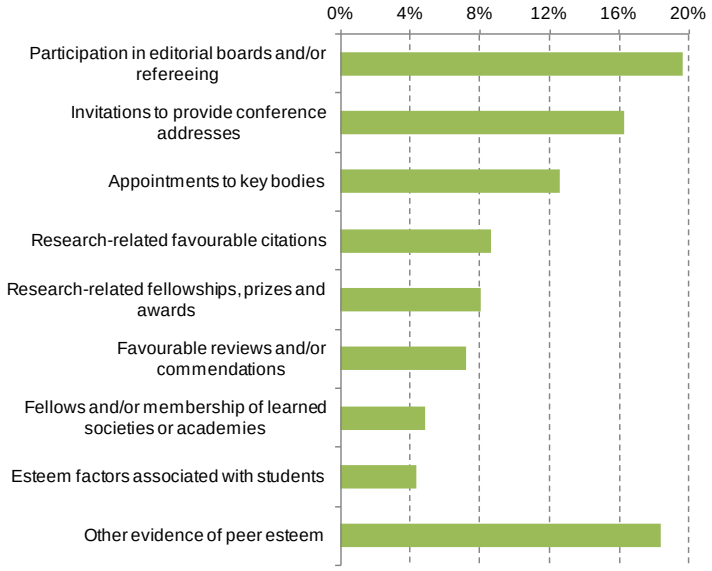
The Humanities and Law Panel had the highest proportion of peer esteem items in the 21-30 range (55 percent), with Biological Sciences having the lowest (28 percent).

Table 21
Number of peer esteem items submitted by panel in 2012

Panel	0-10	11-20	21-30	N
Biological Sciences	42%	30%	28%	778
Business and Economics	29%	28%	43%	789
Creative and Performing Arts	26%	27%	47%	478
Education	31%	32%	37%	665
Engineering, Technology and Architecture	29%	32%	39%	609
Health	38%	32%	30%	490
Humanities and Law	17%	28%	55%	707
Māori Knowledge and Development	37%	24%	39%	145
Mathematical and Information Sciences and Technology	32%	29%	39%	518
Medicine and Public Health	35%	28%	37%	874
Physical Sciences	39%	32%	29%	454
Social Sciences and Other Cultural/Social Sciences	24%	31%	45%	786
All	**31%**	**30%**	**39%**	**7,293**

The distribution of the types of peer esteem submitted by staff in their evidence portfolios is presented in Figure 26. This shows that participation in editorial boards/refereeing was the largest item (excluding the other category), followed by invitations to provide conference addresses. The least submitted category was esteem factors associated with students.

Figure 26
Distribution of peer esteem categories



We present the types of peer esteem submitted in evidence portfolios by PBRF panel in Table 22. Notable among the figures, the Mathematical and Information Sciences and Technology Panel had a high proportion of participation in editorial boards/refereeing (31 percent) and the Humanities and Law panel had a relatively high proportion of items in the research-related favourable citations category (21 percent).

Table 22
Distribution of peer esteem categories by panel in 2012

Panel	Research-related fellowships, prizes and awards	Fellows and/or membership of learned societies or academies	Participation in editorial boards and/or refereeing	Invitations to provide conference addresses	Favourable reviews and/or commendations	Appointments to key bodies	Esteem factors associated with students	Research-related favourable citations	Other evidence of peer esteem	N
Biological Sciences	9%	5%	19%	19%	5%	13%	6%	6%	17%	11,210
Business and Economics	9%	4%	27%	12%	5%	13%	4%	9%	16%	14,136
Creative and Performing Arts	9%	2%	6%	7%	26%	7%	3%	10%	30%	8,764
Education	5%	3%	23%	19%	6%	13%	3%	7%	21%	11,165
Engineering, Technology and Architecture	10%	7%	21%	14%	5%	17%	5%	6%	18%	10,396
Health	8%	6%	21%	18%	4%	15%	5%	4%	18%	7,339
Humanities and Law	5%	3%	15%	15%	12%	7%	3%	21%	18%	14,647
Māori Knowledge and Development	10%	4%	11%	21%	4%	20%	3%	4%	23%	2,437
Mathematical and Information Sciences and Technology	9%	4%	31%	16%	3%	12%	5%	7%	15%	8,664
Medicine and Public Health	10%	8%	17%	23%	4%	16%	4%	5%	13%	14,369
Physical Sciences	9%	6%	17%	18%	5%	14%	6%	7%	18%	6,882
Social Sciences and Other Cultural/Social Sciences	7%	5%	21%	16%	6%	11%	4%	9%	20%	14,433
Total	**8%**	**5%**	**20%**	**16%**	**7%**	**13%**	**4%**	**9%**	**18%**	**124,442**

Table 22 shows that researchers who were considered by the Creative and Performing Arts Panel were much more likely than others to include reviews of their work as evidence of peer esteem but relatively unlikely to cite participation in editorial boards – reflecting the nature of the creative work done by those who submit portfolios for assessment by that panel.

Contribution to the research environment

In this section, we examine the categories of contribution to the research environment submitted by researchers who were in TEOs that participated in both the 2006 and 2012 Quality Evaluations. We include all staff at those TEOs that submitted evidence portfolios in the analysis.

Table 23 shows that 35 percent of staff submitted between 21 and 30 items in their evidence portfolios. Overall, 13 percent submitted the maximum number of 30 items. As was the case for peer esteem, staff in the Humanities and Law Panel area submitted the highest number of items on average.

Table 23

Number of contribution to the research environment items submitted by panel in 2012

Panel	0-10	11-20	21-30	N
Biological Sciences	36%	33%	31%	778
Business and Economics	33%	30%	37%	789
Creative and Performing Arts	41%	28%	31%	478
Education	39%	30%	31%	665
Engineering, Technology and Architecture	29%	33%	38%	609
Health	37%	31%	32%	490
Humanities and Law	25%	32%	43%	707
Māori Knowledge and Development	41%	28%	31%	145
Mathematical and Information Sciences and Technology	34%	30%	36%	518
Medicine and Public Health	31%	32%	36%	874
Physical Sciences	33%	34%	33%	454
Social Sciences and Other Cultural/Social Sciences	27%	35%	38%	786
All	**33%**	**32%**	**35%**	**7,293**

Figure 27 shows the distribution of the types of contribution to the research environment items submitted in evidence portfolios. This shows that supervision of student research was the most common category (21 percent) followed by contributions to the research environment within and outside the TEO (14 percent). We present the distribution of the types of contribution to the research environment items by panel in Table 24.

Figure 27

Distribution of contribution to the research environment categories

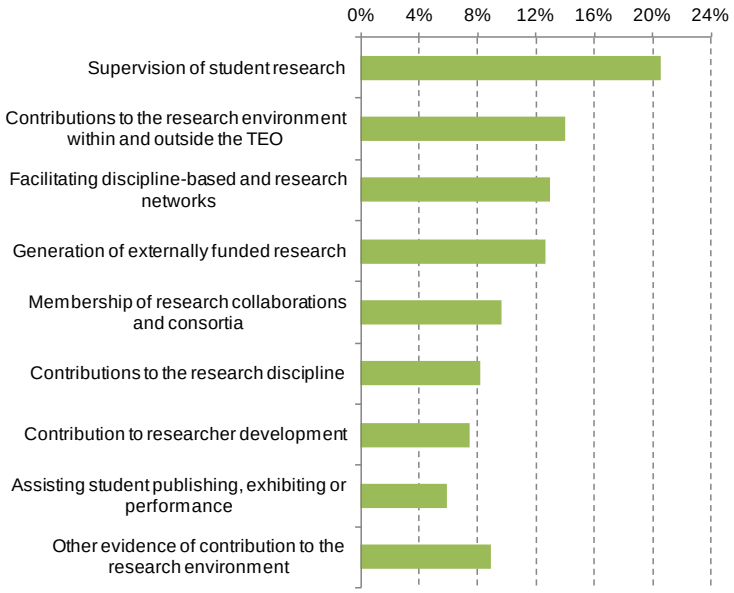


Table 24
Distribution of contribution to the research environment categories by panel in 2012

Panel	Membership of research collaborations and consortia	Contributions to the research discipline	Facilitating discipline-based and research networks	Contributions to the research environment within and outside the TEO	Generation of externally funded research	Contribution to researcher development	Supervision of student research	Other evidence of contribution to the research environment	Assisting student publishing, exhibiting or performance	N
Biological Sciences	11%	7%	12%	10%	19%	7%	20%	8%	6%	12,055
Business and Economics	8%	9%	15%	16%	6%	9%	22%	8%	7%	13,002
Creative and Performing Arts	4%	12%	13%	21%	4%	6%	18%	12%	9%	7,012
Education	9%	8%	13%	17%	9%	10%	20%	8%	5%	9,985
Engineering, Technology and Architecture	10%	6%	15%	11%	16%	5%	23%	8%	6%	10,247
Health	11%	6%	12%	12%	19%	6%	22%	8%	5%	7,469
Humanities and Law	6%	12%	12%	18%	4%	7%	23%	12%	6%	12,734
Māori Knowledge and Development	10%	9%	14%	12%	13%	7%	18%	13%	4%	2,168
Mathematical and Information Sciences and Technology	8%	9%	14%	14%	10%	8%	25%	8%	6%	8,338
Medicine and Public Health	14%	6%	10%	11%	24%	6%	17%	7%	5%	14,414
Physical Sciences	12%	7%	13%	13%	17%	8%	17%	8%	6%	7,170
Social Sciences and Other Cultural/Social Sciences	9%	8%	13%	14%	10%	9%	20%	10%	6%	13,461
Total	**10%**	**8%**	**13%**	**14%**	**13%**	**7%**	**21%**	**9%**	**6%**	**118,055**

As in Table 22 above, the balance of categories in each panel reflects the nature of the research performed by those who submit to each panel. For instance, generation of external research funding was relatively common in Medicine and Public Health, Biological Sciences, Health and Physical Sciences but rare in Creative and Performing Arts and Humanities and Law.

10 THE ASSOCIATION BETWEEN COMPONENT SCORES

KEY POINTS

We analysed the degree of correlation between the three component scores used to help determine quality categories. We found that:

- all three component scores had a high degree of correlation, which generally increased between 2006 and 2012
- the highest correlation was between the peer esteem score and contribution to the research environment score.

Introduction

In this section, we compare the correlation between the research output, peer esteem and contribution to the research environment scores. We do this at the panel level as well as looking at the data overall. For trend analysis purposes, we restrict our analysis to TEOs that participated in both 2006 and 2012 and we also restrict ourselves to those staff who received at least a 'C(NE)' quality category.

We use Pearson's correlation coefficient to measure the linear association between the component scores. The closer the value is to 1, the closer is the linear association between the component scores.

Results

The Pearson's correlation coefficients are presented in Table 25. They show that the highest correlation was between the peer esteem and contribution to the research environment score. In 2012, this correlation was 0.85. This was followed by the correlation between the research output score and peer esteem score (0.79), then the research output score and contribution to the research environment score (0.73).

In 2012, for all but one panel the correlation between research output and peer esteem was at least 0.77. For research output versus contribution to the research environment, the correlation was at least 0.70 in 10 of the 12 panels. For peer esteem versus contribution to the research environment, the correlation coefficient was at least 0.80 in 10 of the 12 panels. So the strong degree of correlation holds across most panels as well as overall.

The level of correlation has increased in the 2012 Quality Evaluation for the comparison of research output score and peer esteem score, as well as the comparison of the research output score with the contribution to the research environment score.

Table 25
Correlation coefficients between component scores and panels in 2006 and 2012

Year	Panel	RO vs PE	RO vs CRE	PE vs CRE
2006	Biological Sciences	0.73	0.66	0.84
	Business and Economics	0.71	0.63	0.84
	Creative and Performing Arts	0.74	0.63	0.77
	Education	0.81	0.77	0.88
	Engineering, Technology and Architecture	0.77	0.71	0.85
	Health	0.74	0.71	0.86
	Humanities and Law	0.77	0.67	0.86
	Māori Knowledge and Development	0.80	0.67	0.83
	Mathematical and Information Sciences and Technology	0.81	0.75	0.87
	Medicine and Public Health	0.77	0.77	0.89
	Physical Sciences	0.82	0.78	0.89
	Social Sciences and Other Cultural/Social Sciences	0.79	0.70	0.83
	All	**0.77**	**0.70**	**0.85**
2012	Biological Sciences	0.80	0.76	0.89
	Business and Economics	0.71	0.66	0.86
	Creative and Performing Arts	0.81	0.68	0.77
	Education	0.82	0.79	0.87
	Engineering, Technology and Architecture	0.77	0.75	0.84
	Health	0.79	0.80	0.88
	Humanities and Law	0.81	0.72	0.84
	Māori Knowledge and Development	0.82	0.72	0.75
	Mathematical and Information Sciences and Technology	0.77	0.70	0.86
	Medicine and Public Health	0.81	0.80	0.89
	Physical Sciences	0.83	0.78	0.89
	Social Sciences and Other Cultural/Social Sciences	0.78	0.74	0.87
	All	**0.79**	**0.73**	**0.85**

Notes: 1. The number of evidence portfolios in the 2006 analysis = 5,739 and in 2012 = 6,733. 2. This analysis is restricted to those TEOs that participated in both the 2006 and 2012 Quality Evaluations.

Another way of examining the association between the component scores is to look at the difference between the scores. In Figure 28 and Table 26 we present the difference between component scores for individual evidence portfolios in 2006 and 2012. For example, in 2006 around 39 percent of evidence portfolios received the same research output and peer esteem scores. Another 43 percent of evidence portfolios exhibited a difference of 1 in their research output and peer esteem scores.

In 2012, the closest association was between the peer esteem and contribution to the research environment score. In 53 percent of cases, the evidence portfolios received an identical score, while around 92 percent of cases were within one point. The relationship was similar in the 2006 Quality Evaluation.

The closeness of the peer esteem and contribution to the research environment scores suggests that these two components may not be making a distinct contribution to the overall research quality category.

Figure 28

Distribution of differences between component scores

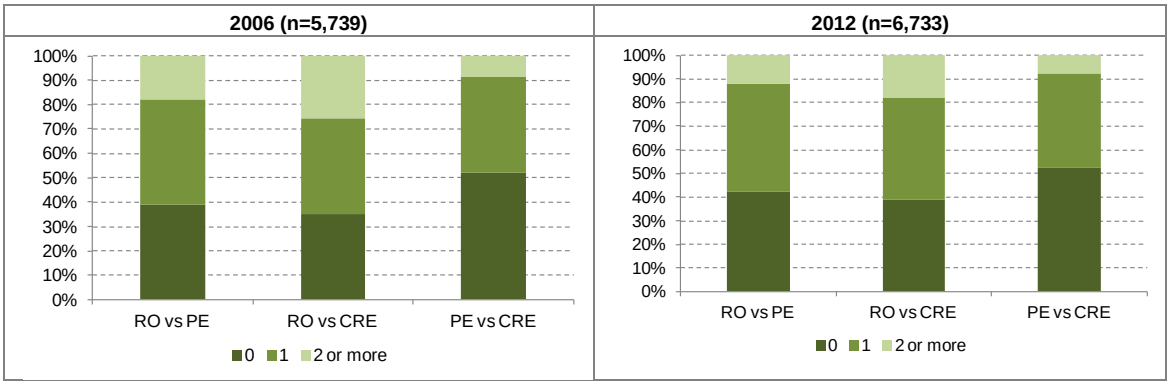


Table 26

Differences in component scores

Year	Difference in score	RO vs PE	RO vs CRE	PE vs CRE
2006	0	39%	35%	52%
	1	43%	39%	39%
	2 or more	18%	26%	9%
	Total	**100%**	**100%**	**100%**
2012	0	42%	39%	53%
	1	46%	43%	40%
	2 or more	12%	18%	8%
	Total	**100%**	**100%**	**100%**

11 A PROFILE OF EVIDENCE PORTFOLIOS SEEN BY THE EXPERT ADVISORY GROUPS

KEY POINTS

- We found that the proportion of evidence portfolios submitted to the Expert Advisory Groups in the 2012 Quality Evaluation was relatively small.
- The evidence portfolios examined for their commercialisation content were mainly from the 'Business and economics', 'Biological sciences' and 'Physical sciences' panels.
- The performance profile of the evidence portfolios accepted by the Expert Advisory Groups was similar to the overall performance of staff in the Quality Evaluation.

Introduction

In this section, we present a profile of the evidence portfolios that were accepted by the Expert Advisory Groups.

Results

Table 27 shows the number of evidence portfolios that were submitted and the number that were accepted for assessment by the Expert Advisory Groups and their associated subgroups. The vast majority of evidence portfolios submitted by TEOs to the Expert Advisory Groups were accepted for assessment. The number of evidence portfolios submitted to the Professional and Applied Expert Advisory Group was a small percentage of all portfolios. This might be because TEOs believed panels already had the expertise to assess the evidence portfolios or that there was not a significant amount of this activity taking place.

Table 27
Acceptance rate for evidence portfolios submitted to Expert Advisory Groups in 2012

	Submitted	Accepted	Acceptance rate	Accepted as % of all evidence portfolios
Commercial	101	101	100%	1.4%
Social	37	31	84%	0.5%
Environmental	96	95	99%	1.3%
Professional practice	114	106	93%	1.6%
Pacific research	145	131	90%	2.0%

We have already seen in Chapter 5 that there was no statistically significant difference in the grades assigned to evidence portfolios that were accepted by the Expert Advisory Groups. In Tables 28 to 32 we present a more detailed profile of the evidence portfolios that were accepted by the Expert Advisory Groups.

Professional and Applied – Commercial

Evidence portfolios assessed were mainly located in three panels ('Business and economics', 'Biological sciences' and 'Physical sciences'). While seven in every eight submissions were by university staff, that reflected the fact that the overwhelming number of all evidence portfolios in the 2012 Quality Evaluation were from university staff. In fact, those who worked outside the universities were more likely to have submitted their portfolios to the Professional and Applied Expert Advisory Group; 13 percent of submissions to the panel came from the 9 percent of non-university researchers.

Other groups with a higher likelihood of submitting to this panel included:

- men
- professors and associate professors
- researchers aged 40 or above.

Table 28
Profile of staff who were assessed by the Professional and Applied Expert Advisory Group Commercial subgroup

		Commercial	All
Panel	Biological Sciences	19%	11%
	Business and Economics	21%	11%
	Creative and Performing Arts	7%	7%
	Engineering, Technology and Architecture	10%	8%
	Health	9%	7%
	Medicine and Public Health	9%	12%
	Physical Sciences	17%	6%
	Other	9%	39%
Subsector	University	87%	91%
	Other	13%	9%
Title	Professor	23%	15%
	Associate professor	18%	14%
	Senior lecturer	29%	39%
	Lecturer	18%	23%
	Other	13%	8%
Quality categories	A	14%	12%
	B	38%	37%
	C	37%	30%
	C(NE)	5%	14%
	R	7%	6%
	R(NE)	0%	2%
Gender	Female	20%	41%
	Male	80%	59%
Age	40 and under	10%	24%
	41-50	31%	29%
	51-60	30%	28%
	61 or over	30%	19%
Total		**101**	**7,324**

The performance of the assessed evidence portfolios was in line with the overall results in the 2012 Quality Evaluation, though there was a lower proportion of evidence portfolios receiving a new and emerging quality category than all submitted evidence portfolios.

Professional and Applied – Social

The evidence portfolios assessed by this subgroup were mainly located in the 'Social sciences' (39 percent) and 'Creative arts' (19 percent) panels. The age distribution of staff who had evidence portfolios assessed by this subgroup tended to be slightly younger than the overall population, but the gender distribution of evidence portfolios was in line with all submitted evidence portfolios.

Table 29

Profile of staff who were assessed by the Professional and Applied Expert Advisory Group Social subgroup

		Social	All
Panel	Creative and Performing Arts	19%	7%
	Social Sciences and Other Cultural/Social Sciences	39%	11%
	Other	42%	83%
Subsector	University	87%	91%
	Other	13%	9%
Title	Professor/Associate professor	32%	29%
	Senior lecturer	23%	39%
	Lecturer	19%	23%
	Other	26%	8%
Quality categories	B	45%	37%
	C	32%	30%
	Other	23%	33%
Gender	Female	39%	41%
	Male	61%	59%
Age	50 and under	45%	53%
	51-60	26%	28%
	61 or over	29%	19%
Total		**31**	**7,324**

Professional and Applied – Environmental

The evidence portfolios assessed by this subgroup were located mainly in the 'Biological sciences' (33 percent) and 'Business and economics' (23 percent) panels. The distribution of quality categories received by the evidence portfolios assessed by this subgroup was broadly in line with the total study population.

Table 30

Profile of staff who were assessed by the Professional and Applied Expert Advisory Group Environmental subgroup

		Environmental	All
Panel	Biological Sciences	33%	11%
	Business and Economics	23%	11%
	Creative and Performing Arts	5%	7%
	Engineering, Technology and Architecture	9%	8%
	Health	5%	7%
	Physical Sciences	7%	6%
	Social Sciences and Other Cultural/Social Sciences	7%	11%
	Other	9%	40%
Title	Professor	25%	15%
	Associate professor	13%	14%
	Senior lecturer	27%	39%
	Lecturer	18%	23%
	Other	17%	8%
Quality categories	A	12%	12%
	B	41%	37%
	C	27%	30%
	C(NE)	13%	14%
	Other	7%	8%
Gender	Female	34%	41%
	Male	66%	59%
Age	40 and under	29%	24%
	41-50	25%	29%
	51-60	27%	28%
	61 or over	18%	19%
Total		**95**	**7,324**

Professional and Applied – Professional Practice

The evidence portfolios assessed by this subgroup were mainly in the 'Engineering, technology and architecture' (22 percent), 'Social sciences' (15 percent) and 'Medicine and public health' (13 percent) panels. The staff who had evidence portfolios assessed by this subgroup were older and more likely to be men than the overall population.

Table 31

Profile of staff who were assessed by the Professional and Applied Expert Advisory Group Professional Practice subgroup

		Professional Practice	All
Panel	Biological Sciences	9%	11%
	Creative and Performing Arts	8%	7%
	Education	9%	9%
	Engineering, Technology and Architecture	22%	8%
	Health	8%	7%
	Humanities and Law	6%	10%
	Medicine and Public Health	13%	12%
	Social Sciences and Other Cultural/Social Sciences	15%	11%
	Other	8%	26%
Subsector	University	92%	91%
	Other	8%	9%
Title	Professor	15%	15%
	Associate professor	18%	14%
	Senior lecturer	46%	39%
	Lecturer	14%	23%
	Other	7%	8%
Quality categories	A	5%	12%
	B	39%	37%
	C	39%	30%
	C(NE)	9%	14%
	Other	8%	8%
Gender	Female	32%	41%
	Male	68%	59%
Age	40 and under	8%	24%
	41-50	26%	29%
	51-60	41%	28%
	61 or over	25%	19%
Total		**106**	**7,324**

Pacific Research

The evidence portfolios assessed by this Expert Advisory Group were situated mainly in the 'Social sciences' (32 percent) and 'Education' (17 percent) panels. They were more likely to be female, and older, than the overall study population.

Table 32

Profile of staff who were assessed by the Pacific Research Expert Advisory Group

		Pacific Research	All
Panel	Biological Sciences	6%	11%
	Creative and Performing Arts	6%	7%
	Education	17%	9%
	Engineering, Technology and Architecture	7%	8%
	Humanities and Law	11%	10%
	Medicine and Public Health	11%	12%
	Physical Sciences	5%	6%
	Social Sciences and Other Cultural/Social Sciences	32%	11%
	Other	5%	27%
Subsector	University	92%	91%
	Other	8%	9%
Title	Professor	11%	15%
	Associate professor	11%	14%
	Senior lecturer	42%	39%
	Lecturer	27%	23%
	Other	9%	8%
Quality categories	A	8%	12%
	B	38%	37%
	C	36%	30%
	C(NE)	13%	14%
	Other	5%	8%
Gender	Female	54%	41%
	Male	46%	59%
Age	40 and under	11%	24%
	41-50	28%	29%
	51-60	31%	28%
	61 or over	29%	19%
Total		**131**	**7,324**

APPENDIX

Quality Evaluation

The assessment of research quality – Quality Evaluation (QE) – is undertaken by interdisciplinary peer review panels consisting of disciplinary experts from both within New Zealand and overseas. These panels provide expert coverage of the subject areas within each panel's respective field of responsibility.

Each researcher presents their research in the form of an evidence portfolio (EP). The EP has three components:

- Research outputs: the outputs of a staff member's research (each staff member nominates up to four of their best research outputs for primary consideration by the panel, and up to 30 other research outputs (OROs))

- Peer esteem: an indication of the quality of the research of the staff member, as recognised by their peers in the form of fellowships, prizes, awards, memberships of learned societies, participation in editorial boards, invitations to present at conferences, favourable reviews, etc (each staff member determines their top 30 examples, providing a list and details to the peer review panel)

- Contribution to the research environment: the staff member's contribution to a vital, high-quality research environment, both within the TEO and beyond it, as evidenced by membership in research consortia, generation of external research income, supervision of student research, etc (each staff member determines their top 30 examples, providing a list and details to the peer review panel).

In assessing the EP, the scores assigned to each component are weighted to calculate a weighted total score, which corresponds to a quality category. There are six quality categories:

- Quality category 'A': For an EP to be assigned an 'A', it would normally be expected that the staff member has, during the assessment period in question, produced research outputs of a world-class standard, established a high level of peer recognition and esteem within the relevant subject area of their research, and made a significant contribution to the New Zealand and/or international research environments

- Quality category 'B': For an EP to be assigned a 'B', it would normally be expected that the staff member has, during the assessment period in question, produced research outputs of a high quality, acquired recognition by peers for their research at least at a national level, and made a contribution to the research environment beyond their institution and/or a significant contribution within their institution

- Quality category 'C': For an EP to be assigned a 'C', it would normally be expected that the staff member has, during the assessment period in question, produced a reasonable quantity of quality-assured research outputs, acquired some peer recognition for their research, and made a contribution to the research environment within their institution. (This Quality Category is available for the EPs of all PBRF-eligible staff members except new and emerging researchers.)

- Quality category 'C(NE)': For an EP to be assigned a 'C(NE)', a new or emerging researcher would normally be expected, during the assessment period in question, to have produced a reasonable platform of research, as evidenced by having: either (a) completed their doctorate

or equivalent qualification and produced at least two quality-assured research outputs, or (b) produced research outputs equivalent to a doctorate and at least two quality-assured research outputs. (This Quality Category is available for the EPs of new and emerging researchers only.)

- Quality category 'R': An EP will be assigned an 'R' when it does not demonstrate the quality standard required for a 'C' Quality Category or higher. (This Quality Category is available for the EPs of all PBRF-eligible staff members except new and emerging researchers.)

- Quality category 'R(NE)': An EP will be assigned an 'R(NE)' when it does not demonstrate the quality standard required for a 'C(NE)' Quality Category or higher. (This Quality Category is available for the EPs of new and emerging researchers only.)

EPs are evaluated through a rigorous, collaborative process. EPs are assigned to a primary and secondary panellist who independently assess the EP and then agree an initial score together. This score is then discussed at the panel meeting and a final score is decided. All the scores are moderated by that panel and then between the other panels.

Funding in relation to the QE is based on:

- the quality categories assigned to EPs

- the funding weighting for the subject area to which EPs have been assigned

- the full-time equivalent (FTE) status of the participating TEO's PBRF-eligible staff as at the date of the PBRF Census.

QEs are conducted every six years. However, given the need for a managed transition, the second QE round took place three years after the first, but that was a partial round. Thus, QEs have taken place in 2003 and 2006 (partial). The third QE took place in 2012.

Table 33 shows the 12 peer review panels that assess EPs and the subject areas that each panel is responsible for assessing.

Table 33
PBRF EP assessment panels and subject areas

Panel	Subject area
Biological Sciences	Agriculture and other applied biological sciences
	Ecology, evolution and behaviour
	Molecular, cellular and whole organism biology
Business and Economics	Accounting and finance
	Economics
	Management, human resources, industrial relations, international business and other business
	Marketing and tourism
Creative and Performing Arts	Design
	Music, literary arts and other arts
	Theatre and dance, film and television and multimedia
	Visual arts and crafts
Education	Education
Engineering, Technology and Architecture	Architecture, design, planning, surveying
	Engineering and technology
Health	Dentistry
	Nursing
	Other health studies (including rehabilitation therapies)
	Pharmacy
	Sport and exercise science
	Veterinary studies and large animal science
Humanities and Law	English language and literature
	Foreign languages and linguistics
	History, history of art, classics and curatorial studies
	Law
	Philosophy
	Religious studies and theology
Māori Knowledge and Development	Māori knowledge and development
Mathematical and Information Sciences and Technology	Computer science, information technology, information sciences
	Pure and applied mathematics
	Statistics
Medicine and Public Health	Biomedical
	Clinical medicine
	Public health
Physical Sciences	Chemistry
	Earth sciences
	Physics
Social Sciences and Other Cultural/Social Studies	Anthropology and archaeology
	Communications, journalism and media studies
	Human geography
	Political science, international relations and public policy
	Psychology
	Sociology, social policy, social work, criminology and gender studies

There are two key principles underpinning the eligibility of a TEO's staff member to participate in a QE:

- The individual is expected to contribute to the learning environment at the degree level

  and/or

- The individual is expected to make a sufficiently substantive contribution to research activity.

Other elements underpinning the staff participation criteria are:

- The staff member has an explicit requirement to teach and/or undertake research as one of their employment functions, as at the date of the PBRF Census.

- A sufficiently substantive contribution is determined by applying the substantiveness test.

- The full-time equivalent (FTE) counted in the QE for each PBRF-eligible staff member is generally that contained in their employment agreement.

- Employment history in the 12-month period prior to the PBRF Census date is to be apportioned on an FTE basis to ensure fair representation of staff time, and to minimise 'poaching'.

- Staff employed in wholly owned subsidiaries and fully controlled trusts of the TEO are PBRF eligible, since these bodies operate under the control of the participating TEO.

- Provision has been made to allow staff members based overseas, and staff members sub-contracted to TEOs by non-TEOs, to be PBRF eligible under certain conditions.

REFERENCES

Smart, W. (2005) *What determines the research performance of staff in new Zealand's tertiary education sector? An analysis of the Performance-Based Research Fund Quality Evaluation*, Ministry of Education: Wellington.

Smart, W. (2008a) *Research par excellence*, Ministry of Education: Wellington.

Smart, W. (2008b) *Trends in measured research quality: an analysis of the PBRF Quality Evaluation results*, Ministry of Education: Wellington.

Smart, W. (2013) *Analysis of the impact of the PBRF – interim findings*, Ministry of Education: Wellington.

Tertiary Education Commission (2007) *Performance-Based Research Fund: evaluating research excellence. The 2006 Assessment*, Tertiary Education Commission: Wellington.

TEC (2013) *Performance-Based Research Fund: evaluating research excellence. The 2012 Assessment*, Tertiary Education Commission: Wellington.

Wensvoort, M. (2012) *2011 Tertiary education enrolments*, Ministry of Education: Wellington.

MINISTRY OF EDUCATION

Te Tāhuhu o te Mātauranga

newzealand.govt.nz