Broadening the Criteria: Lessons from the Australian Rankings

Ross Williams
Melbourne Institute
University of Melbourne

****

Paper Presented at the First International Conference on World Class Universities, Shanghai Jiao Tong University, June 16-18, 2005

1. Introduction

Why has the measurement of the international standing of universities become important? Why did the pioneering publication in 2003 of the ranking of world universities by the research team led by Professor Liu in the Institute of Higher Education at Shanghai Jiao Tong University (SJTU) create so much worldwide interest? The answer lies in the demand for such information arising from the rapid internationalisation of higher education and, on the supply side, the ready availability of data on the worldwide web.

Globalisation has meant an increased demand from students, employers, and academics for indicators of the international standing of universities. Decisions about where to study, whom to employ, or where to seek professional expertise should be based on quantitative and qualitative information but it is often difficult for the decision maker to obtain this information directly. Global companies recruit globally; international agencies seek expertise wherever it exists. Students intending to study abroad must choose from a vast array of universities, a decision for which much information is required.

A range of end users implies the need for a range of performance and reputational measures. It is essential in developing and using indicators of institutional performance that the target group of users be identified. Thus, for example, measures designed to assist in the choice of undergraduate studies would not be identical to measures designed to assist in the choice of an institution for doctoral studies.

In this paper we are concerned only with measures of international academic standing. We ask the question: How well does a university perform in the areas that create an international reputation?

The paper examines the issues involved in measuring international standing and illustrates them by reference to the study undertaken at the Melbourne Institute (see Williams and Van Dyke (2004, forthcoming)). The Melbourne Institute Index of the International Standing of Australian Universities is designed in a manner which permits extension to other countries. The Melbourne Institute project also provides information, based on a worldwide survey of academic leaders, on the weights that might be allocated to the different attributes that determine international standing.

2. Conceptual Issues in Measuring International Standing

Two levels of decision making are required in selecting the determinants of international standing: conceptual issues and the choice of data. There is reasonably broad agreement that it is the international standing of staff as measured by research influence that is the major determinant of international standing. However, a number of other attributes, such as quality of teaching and research training, need also to be considered. Broadening the range of attributes does, however, make the task of evaluation across national frontiers much more difficult.

In this section I examine a number of conceptual issues which must be faced in attempts to measure the international standing of universities.

2.1 Reputation versus current performance

International standing is a combination of current performance and reputation. The latter can be thought of as a combination of current and past performance. Reputation is in this sense the wider measure, but it lags current performance and also favours older institutions.

Reputation can be measured by surveys, but surveys of whom? The potential spectrum ranges from leading academic researchers and administrators, through beneficiaries of research, funding agencies including government, employers of graduates, and the general public. The reputation of an institution will more closely correlate to current performance the more informed is the respondent. For example, an academic who commands world-wide respect for her output is more likely to equate reputation with current performance than is a member of the general public. In order to more closely approximate current performance, surveys are best conducted amongst those who are relatively well informed.

Quantitative measures also have built-in time lags, which may lead to ratings not matching current performance. Publications and citations are typically over a number of past years. The use of data on Nobel Prize winners and such like poses particular problems of time lags.

2.2 Gross versus value-added measures

In looking at the international standing of universities the measures used are almost always gross measures of performance which make no allowance for factors such as the facilities and resources available to researchers or the nature of student intake. Conventional measures do not recognise that an institution may, for example, take in significant numbers of ill-prepared students and train them up to produce significant research. If institutions A and B produce similar research output but A has significantly poorer resources should not institution A be measured as performing better? Conceptually, at least, it would be possible to deflate research output by the dollar cost of producing that output.

I concur with the conventional the view that international standing is primarily achieved through gross measures of performance, which are largely unrelated to how that gross output was achieved. Even for student choice, gross measures are often appropriate. As Gormley and Weimar (1999, pp. 65-66) point out:

> Potential clients of organisations may sometimes care only about the gross levels of outcomes they achieve. Students considering MBA study, for example, may find rankings of programs based on reputation in the business world, such as those of Harvard University and Stanford University, derive much of their success from being able to recruit exceptional student bodies rather than from the additions they make to the capabilities of their students.

There is much to be said, however, for developing other rankings or ratings based on value added criteria to encourage and reward innovation and diversification in universities.

2.3 Size of institution

How to allow for size of institution poses similar issues to those raised by the 'gross versus value added' debate. If all that matters is the total international impact of an institutions activities then size is irrelevant. But by this measure, a very large institution can gain high international standing and still employ significant numbers of less able staff, a luxury not afforded smaller institutions. In other words, to achieve a given level of international standing, the smaller institution must have a higher quality staff overall.

What allowance should be made for size? In the Melbourne Institute work we treated this as an empirical question. We had two rankings of Australian universities: one based on surveys of CEOs (presidents) of the world's leading universities and Australian deans, and one based on quantitative measures of performance. If we assume survey measures are a subjective measure of international reputation and quantitative measures are a measure of actual performance it is possible to match the two measures to derive an estimate of the implicit weight that peers place on total performance versus performance

adjusted for size of institution. We did this by constructing a weighted average of performance in levels and another of performance per academic where the weights were obtained by maximising the rank correlation coefficient between this composite index and the index obtained from the survey of peers.[1] This yielded a weight of 0.84 on total performance and 0.16 on performance per academic. These weights are quite similar to the weights of 0.9 (levels) and 0.1 (size adjusted) used by SJTU. Total output does seem to be the major determinant of international standing.

2.4 Discipline mix

Readily available measures of research performance such as those compiled by the Thompson organisation (International Science Indicators, ISI, and the subset Essential Science Indicators, ESI) are biased towards output from the laboratory sciences. To a large extent this is merely a reflection of the fact that the results of scientific research are of interest across national frontiers whereas much research in the humanities is primarily of national interest. For this reason, an institution which is science intensive is more likely to have international standing than one which specialises in the humanities and social sciences.[2] Nevertheless, the measures themselves create discipline biases: the ISI data bank excludes books, for example.

The most obvious manifestation of the bias towards the sciences in the ISI data is the relatively lowly place occupied by the London School of Economics in the SJTU rankings, despite attempts to allow for its distinctive specialisation in the social sciences.

The ESI data base provides a useful indicator of the relative discipline strengths of an institution.[3] In table 1, citations over the period 1 January1994 – 31 December 2004 taken from this data base are presented for the top 25 universities in the SJTU 2004 ranking. A comparison of the ranking by total citations and by citations in business and the social sciences shows a number of major changes in the rankings: Chicago is ranked twenty-first in total citations but is fifth in the social sciences; Washington is ranked fourth in total citations but eleventh in the social sciences.

---

[1] In each case the measures included data from the six categories of performance as described in section 3.1. The only data which needed to be adjusted for size of institution was that relating to research performance. We followed usual practice in choosing measures of performance in teaching that are expressed in ratio form. Nevertheless, it might be argued that for these measures a large institution that scores the same as a small institution is really doing better. Thus if two institutions have common scores on variables such as entrance levels, progression rates and student evaluations, because the larger institution has a greater number of students at a given level or score then in some sense its total performance is superior to that of the smaller institution.

[2] van Raan (2005) also points out that international data banks are typically biased towards English language publications, which also favours the sciences (as well as more obviously universities in English speaking countries.)

[3] A feature of the ESI data bank is that it uses a threshold for inclusion in each of its 22 discipline categories. The threshold is that citations for an institution must represent at least 1 per cent of total citations.

Ratings based on research performance are also influenced substantially by whether the university has a clinical medical school. Data on research output in clinical medicine is influenced by the institutional arrangements which determine the affiliation given on publications, for example, whether or not only a hospital affiliation is given rather than a university affiliation. Citations in clinical medicine are also likely to be particularly favoured by teams of researchers cross-referencing colleagues.[4]

The importance of a medical school can be illustrated by looking at publications in the category "clinical medicine" in the ESI data base. For the top 25 universities in the 2004 SJTU rankings, citations to papers in clinical medicine represent 26 per cent of the total, but the percentage share ranges from 1 per cent at Caltech which has no medical school, to over a third for a number of institutions with strong medical schools (Harvard, Imperial College, Johns Hopkins, Toronto, Washington and UC San Francisco).

In table 1, the top 25 SJTU universities are re-ranked excluding citations in clinical medicine. As expected, because of the large share of citations in clinical medicine, the re-ranking produces significant changes: Berkeley, for example, with no clinical medical school moves up six places; Johns Hopkins with a strong medical school moves down six places.

The exclusion of citations in clinical medicine also results in a compression in the relativities across institutions. This is shown in table 2, where the list of universities is extended to include twelve universities from the Asia-Pacific region. In this table the citations counts are given as a percentage of the top-ranked institution (which for all three rankings is Harvard).

Alternative ways of dealing with data deficiencies in the nonSciences include:

- the introduction of additional internationally available measures of output in the nonSciences such as those provided by the Social Science Research Network (www.ssrn.com) which has commenced the ranking of universities in areas such as Law, Economics and Management;
- deflating science-biased measures of performance by a measure of the science intensity of the institution, e.g. deflating publications or citations in science by the number of academic researchers in science.

In the Melbourne Institute study we grouped the discipline categories in the ESI data base into "nonLab" (Economics and Business, Social Sciences (general)) and "Lab" (all other categories) and then gave weights to each which was a better reflection of the total output mix. We also used official national data on all refereed research output, including books.

[4] Katz (2000) shows that in the sciences citations increase more than proportionately to publications at both the institutional and national level with an exponent of around 1.25.

Table 1: Rankings by ESI citations, 1994-2004

Rank	Total citations	Citations excluding clinical medicine	Citations in business and social sciences
1	Harvard	Harvard	Harvard
2	Johns Hopkins	Stanford	Michigan
3	Stanford	Berkeley	UC Los Angeles
4	Washington	MIT	Pennsylvania
5	UC San Francisco	Tokyo	Chicago
6	UC Los Angeles	Washington	Stanford
7	UC San Diego	UC San Diego	Wisconsin
8	Michigan	Johns Hopkins	Berkeley
9	Berkeley	Cambridge	Columbia
10	Yale	Yale	Yale
11	Pennsylvania	UC Los Angeles	Washington
12	Tokyo	Michigan	MIT
13	Columbia	UC San Francisco	Cornell
14	MIT	Wisconsin	Johns Hopkins
15	Toronto	Columbia	Toronto
16	Cambridge	Pennsylvania	Princeton
17	Cornell	Cornell	UC San Diego
18	Wisconsin	Oxford	UC San Francisco
19	Oxford	Toronto	Oxford
20	Kyoto	Caltech	Univ College
21	Chicago	Kyoto	Cambridge
22	Univ College	Chicago	Caltech
23	Caltech	Univ College	Tokyo
24	Princeton	Princeton	Kyoto
25	Imperial College	Imperial College	Imperial College

Table 2: Scores on ESI citations (1994-2004): Selected Universities

Total citations		Citations excluding clinical medicine		Citations in business and social sciences	
100.0	Harvard	100.0	Harvard	100.0	Harvard
44.2	Johns Hopkins	47.9	Stanford	59.4	Michigan
41.7	Stanford	47.1	Berkeley	50.8	UC Los Angeles
40.4	Washington	42.0	MIT	49.8	Pennsylvania
37.7	UC San Francisco	41.7	Tokyo	49.5	Chicago
36.6	UC Los Angeles	40.7	Washington	46.5	Stanford
34.3	UC San Diego	40.2	UC San Diego	42.7	Wisconsin
33.8	Michigan	39.0	Johns Hopkins	42.6	Berkeley
32.6	Berkeley	38.8	Cambridge	41.8	Columbia
32.4	Yale	38.4	Yale	33.7	Yale
32.2	Pennsylvania	38.2	UC Los Angeles	32.1	Washington
31.8	Tokyo	35.7	Michigan	30.7	MIT
30.9	Columbia	35.2	UC San Francisco	23.8	Cornell
30.1	MIT	34.7	Wisconsin	22.4	Johns Hopkins
29.6	Toronto	34.4	Columbia	20.8	Toronto
28.4	Cambridge	34.3	Pennsylvania	19.4	Princeton
27.6	Cornell	32.9	Cornell	18.2	UC San Diego
27.6	Wisconsin	32.2	Oxford	16.0	LSE
25.5	Oxford	29.9	Toronto	15.5	UC San Francisco
21.8	Kyoto	28.5	Caltech	15.0	Oxford
20.4	Chicago	28.3	Kyoto	12.7	Univ College
19.6	Osaka	23.1	Chicago	12.0	Cambridge
19.3	Univ College	23.0	Osaka	8.9	Aust Nat Univ
18.7	Caltech	22.1	Univ College	5.7	Melbourne
13.8	Princeton	20.5	Princeton	5.4	Nat Univ Singapore
8.3	Melbourne	11.1	Aust Nat Univ	3.9	Caltech
7.7	Aust Nat Univ	9.5	Melbourne	1.3	Tokyo
4.9	Nat Univ Singapore	6.4	Nat Univ Singapore	1.1	Kyoto
4.8	Seoul National Univ	5.9	Seoul National Univ	0.9	National Taiwan Univ
4.5	National Taiwan Univ	4.6	National Taiwan Univ	0	Beijing
2.0	Indian Inst Science	3.0	Indian Inst Science	0	Fudan
1.7	Beijing	2.6	Beijing	0	Imperial College
1.4	Tsing Hua	2.1	Tsing Hua	0	Indian Inst Science
1.0	Fudan	1.5	Fudan	0	Osaka
0.5	LSE	0.8	LSE	0	Tsing Hua
0.3	SJTU	0.5	SJTU	0	Seoul National Univ
0.3	Imperial College	0.0	Imperial College	0	SJTU

2.5 Weighting Attributes

In order to produce a ranking of institutions different attributes must be combined into a single measure. This requires weights, the choice of which is subjective. The final ranking is less sensitive to weights when the criteria used are relatively homogeneous. Thus rankings based solely on research performance (output, citations etc) will be relatively insensitive to the choice of weights owing to the high correlation between various measures. (The exception in the SJTU rankings would be the weights given to Nobel Prize winners.)

If a range of measures are used, then the choice of weights becomes critical. In constructing the Melbourne Institute Index of the International Standing of Australian Universities we asked respondents to place weights on our groups of attributes thus removing subjective choice by the researchers. (See section 3 for details.)

2.6 Choice of Data

In choosing data to measure concepts, the following criteria should be used: the data should (i) come either from international data banks or conceptually have their counterparts in other countries, (ii) have been collected on a consistent basis by an external agency and (iii) avoid undue complexity.

3. Determinants of International Standing

3.1. Six groups of attributes

We have found it convenient to group the attributes that determine international standing under six headings: quality/international standing of academic staff, quality of graduate programs, quality of undergraduate intake, quality of undergraduate programs, resource levels, and opinions of stakeholders. We recognise, however, that there is some overlap between these groupings. It is not possible to have a good graduate program without high quality academic staff, although the affiliation of leading researchers with an institution does not guarantee a high quality graduate program.

Some attributes of a university, such as location and extracurricular options, which make it attractive to undergraduate students are not particularly relevant for international standing.

I now discuss possible measures of international standing under our six headings. For some components care must be taken to translate national measures into international comparable measures and this matter is also addressed. The measures and issues are illustrated by reference to our study of Australian universities (further details are given in the Appendix).

(i) Quality/International Standing of Academic Staff:

The variables commonly used to measure academic staff quality are as follows: the quantity and quality of research publications; citations to the research output; research income, especially from competitive grants; recognition of academic standing through election to academies, receipt of prestigious awards, or invitations to international conferences; status of degrees held by academic staff. These measures broadly translate across national frontiers, although there is a need to standardise membership of national academies, possibly by linking it back to national contribution to publications or citations.

In our study of Australian universities we preferred to use research output over degrees held as it is a better measure of current performance. Our data sources were ESI, academy lists, and government statistics on research performance. The various measures of current research performance were quite highly correlated but membership of academies was less correlated with current performance.

(ii)*Quality of Graduate Programs*

Measures of the quality of graduate programs include the degree of selectivity in intake, number of completions, progression rates, placement on graduation, and student evaluation. In order to make international comparisons, emphasis should be placed on the PhD program as it is recognised internationally.

In Australia there is available on a national basis the results of a course experience questionnaire sent to all students on graduation. This is not available in many countries.

(iii) Quality of Undergraduate Intake

Quality of intake falls between a reputation measure and a performance measure. While the lags can be of varying length, a poor undergraduate program will in time be reflected in a decline in the quality of intake. Measures typically used are scores on national entrance examinations and acceptance rates. Geographic diversity of students is also sometimes used as measure of standing. Acceptance rates can suffer from the limitation, as in Australia, that students tend to apply only for courses that they believe they have a good chance of entering. In our study we were able to use nationally standardised tertiary entrance scores.

(iv) Quality of Undergraduate Programs

The most common measures of the quality of undergraduate programs are progression rates, outcomes on graduation such as employment or progression to a higher degree, awards won or honours grades achieved, student evaluations, diversity of the student body, and class size. The weakness of class size is that it assumes a given technology of teaching. The ratio of students to academic staff is a more neutral measure, although

with the increased role of teaching through the internet the contribution of technical staff is also relevant. We reject unemployment or 'not employed' as an appropriate negative indicator of success. It is subject to regional and national variations in employment rates, and overlooks the increased tendency for students to take time off between study and work. Honours grades and awards do not readily translate internationally.

In our evaluation of Australian universities, wherever data permitted, we disaggregated by discipline into "Lab" and "NonLab" and used separate variables for domestic and foreign students.

(v) Resource Levels

These can be measured in physical units or monetary amounts. The use of physical units facilitates international comparisons. Examples used in previous studies are support staff, library volumes and IT facilities. With technological developments, however, it is increasingly the enabling ability which matters in retrieving information. Monetary measures include total revenue per student or staff member, and expenditure on IT and libraries. Alumni giving and salary levels are sometimes included as measures of standing. Monetary variables, unless expressed in ratio form, suffer from purchasing power problems when comparing across countries.

Data limitations restricted the variables used in our Australian study to revenue per student and revenue per academic staff member. Consistent with our view that gross measures are the relevant ones we equate high levels of resources with better teaching and research training, and a measure of standing.

(vi) Subjective Assessment: In a number of rankings of universities and disciplines, people are asked to rate universities using various criteria. If the survey is international in scope, it provides a way of linking national rankings into world rankings.

In the Melbourne Institute study we surveyed CEOs (presidents) of the world's leading universities and Australian deans. Respondents were asked to rate each Australian university relative to universities in the respondents own continent, with provision in the questionnaire for "not well enough known". As expected, response rates were higher from institutions with historical connections to Australian universities.

3.2 Weighting the groups

In our study we asked CEOs of the world's leading universities as judged by SJTU, plus deans in Australian universities, to place weights on our six groups of attributes. Useable replies were received from 40 foreign CEOs and 80 deans.[5] An unexpected finding was that foreign CEOs and domestic deans placed very similar weights on the attributes. The international standing of staff was given a weight of 40 per cent and the other five characteristics shared the remaining 60 per cent fairly evenly. The results are given in table 3.

Table 3: Mean responses to question, "If you were to evaluate the international standing of a university, what percentage weight would you place on each of the following?"

Attribute	**Foreign CEOs**	**Australian Deans**	**Chosen weights**
Quality of Academic Staff	39.8%	39.5%	40%
Quality of Graduate Programs	17.1%	14.1%	16%
Quality of Undergraduate Intake	10.4%	11.8%	11%
Quality of Undergraduate Programs	13.4%	14.3%	14%
Resource Levels	10.7%	12.3%	11%
Peer opinion	8.7%	8.0%	8%

[5] We aimed to achieve around 40-50 replies from foreign CEOs. Preliminary investigation indicated that a low response rate could be expected. Because of the pressures on CEOs time we decided to over-sample rather than to follow up nonrespondents. The sample size of 172 was arrived at by including the following universities: those included in the top 100 universities in the SJTU study for 2003 or 2004, those ranked 101-152 in the 2004 SJTU study, UK and Canadian universities ranked 153-201in the 2004 study, all members of the Russell Group of research universities in the UK, and, in order to increase country representation from the Asia-Pacific region, all non-Japanese universities (which are well represented under other criteria) included in the top 300 in either of the two Shanghai studies, plus all eight New Zealand Universities. Not all of these categories are mutually exclusive. About 50 per cent of replies were from CEOs of European universities. Questionnaires were sent to 200 deans.

4. Comparisons of Ranking of Australian Universities.

The rankings of the top 15 Australian universities by three methodologies (Melbourne Institute, SJTU and Times Higher Education Supplement (THES)) are given in table 4. In the Melbourne Institute study, the highest-ranked fourteen universities are the same fourteen Australian universities (though with some differences in ordering) as appear in the 2004 SJTU list of the top 500 world universities (see table 5). These remarkably similar results arise from quite different methodologies. The SJTU study is based almost entirely on what we term the international standing of academic staff as measured by research performance. The only common data is that provided by ISI which comprises only 16 per cent of our index. The similarity in results is due to three factors: (i) in our study the variation between universities is much greater for research standing than for other attributes, (ii) the measurement of research performance is relatively insensitive to the precise variables chosen, and (iii) both studies adjust for institutional size in a similar manner.

The results from our survey of foreign CEOs are also quite consistent with the 2004 SJTU rankings: about 50 per cent placed ANU and Melbourne in the top 80 institutions in the world[6] a clear majority placed Sydney, Queensland and New South Wales in the top 200, and about 50 per cent placed Monash in this category.

The rankings from the THES differ significantly from ours and those of SJTU. The THES rankings give a weight of 50 per cent to peer review and 25 per cent to quality of academic staff (as measured by citations per academic staff member and international character of academic staff).

5. The Way Forward

If terms of conceptual difficulty the spectrum ranges from, at the easiest end, ranking performance by discipline across universities in a given country to, at the most difficult end, ranking universities internationally considered as a single entity. Ranking by discipline does have its place, but in recent years the growth in interdisciplinary research and quality controls on departments has meant that the variability in the quality of departments within an institution has fallen. Bad departments in good universities are becoming rarer.

Doubts are expressed by some about the usefulness of trying to rank universities internationally. I take the view that there exists what might be called a university brand-name effect, the value of which can, at least conceptually, be quantified. Universities themselves certainly act through their media and public relations departments as though their brand name is important to them.

[6] The exact figures were 53 per cent for ANU and 48 per cent for Melbourne.

Table 4: Comparison of Rankings: THES, SJTU and Melbourne Institute

THES	World Rank	Australian Rank	SJTU	World Rank	Australian Rank	Melbourne Institute	Australian Rank
Australian National University	16	1	Australian National University	53	1	Australian Natitional University	1
University of Melbourne	22	2	University of Melbourne	82	2	University of Melbourne	1
Monash University	33	3	University of Queensland	101-152	3	University of Sydney	3
University of New South Wales	36	4	University of Sydney	101-152	3	University of Queensland	4
University of Sydney	40	5	University of New South Wales	153-201	5	University of New South Wales	5
University of Queensland	49	6	University of Western Australia	153-201	5	Monash University	6
RMIT University	55	7	Monash University	202-301	7	University of Western Australia	6
University of Adelaide	56	8	University of Adelaide	202-301	7	University of Adelaide	8
Macquarie University	68	9	Macquarie University	302-403	9	Flinders University of South Australia	9
Curtin University	76	10	University of Newcastle	302-403	9	La Trobe University	10
University of Western Australia	96	11	University of Tasmania	302-403	9	Macquarie University	11
University of Technology, Sydney	113	12	Flinders University of South Australia	404-502	12	University of Tasmania	12
La Trobe University	142	12	La Trobe University	404-502	12	University of Newcastle	13
University of Tasmania	161	14	Murdoch University	404-502	12	Murdoch University	14

Sources: Times Higher Education Supplement (THES), www.thes.co.uk
Shanghai Jiao Tong University (SJTU): http://ed.sjtu.edu.cn/ranking.htm
Melbourne Institute:Williams and Van Dyke (2004, forthcoming)

Nevertheless, the discipline mix of a university does influence international rankings significantly. If rankings are defined as international standing or international recognition, then it is inevitable that some disciplines will be more important than others. But in an age of globalisation, this is not to say that researchers in most areas cannot obtain an international influence. It is not only the sciences where a deep international market in ideas exists. Disciplines such as finance, economics, and philosophy are international in nature. Even in the humanities, great scholarship tends to be recognised wherever it occurs. In my own university we have an international renowned art scholar in the history of Italian renaissance painting, and another in nineteenth century French history. It is only where institutional factors dominate, such as in areas of law, where international standing and domestic standing may be in conflict.

The greatest limitation to measuring international standing in the humanities and social sciences is the absence of suitable data bases. The publication and citation of monographs, including their evaluation by reviewers, is probably the greatest omission from readily available data banks.

A number of issues discussed in section 2 come down to the issue of how do we compare *extreme* excellence in one area, whether it be in clinical medicine or in social sciences or in some other area, with excellence across most fields of study? A university can achieve international fame on the basis of top ranking in only one or two fields. One solution is to construct alternative rankings by broad discipline fields, the broadest being a three-fold division into (i) clinical medicine, (ii) science, and (iii) humanities and social sciences. These three sets of rankings could then be both calculated separately and also combined using an appropriate set of weights, which might include giving a weight to humanities and social sciences which attempts to overcome current data deficiencies in these areas. This approach would serve the dual purpose of providing a broad discipline ranking and on overall ranking.

Appendix: Variables and Weights used in Measuring the International Standing of Australian Universities

The period of coverage of the ESI data is January 1994 to February 2004. The ISI data for highly cited researchers covers the period 1981-1999.

The shorthand expression *NonLab* is used to refer to performance in the social sciences, business and the humanities; the expression *Lab* refers to performance in science, engineering, medicine and related areas.

Within each category the subcategory weights are stated (they sum to one). In the empirical work, the results for each subcategory are standardised with the score for the highest performing institution set at 100. The scores for each subcategory are then added using the weights given below and the results rescaled such that the score for the highest performing institution is set at 100. The process is repeated using the scores out of 100 for the six categories to obtain an overall index.

1. Quality and International Standing of Staff (40 per cent)

(a) Measures of total performance

- publications (0.25): comprising Lab ESI articles (0.10), NonLab ESI articles (0.05), and Department of Education, Science and Technology (DEST) audited publications, average 2001 and 2002 (0.10)
- citations (0.25): comprising Lab ESI (0.17), NonLab ESI (0.08)
- peer recognition (0.25): Academy membership at June 2004 (0.20), ISI highly cited researcher (0.05)
- research income, average 2002-2003 (0.25): National Competitive Grants (0.15), total research income (0.10)

(b) Measures of performance scaled for size

- publications (0.25): DEST scores per academic staff member (0.10), ESI count per head (0.15)
- citations per article (0.25): Lab ESI (0.17), NonLab ESI (0.08)
- peer recognition (0.25): Academy membership per academic staff member (0.20), highly cited researchers per head (0.05)
- research income (0.25): National Competitive Grants per academic (0.15), total income per academic (0.10)

2. Quality of Graduate Program (16 per cent)

- PhD Completions, average 2001-2002 (0.35)
- postgraduate completion rates, 2002 (0.30): domestic students (0.15), foreign students (0.15)
- student evaluation of their PhD program, average 2000, 2002, 2003 (0.35)

3. Quality of Undergraduate Intake (11 per cent)

Measure uses average 2002 and 2003 data

- Median Tertiary Entrance Score

4. Quality of Undergraduate Programs (14 per cent)

Measures use data for 2002

- progression rates (0.25): domestic students(0.125), foreign students (0.125)
- ratio of academic staff to students (0.25): Lab (0.125), NonLab (0.125)
- student evaluation of course on completion (0.25): Lab (0.125), NonLab (0.125)
- percentage of students graduating with a pass or honours bachelors degree who in the next year are enrolled in a higher degree (0.25): masters by coursework (0.10), honours degree, masters by research or PhD (Lab 0.075, NonLab 0.075)

5. Resource Levels (11 per cent)

Measures of revenue, average 2001-2002

- revenue per academic member of staff (0.75)
- revenue per student (0.25)

6. Peer Opinion (8 per cent)

Survey undertaken in 2004

- ratings by CEOs (presidents) of foreign universities (0.50)
- ratings by Australian deans (0.25)
- rankings by Australian deans and New Zealand vice-chancellors (0.25)

References

Gormley Jr, William T. and Weimer David L. (1999). *Organisational Report Cards.* Cambridge, Mass: Harvard University Press.

Katz J. Sylvan (2000), 'Scale-independent Indicators and Research Evaluation', *Science and Science Policy*, 27(1), 23-36.

van Raan, Anthony F.J. (2005). 'Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities by Bibliometric Methods', *Scientometrics*, 62(1), 133-143.

Williams, Ross and Nina Van Dyke (2004), *The International Standing of Australian Universities*, Melbourne Institute, University of Melbourne (www.melboureinstitute.com)

Williams, Ross and Nina Van Dyke (forthcoming), 'Measuring the International Standing of Universities with an Application to Australian Universities', *Higher Education.*