



MINISTRY OF EDUCATION

Te Tāhuhu o te Mātauranga

The measurement of the research performance of Tertiary Education Organisations

An analysis of the impact of weightings in the Performance-Based
Research Fund

Report

The measurement of the research performance of Tertiary Education Organisations. An analysis of the impact of weightings in the Performance-Based Research Fund.

Author

Warren Smart, Senior Research Analyst
Tertiary Sector Performance Analysis & Reporting, Ministry of Education
Email warren.smart@minedu.govt.nz
Telephone 64-4-463-8035
Fax 64-4-463-8526

Acknowledgements

The author gratefully acknowledges the support Roger Smyth, Jonathan Boston and Brenden Mischewski in their role as peer reviewers.

Published by:

MINISTRY OF EDUCATION
© Crown Copyright
All rights reserved
All enquiries should be made to the publisher

June 2005

This report is available from the Ministry of Education website
www.minedu.govt.nz/goto/tertiaryanalysis

ISBN (Web) 0-478-13316-2

Contents

TABLES AND FIGURES.....	3
1 INTRODUCTION.....	4
2 MEASURING THE RESEARCH PERFORMANCE OF PBRF ELIGIBLE STAFF.....	5
3 THE IMPACT OF WEIGHTING QUALITY CATEGORIES ON TEO PERFORMANCE... 	6
4 THE IMPACT OF METHODOLOGY AND WEIGHTING IN THE PANEL ASSESSMENT	8
COMPARING TEO PERFORMANCE USING THE WEIGHTED QUALITY CATEGORY SCORE AND THE OQS ..	8
COMPARING TEO PERFORMANCE USING THE WEIGHTED QUALITY CATEGORY SCORE, OQS (UNWEIGHTED) AND DEA	10
5 CONCLUSION.....	11
APPENDIX.....	13
BIBLIOGRAPHY	14

Tables and figures

Figure 1: TEO quality category score per FTE on a weighted and unweighted basis as a proportion of the University of Auckland score (all PBRF eligible staff)	7
Figure 2: TEO average weighted quality category score and OQS as a proportion of the University of Auckland score (panel assessed staff only).....	9
Figure 3: TEO average weighted quality category score, OQS, OQS (unweighted) and DEA ratio, as a proportion of the University of Auckland score (panel assessed staff only).....	11
Table 1: FTE staffing, PBRF quality scores and DEA ratio by TEO	13

1 Introduction

The implementation of the Performance-Based Research Fund (PBRF) in 2004 was designed to reward excellence in research. As a result, for the first time in New Zealand, a comprehensive measurement of the relative research performance of Tertiary Education Organisations (TEOs) was undertaken and then reported.

The measure used to rank the research performance of PBRF eligible staff at TEOs in the 2003 assessment, quality category score per full-time equivalent (FTE) staff member, indicated a significant degree of variation in performance¹. Although variation in performance is to be expected, the degree to which this variation is a reflection of the actual research performance of TEOs, or is influenced by the methodology used to calculate the quality category score, is the subject of this paper. Generally, universities were the top performers, while polytechnics, colleges of education and some private training establishments received significantly lower quality scores in comparison.

Weightings are applied in several areas of the PBRF to formulate funding allocations for TEOs². The key weighting involves the allocation of 60 percent of the total PBRF funding to the quality evaluation of researcher performance, 25 percent to research degree completions and 15 percent for external research income. Weightings are also applied in other areas. These include weighting graduate completions by subject area and the ethnicity of the graduate and the weighting of the various performance measures that make up the quality evaluation³.

In this paper, two areas where weightings were applied in the PBRF are investigated to measure the degree to which they impact on the relative performance of TEOs.

Firstly, data for all PBRF eligible staff is used to analyse the impact of the weighting of A, B, C and R staff on the relative performance of TEOs in the quality evaluation. This analysis involves comparing the relative performance of TEOs using the published weighted measure of quality category score per FTE, with a measure where the weightings are removed.

Then, we analyse the impact of the methodology used by the peer review panels in assigning quality categories to those staff that were assessed by a PBRF assessment panel in the 2003 quality evaluation. This includes analysing the effect of weighting research outputs, peer esteem and contribution to research environment within the quality evaluation and the process used to classify staff into their various quality categories.

¹ For example, the best performing TEO, the University of Auckland achieved an average quality category score of 3.94 per PBRF eligible FTE, whereas the lowest performing TEO, Bethlehem Institute of Education, received a score of 0.

² See Tertiary Education Commission (2003) *Performance-Based Research Fund, Evaluating Research Excellence, the 2003 assessment*, for a detailed explanation of the weightings and methodology used in the PBRF.

³ See Ministry of Education (2005) *An analysis of funding allocations for staff and research degree completions in the Performance-Based Research Fund*, for an analysis of how the weightings applied in the PBRF impacted on the funding allocations to staff and research degree completions.

The analysis of relative TEO performance, using solely information relating to those whose evidence portfolios were quality assessed, is useful in investigating the methodology used by the peer review panels. However, as not all staff had evidence portfolios forwarded to the peer review panels, those TEOs that had a smaller proportion of non-panel assessed staff will be advantaged, as they will receive research scores that are inflated on a per FTE basis⁴. Therefore, the relative performance of TEOs, using only panel assessed staff data, should not be compared with measures of TEO performance that include all PBRF eligible staff.

2 Measuring the research performance of PBRF eligible staff

In the first stage of the 2003 quality evaluation, TEOs rated the research performance of their staff in a self-assessment exercise. The evidence portfolios of staff that were considered to be of an A, B or C quality category standard were forwarded to the subject panels for peer review. The peer review panels then assigned final quality categories to these staff. Staff that did not have evidence portfolios forwarded to the panels were assigned an R quality category automatically.

The quality category of each PBRF eligible staff member was determined by three measures of staff research performance. These were: the quality of research outputs (RO), the esteem with which the staff member was held by their peers (PE) and their contribution to the research environment (CRE). For each of these measures, a staff member was assigned a score between 0 and 7, with 7 representing the highest level of performance. To obtain an overall measure of research performance, the overall quality score (OQS), a weighting of 70 percent was then applied to the RO score, 15 percent to the PE score and 15 percent to the CRE score. As it is the key measure of staff research performance, the RO score received the highest weighting. The PE and CRE scores are valued as indicators of the extent to which the researcher contributes to the creation of a research culture in the organisation. From a policy perspective, these indicators are less important than the RO score and therefore were allocated lower weightings.

The OQS was calculated using the following formula and results in a score between 0 and 700:

$$\text{OQS} = 70 \times \text{RO} + 15 \times \text{PE} + 15 \times \text{CRE}$$

The OQS was then used as an aid in determining the quality category of staff. The peer review panels used a holistic approach in assigning quality categories and a result could be altered from that indicated by the OQS⁵.

Generally, for those evidence portfolios that were peer reviewed, an OQS score between 0 and 199 would earn an R quality category for a staff member, a score between 200 and 399 would earn a C, a score between 400 and 599 a B and a score between 600 and 700 an A. The range of score between 0 and 700 may give the

⁴ For universities, the proportion of PBRF eligible staff that were not panel assessed ranged from 66 percent for the Auckland University of Technology to 8 percent for the University of Canterbury.

⁵ Approximately 1 percent of staff that submitted evidence portfolios to the review panels received a quality category different from that indicated by their OQS.

impression that there is a high degree of accuracy in these scores that is more apparent than real. Therefore, the application of a cruder scale is perhaps a more robust way to report on performance.

In the funding formula, staff who achieved an A quality category received a weighting of 5, B staff a weighting of 3 and C staff a weighting of 1. The higher weighting assigned to A and B level staff was designed to reflect the policy goal of rewarding excellence in research⁶.

The published average quality category scores for TEOs doubled the weightings that had applied for funding purposes. Therefore, staff assessed as achieving quality category A received a score of 10, B level staff a score of 6, C level staff a score of 2 and R level staff a score of 0. These scores were then aggregated for each TEO and an average quality score per FTE calculated for comparison purposes.

3 The impact of weighting quality categories on TEO performance

To analyse the impact of the weightings on relative TEO performance, the quality category scores attached to the performance of staff had the weightings reduced to be more in line with the relative level of performance indicated by the OQS classifications. To achieve this, an A level staff member was assigned a score of 3, a B level staff member a score of 2, a C level staff member a score of 1 and an R level staff member a score of 0⁷. This score is then weighted by the FTE status of researchers and presented as a per FTE figure. For the remainder of this report this score is referred to as the unweighted quality category score.

In order to compare the weighted quality category score with the unweighted score, each TEO's performance is expressed as a proportion of the score achieved by the University of Auckland, the top performer in the 2003 quality evaluation (see Figure 1).

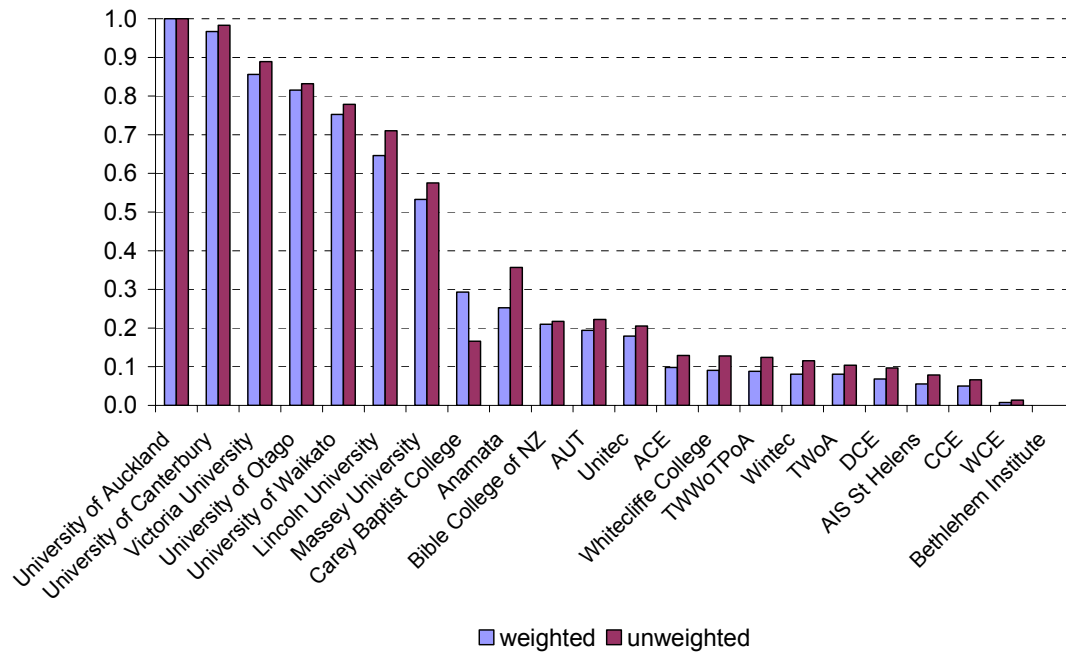
The effect of softening the weightings was to remove the advantage that many A and B quality category staff give to a TEO. With the exception of Carey Baptist College, all the TEOs improved performance, relative to the University of Auckland, when the weightings on the quality categories were reduced. On average, the relative performance of TEOs, as a proportion of the score achieved by the University of Auckland, improved from 0.30 to 0.32. Of the major TEOs, the largest improvement in relative performance was by Lincoln University, which improved from 0.65 to 0.71.

Therefore, to a degree, the use of weightings exaggerated the variation in the performance of the TEOs. Although the ranking order of the major TEOs was unaffected, the impact of the weightings on relative performance has implications for the anticipated outcomes of the publication of PBRF results.

⁶ See Ministry of Education and Transitory Tertiary Education Commission (2002) *Investing in Excellence, The Report of the Performance-Based Research Fund Working Group*, pp 7.

⁷ As the OQS interval for A level staff is only 100 points wide compared with 200 for the other quality categories, A level staff effectively still receive a higher weighting.

Figure 1: TEO quality category score per FTE on a weighted and unweighted basis as a proportion of the University of Auckland score (all PBRF eligible staff)



Notes:

1. 'AUT' is the Auckland University of Technology, 'ACE' is the Auckland College of Education, 'TWVoTPoA' is the Te Whare Wānanga o Te Pihopatanga o Aotearoa, 'TWoA' is the Te Wānanga o Aotearoa, 'DCE' is the Dunedin College of Education, 'CCE' is the Christchurch College of Education and 'WCE' is the Wellington College of Education.
 2. For the 'weighted' scores, A level staff were assigned a score of 10, B staff a score of 6, C staff a score of 2 and R staff a score of 0.
 3. For the 'unweighted' scores, A level staff were assigned a score of 3, B staff a score of 2, C staff a score of 1 and R staff a score of 0.
- Sources: Ministry of Education, Tertiary Education Commission.

One of the key desired outcomes of the PBRF was the production of information about the relative quality of providers for interested stakeholders. In particular, students making enrolment decisions and private sector entities looking for research partners would be able to use the published figures to make informed decisions⁸.

If stakeholders are to make decisions based on TEO research performance, then it might be argued that an unweighted measure of performance would provide them with a more accurate picture of the relative standings of TEOs. The higher weighting of A and B researchers could still apply for funding purposes and therefore reward excellence in this manner.

As it currently stands, the TEOs that received a higher relative published score as a result of the weightings stand to receive the additional benefits of potential increases in student enrolments and external research contracts.

The approach used in the Research Assessment Exercise (RAE) in the United Kingdom weights research performance for funding purposes, but not for publication purposes. In the RAE, academic departments in universities are rated according to a 7

⁸ See Ministry of Education and Transitory Tertiary Education Commission (2002) *Investing in Excellence, The Report of the Performance-Based Research Fund Working Group*, pp 22.

number scale that ranges from 1 - 5*, with the latter being the highest level of performance. In the latest RAE in 2001, a department that achieved a score of 5* received a weighting 8 times that of a department that received a 3a, the lowest category to receive funding. However, in the publication of the results, the RAE presented the scores of departments in an unweighted manner.

4 The impact of methodology and weighting in the panel assessment

The analysis in this section uses data for panel assessed staff only. Therefore it should not be compared with analysis of TEO performance that includes all PBRF eligible staff. The exclusion of non-panel assessed staff favours those institutions with a low proportion of non-panel assessed data.

There are four issues that arise as a result of the methodology used by the peer review panels to assign quality categories to staff. Firstly, the weightings applied to the three components of the quality evaluation place a greater emphasis on research outputs. Secondly, staff who achieved an OQS towards the bottom of a classification were given the same quality score as staff towards the top. Thirdly, staff just above an OQS classification division received a much higher quality score than staff just below that point. Finally, staff who achieved an OQS score between 0 and 199 did not receive any points in the quality category score.

To illustrate the above points, consider the example of a staff member who received an OQS of 585. This person was assigned a B grade and a score of 6 in the quality evaluation. By contrast, a staff member who achieved an OQS score of 600 received an A and a score of 10. Therefore, the relative performance of the B staff member decreased from 97.5 percent of that achieved by the A staff member using the OQS to just 60 percent using the quality category score.

Further, the staff member who achieved the OQS of 585 received the same quality category score (6) as a staff member who scored 400. Therefore the relative performance of the staff member who had the OQS of 400 went from being 68 percent of that by the staff member who received an OQS of 585, to being assessed as having the same level of performance using the quality category score.

To analyse the impact of the methodology used to assign the quality category scores, the scores that a TEO would have received using different systems of weighting are calculated. Firstly, the weighted average quality category score of TEOs is compared with their OQS. Secondly the RO, PE and CRE scores have different weightings applied to them and the resulting relative performance of TEOs is compared with their weighted average quality category score and OQS.

Comparing TEO performance using the weighted quality category score and the OQS

In order to analyse the impact of the weightings on the measurement of panel assessed staff performance, the average weighted quality category score per FTE is compared with the average OQS score per FTE. For means of comparison, the measures have

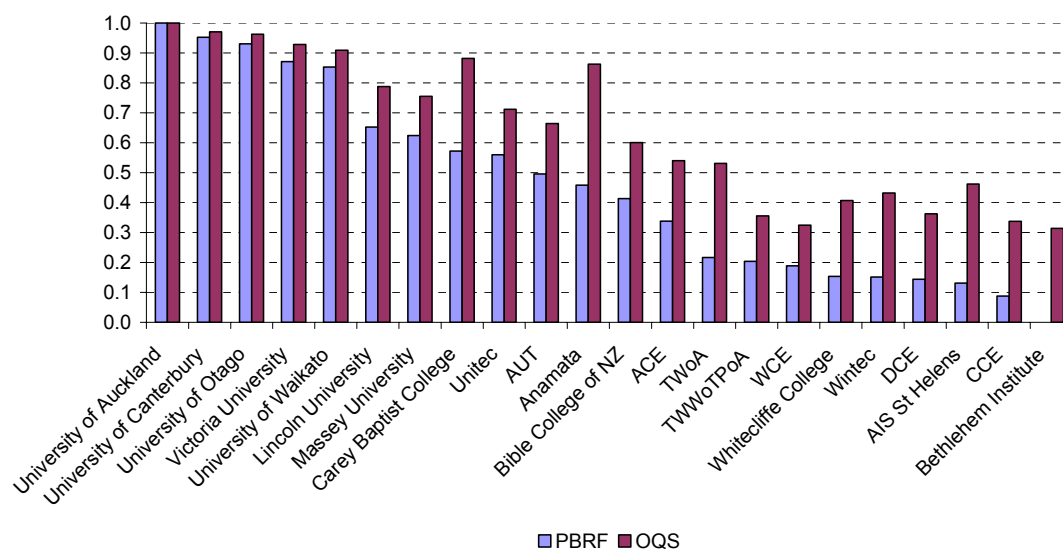
been expressed as a proportion of the score achieved by the University of Auckland, the best performing TEO in the 2003 quality evaluation.

As illustrated in Figure 2, all TEOs exhibited higher relative performance, compared with the University of Auckland, when using the OQS as the measure of performance. Of the universities, Massey, Lincoln and AUT showed the largest improvement in relative performance compared with the University of Auckland. This reflects the higher weighting applied to A and B quality category staff in the quality category score as well as the impact of staff who obtained an OQS between 0 and 199 now having this score counted.

Some of the largest gains in relative performance were by TEOs that had performed poorly when using the weighted quality category score as the unit of measurement. For example, Bethlehem Institute of Education, which did not even receive a weighted quality category score, achieved an OQS per FTE that was just over 30 percent of that achieved by the University of Auckland.

The apparent large gains in performance for Anamata and Carey need to be placed in perspective. These TEOs had 2 and 8.6 PBRF eligible FTEs respectively and therefore the results are particularly sensitive to changes in weightings.

Figure 2: TEO average weighted quality category score and OQS as a proportion of the University of Auckland score (panel assessed staff only)



Notes:

1. 'AUT' is the Auckland University of Technology, 'ACE' is the Auckland College of Education, 'TWwTPoA' is the Te Whare Wānanga o Te Pihopatanga o Aotearoa, 'TWoa' is the Te Wānanga o Aotearoa, 'DCE' is the Dunedin College of Education, 'CCE' is the Christchurch College of Education and 'WCE' is the Wellington College of Education.
2. 'PBRF' is calculated by dividing the average weighted quality category score per FTE for each TEO by the average score achieved by the University of Auckland.
3. 'OQS' is calculated by dividing the average OQS per FTE for each TEO by the average OQS per FTE achieved by the University of Auckland.
4. It is important to note that Anamata and Carey have staffing of just 2.0 and 8.6 PBRF eligible FTEs respectively and their performance is especially sensitive to changes in the weightings.
5. The relative performance of TEOs in this graph should not be compared with measures of performance that includes data for all PBRF eligible staff. TEOs with a small proportion of non-panel assessed staff will be advantaged with an inflated level of performance.

Sources: Ministry of Education, Tertiary Education Commission.

Although the order of ranking of the universities was not affected by using the OQS as the unit of measurement, the ranking of some of the remaining TEOs was altered slightly. Of the larger TEOs, Wintec experienced the largest improvement in position, from 18th to 16th.

Comparing TEO performance using the weighted quality category score, OQS (unweighted) and DEA

In the calculation of the OQS measure, the higher weighting applied to research outputs (70 percent) compared with peer esteem (15 percent) and contribution to the research environment (15 percent), favoured those TEOs that performed relatively well in terms of their RO measure. There are two methods employed to compare how TEO relative performance was affected by these weightings. Firstly, each of the three performance measures is given equal weighting to derive an unweighted OQS measure. Then, Data Envelopment Analysis (DEA) is used to apply optimal weightings to the performance measures so that TEO performance can be seen in its best possible light⁹. This is a useful approach in that it removes the arbitrary nature of assigning weightings to the three performance measures¹⁰.

As shown in Figure 3, the unweighted OQS resulted in improved relative performance for some TEOs and diminished performance for others, when compared with their OQS. Providers that performed well in terms of their RO score, compared with their PE and CRE scores, were penalised by the equal weightings assigned to the three measures. Examples of these were Massey University, the University of Canterbury, AUT and Unitec. Universities such as Waikato and Otago improved their relative performance as a result of their relatively strong performance in their PE and CRE scores.

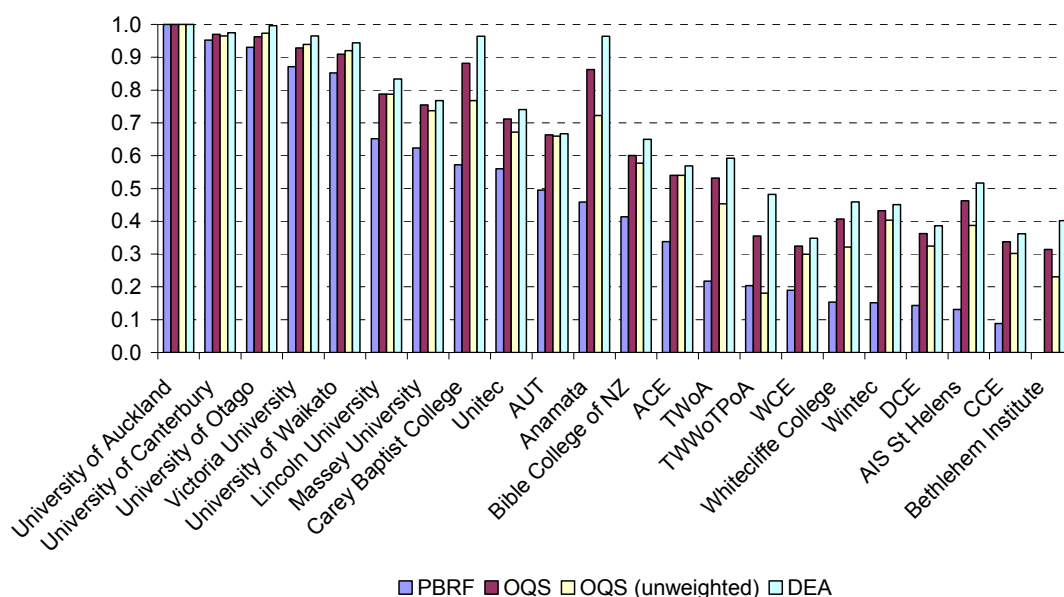
The use of DEA allows for an optimum measure of relative performance for the TEOs to be calculated. As can be seen in Figure 3, the relative performance of TEOs is at their highest level when using the DEA ratio, compared with the University of Auckland. This result is not surprising, in that, the DEA methodology provides the optimal weightings to each of the three performance measures and therefore TEO relative performance is seen in the best possible light.

A notable result was that the University of Otago improved its performance using this measure to the stage where it was only slightly behind the University of Auckland in terms of relative performance. As a result, Otago moved from 4th in the rankings to 2nd using this measure.

⁹ DEA is a non-parametric linear programming methodology useful in scenarios where there are multiple inputs and outputs.

¹⁰ It could be argued that different TEOs may place a higher importance on different performance measures. The DEA methodology allows for this difference in focus by assigning weightings that optimise the relative performance of the TEO.

Figure 3: TEO average weighted quality category score, OQS, OQS (unweighted) and DEA ratio, as a proportion of the University of Auckland score (panel assessed staff only)



Notes:

1. 'AUT' is the Auckland University of Technology, 'ACE' is the Auckland College of Education, 'TWVoTPoA' is the Te Whare Wānanga o Te Pihopatanga o Aotearoa, 'TWoA' is the Te Wānanga o Aotearoa, 'DCE' is the Dunedin College of Education, 'CCE' is the Christchurch College of Education and 'WCE' is the Wellington College of Education.
 2. 'PBRF' is calculated by dividing the average weighted quality category score per FTE for each TEO by the average score achieved by the University of Auckland.
 3. 'OQS' is calculated by dividing the average OQS per FTE for each TEO by the average OQS per FTE achieved by the University of Auckland.
 4. 'OQS (unweighted)' is calculated by dividing the combined RO, PE and CRE score per FTE for each TEO and dividing it by the combined score per FTE achieved by the University of Auckland.
 5. 'DEA' is the relative technical efficiency measure for each TEO.
 6. The relative performance of TEOs in this graph should not be compared with measures of performance that includes data for all PBRF eligible staff. TEOs with a small proportion of non-panel assessed staff will be advantaged with an inflated level of performance.
- Sources: Ministry of Education, Tertiary Education Commission.

5 Conclusion

The calculation of an unweighted average quality category score for TEOs showed that the actual variance in research performance of TEOs in the 2003 quality evaluation was smaller than was indicated by the published measure of performance.

The large variation in relative performance, as reported in the 2003 PBRF assessment, is therefore partly a reflection of the scoring system used, rather than a TEO's actual level of relative performance.

The desirability of this outcome depends on whether the high performing TEOs should receive the reporting benefits from the publication of a higher weighted score, in addition to higher levels of allocated funding. For stakeholders judging the relative performance of TEOs, it could be argued that reporting results with an unweighted measure would be more appropriate, as is used in the RAE in the United Kingdom.

Analysis of peer reviewed staff data showed that using the OQS as a performance measure resulted in a smaller variation in the relative performance of TEOs, compared with the weighted quality category score. Further analysis, that removed the weightings applied to the RO, PE and CRE measures, showed that some TEOs improved relative performance while others showed a decrease in relative performance.

DEA analysis showed that applying an optimal set of weightings for each TEO to the RO, PE and CRE scores, resulted in improved performance of all the TEOs, relative to the University of Auckland. In a number of cases, most notably the University of Otago, the ranking of TEOs was altered.

Although the approach of removing weightings or applying optimal weightings on the RO, PE and CRE scores is interesting from an analytical perspective, in that, it illustrates the sensitivity of the methodology to the weightings used, it must be remembered that the weightings applied to the three measures reflect the fact that research output is the key performance measure. Thus from a policy perspective the RO score warrants a high weighting.

Appendix

Table 1: FTE staffing, PBRF quality scores and DEA ratio by TEO

TEO	Panel assessed FTE	PBRF eligible FTE	PBRF quality score	Unweighted PBRF quality score	OQS	RO	PE	CRE	DEA ratio
University of Auckland	1,284.2	1,411.8	5,607	1,979	506,142	5,326	4,520	4,370	1
University of Canterbury	542.9	590.1	2,257	813	207,580	2,193	1,864	1,740	0.975
Victoria University of Wellington	516.8	579.3	1,966	722	189,136	1,973	1,754	1,645	0.964
University of Otago	932.5	1,174.9	3,788	1,370	353,896	3,695	3,189	3,161	0.996
University of Waikato	429.8	536.3	1,601	585	154,042	1,607	1,391	1,381	0.944
Lincoln University	175.5	195.3	500	194	54,484	573	458	498	0.834
Massey University	945.4	1,225.8	2,574	988	281,370	3,010	2,431	2,279	0.768
Auckland University of Technology	201.5	567.7	436	177	52,723	557	472	441	0.667
Unitec	100.2	345.8	245	100	28,108	308	243	194	0.741
Wintec	52.9	108.0	35	18	9,009	99	73	63	0.451
AIS St Helens	7.0	18.2	4	2	1,275	15	8	7	0.517
Anamata	1.0	2.0	2	1	340	4	2	2	0.964
Auckland College of Education	46.9	174.2	69	32	9,971	105	94	82	0.569
Bethlehem Institute of Education	1.8	17.2	0	0	223	3	1	1	0.402
Bible College of New Zealand	8.3	17.9	15	5	1,954	21	19	13	0.650
Carey Baptist College	4.0	8.2	10	4	1,390	16	10	8	0.964
Christchurch College of Education	87.9	170.4	34	16	11,688	132	86	75	0.362
Dunedin College of Education	28.7	66.4	18	9	4,099	46	28	29	0.386
Te Wānanga o Aotearoa	22.8	67.2	22	10	4,774	56	32	26	0.592
Te Whare Wānanga o Te Pihopatanga o Aotearoa	4.5	11.5	4	2	630	9	0	0	0.482
Wellington College of Education	17.3	109.7	14	2	2,214	25	19	14	0.348
Whitecliffe College of Arts and Design	10.5	16.3	7	3	1,687	20	10	7	0.459

Note: The quality scores are weighted on an FTE basis.

Sources: Ministry of Education, Tertiary Education Commission.

Bibliography

Ministry of Education (2005) *An analysis of funding allocations for staff and research degree completions in the Performance-Based Research Fund*, Tertiary Sector Performance Analysis and Reporting, Ministry of Education, Wellington.

Ministry of Education and Transition Tertiary Education Commission (2002) *Investing in Excellence, The Report of the Performance-Based Research Fund Working Group*, Wellington.

Tertiary Education Commission (2003) *Performance-Based Research Fund, Evaluating Research Excellence, the 2003 assessment*. Tertiary Education Commission, Wellington.