

**AUSTRALIAN
UNIVERSITIES QUALITY
AGENCY**

**Ranking of
Higher Education Institutions**

August 2006

Dr Antony Stella and Dr David Woodhouse

AUQA Occasional Publications Number 6

ISSN 1446-4268

ISBN 1 877090 57 3

© 2006 Australian Universities Quality Agency

Level 10, 123 Lonsdale Street

Melbourne, VIC 3000

Ph 03 9664 1000

Fax 03 9639 7377

admin@auqa.edu.au

www.auqa.edu.au

The Australian Universities Quality Agency receives financial support from the Commonwealth, State and Territory Governments of Australia.

PREFACE

This analysis, *Ranking of Higher Education Institutions*, was originally written in March 2006 for the AUQA Board, which asked for an overview of ranking systems in use worldwide and the issues surrounding their use. Subsequently, AUQA decided to make a slightly edited and updated version publicly available, as we felt the paper would be of general interest to those working in quality assurance in higher education both within Australia and internationally.

Rankings of higher education institutions are attractive to consumers, students, benefactors, alumni and governments: they promise to reduce uncertainty for those who have to make choices in an environment where full information is not available and quality cannot be equated to price. The number of organisations that issue rankings is now substantial enough to have formed its own expert group and to be preparing advice to UNESCO.

As well, academic articles on higher education rankings are now starting to appear. Some of these articles are largely descriptive, while others are more analytical or critical. This paper begins with an outline of major rankings, or 'league tables' produced by print media organisations, focusing primarily on English-language media. It then considers reactions to rankings published by these media organisations and a range of criticisms which have been levelled at them. The next sections of the paper consider some of the rankings and ratings used in academia, including those published by particular organisations, such as the Academic Ranking of World Universities by Shanghai Jiao Tong University, and the criticisms made of these rankings.

Concluding sections consider the users of rankings, the potential for misinformation, and actions that can be taken to encourage the use by consumers and other external parties of relevant, appropriate measures of quality outcomes. If, as seems likely, ranking systems are here to stay, there are opportunities to guide both their evolution and the capacity of end-users to interpret and moderate the information they are receiving. The authors conclude that the most urgent task is to minimise the harm that results from irrelevant or inappropriate measurement. And, we must keep in mind that any measure can only be an aggregate indicator of a wide and varying reality.

In summary, the paper provides an overview of the state of higher education rankings in mid-2006, the choices which are available to various actors and encouragement to use rankings wisely.

Jeanette Baird
Series Editor
AUQA Occasional Publications

1. Introduction

Measuring institutional quality is gaining prominence in higher education due to the interplay between many factors. Some of the factors that trigger this interest among stakeholders are: shrinking resource allocation for higher education from public funds, increasing competition among higher education institutions (HEIs) and growing awareness about value for money among the public. The stakeholders pay attention to institutional quality with one or more interests: families for worth of personal investment in education; governments for accountability and policy-making; funding agencies for funding decisions; society for value of tax payers' money; industry for institution–industry partnership; and employers for graduate recruitment. To understand institutional quality, these stakeholders might turn to various sources of data. Rankings based on certain institutional aspects are data some stakeholders might be interested in.

National and international rankings are gaining attention and popularity, and not necessarily for academic reasons. Media, academic groups and the governments have indulged in various ranking efforts for various reasons. Media rankings are guided by 'what sells in the market' rather than the rigorous quality assurance practices of academic bodies. Similarly, there have been instances when governments have taken efforts akin to ranking for specific purposes and the scope of those efforts has been very limited. However, in recent times there is a growing tendency in the academic discussions to misinterpret 'public popularity' as 'academic credibility' and 'limited scope' as 'need of the hour' that misguides many to believe that ranking efforts have to be considered seriously.

To keep a check on this tendency a few questions have to be raised: How sound are the methods used? How reliable are the data available to groups that do rankings? Do rankings reflect institutional quality or an aspect(s) of it adequately? How intensive would efforts have to be to ensure the rigour of the ranking process? Would the efforts match the benefits? Who are the stakeholders who would benefit from the rankings? How competent are the stakeholders in appropriate usage of the available ranking data? How have the stakeholders responded to rankings so far? Is there a better way to promote meaningful use of data on institutional quality?

This paper argues that there are serious problems with most if not all of the whole-of-institution linear rankings. Some of the same problems pertain to discipline or program rankings.

2. Current interest

In June 2002, the first international meeting on ranking systems was held in Poland. There were 40 participants from 12 countries, representing publications that publish rankings. In December 2004, there was a meeting in Washington

DC of twenty-five ‘experts’ on the production of rankings. Those present at this meeting created an International Rankings Expert Group (IREG) to pursue work on this topic. IREG will advise UNESCO-CEPES and the Institute for HE Policy (IHEP) on useful tasks in this area (Merisotis & Sadlak 2005). The journal *Higher Education in Europe* had a special issue (vol. 30, no. 2) in July 2005 drawing on the results of the Washington meeting.

The IREG met again in Berlin in May 2006 to produce ‘Criteria and Standards for International University Ranking’. The meeting resulted in developing the Berlin Principles on Ranking of Higher Education Institutions. The IREG, in its Berlin Principles, lists many of the problems with rankings, however, our fascination with league tables is such that they couldn’t bring themselves to criticise rankings, but said ‘rankings are all right if you avoid these problems’. It is rather as if a set of warnings about the dangers of smoking were converted to a document that said ‘it’s all right to smoke provided you avoid these dangers’!

3. Rankings and league tables by media

Rankings and league tables use indicators or characteristics of institutions and present comparative information for members of the public. Some tables are basically descriptive while others use a variety of quantitative and semi-quantitative indicators to compare institutional performance in teaching, research, the provision of resources and/or student support. Most league tables use weighted combinations of scores. The league tables employ various methods of collecting data. In general, the *US News and World Report (USNWR)*, *The Times Higher Education Supplement (THES)*, *Asiaweek*, and the *Australian Good Universities Guide (GUG)*, each use(d) previously published data readily available in the public domain supplemented, in some cases, by university surveys.

In the United States, academic rankings first appeared in the 1870s. In 1911 the Bureau of Education published a rating of 344 institutions. The ranking group, objective, methodology and the audience changed over the years and by the 1980s it garnered mass appeal. Since their first appearance in 1983, the *US News and World Report* rankings of colleges and graduate schools have received a lot of public attention. The *USNWR* annually publishes ‘America’s best graduate schools’ based on six major indicators and ten sub-variables that include academic reputation, student selectivity, faculty resources, financial resources, retention rate, and alumni satisfaction. Rankings have been categorised by subject area — Business, Law, Medicine and Health, Education, Engineering, Library Science and PhDs. Included is a directory of over 1000 graduate programs by subject and state, the methodology of their rankings and an index. *Money* magazine’s ‘Best College Values’ is also popular in the United States.

The *Sunday Times* first published its rank of United Kingdom universities in 2001. The UK *THES* uses data from:

- Higher Education Statistics Agency
- Higher Education Funding Council
- Quality Assurance Agency
- Office for Standards in Education, and
- individual university surveys.

THES collects data on ten indicators that include entry grades, student:staff ratio, accommodation, completion rates, number of students who get first class, value addition, library spending, postgraduate student numbers, and employment.

The Australian *Good Universities Guide* uses data from five main sources:

- Department of Education, Science and Training
- Graduate Careers Australia
- tertiary admissions centres in each state
- various other national data sets, and
- individual university surveys.

The *GUG* collects data on 16 indicators that include status and standing, entry-level profile, international activities, teaching and courses, employment and student population.

In Canada, the popular weekly magazine *Macleans* collects data on 22 indicators by sending out a detailed questionnaire for the universities to fill out and return. Data includes student body, classes, faculty, finances, library and reputation. *Macleans* first published its ranking in 1991.

The *Gourman Report of Graduate Programs* (1997) ranks top graduate and professional programs in over 100 academic areas. Separate sections cover schools of law, medicine, and health-related professions. Also included are lists of 'approved' engineering and business management schools, a rating of US research libraries and overall rankings of US and international graduate schools. *Gourman Report of Undergraduate Programs* (c.1989) ranks undergraduate programs in over 100 individual disciplines, as well as the top universities in the broader realm of pre-legal and pre-medical education. It includes a section on university administrative areas (e.g. libraries, alumni associations) and international universities. The *Gourman* reports have been being published annually since 1967.

While most rankings are for a national audience, recently there have been developments in international ranking. The magazine *Asiaweek* published its report 'Asia's best universities' from 1997 to 2000. *Asiaweek* used questionnaires sent out to universities as well as a variety of officially published data to complete its survey. It used data on five indicators: academic reputation, student selectivity, faculty resources, research output and financial resources.

The magazine ceased publication in 2001. In Europe, in April 1998, the magazine *Der Spiegel* published a ranking of European universities in various disciplines which was severely criticised by the Association of Universities of Netherlands (VSNU).

THES published its first 'World University Rankings' in 2004 based on five factors — survey of 1,300 academics in 88 countries (50%), percentage of overseas staff (5%), percentage of overseas students (5%), student:staff ratio (20%), citations per staff (20%). *THES* asked academics to name the top institutions in the areas and subjects on which they felt able to make an informed judgement.

4. Response to media ranking

Public

In general, media ranking has become popular among the general public. The rankings are among the top issues of the media in terms of sales generated. For example, the annual 'best universities' issue was one of the year's biggest sellers for *Asiaweek*. An estimate by *Times* magazine in 1998 revealed that prospective college students and their parents spend \$400 million per year on college-prep products, which includes ranking publications. Proponents claim that the rankings provide some insight for secondary school leaving students facing the daunting task of choosing an educational institution for post-secondary studies. Providing data that has not previously been accessible in the public domain is perceived as one of the positive aspects of the rankings. For example, in the case of *Macleans*, the data collected is not readily available anywhere else as the magazine goes directly to the universities for the information. However, many studies have also revealed that the rankings do not go beyond 'popularity' and 'media hyperbole' and that the public do not base decisions solely on media rankings.

Academics

The reliability of media rankings has been criticised by academic groups. Most of the media rankings attempt to get an overall view about the quality of the HEIs through opinion polls and surveys, and this method has come under severe criticism by academic groups. In some cases, particularly in the United States, inaccuracies in the data used for media ranking have been observed. Inconsistencies between the information given by the institutions to ranking groups and to other agencies have been well documented. Critics assert that the HEIs manipulate the data in order to achieve a higher rank (Dill & Soo 2004). Consequently, academics are concerned about the way the media rankings may be used by stakeholders or misinterpreted by the uninformed.

Institutions

Institutional reaction to media rankings has been predictably mixed, with some supporting their use, others opposing it, and most urging caution. From university boycotts to intense scrutiny by academics there has always been strong opinion on each side of the debate. Provan and Abercromby (2000) point out that universities feel compelled to contribute to rankings and league tables possibly because it affords them a large amount of free publicity and may have an effect on a university's reputation:

Universities value international rankings such as the *THES*'s survey of the world's top universities, and the Shanghai Jiao Tong University's top 100. A good ranking gets a lot of free overseas publicity. (Yerbury 2006)

Institutions unhappy with the method or with the general premise of league tables may choose to boycott the entire process. In Canada, Australia and the UK, boycotting of the rankings has been sporadic to non-existent. In 1999, 35 universities refused to participate in the *Asiaweek* survey. Nineteen of these institutions came from China, and the University of Tokyo (top of the league in 1997 and 1998) was the only university not to participate from Japan. President Hasumi Shigehiko of the University of Tokyo explained the university's rationale for not joining the survey:

The quality (of our) education and research cannot be compared with that of other universities. Such characteristics are profoundly individual and extremely difficult to quantify. (*Asiaweek* 1999, quoted in Provan & Abercromby 2000)

According to *Asiaweek* the high level of non-participation from Chinese institutions was politically motivated.

Even those who claim to oppose rankings by media often only pay lip service to anti-ranking campaigns while figuring out how to boost their own positions on the list. In New Zealand, when Victoria University of Wellington went down to number 68 from 38 in media ranking, it chose to include the improvement of its *Asiaweek* ranking among its 'top strategic goals' for the coming decade. Similarly, the Institut Teknologi Bandung in Indonesia has set its appearance in the *THES* ranking as a goal (though the VC of the University of Malaysia seemed refreshingly unmoved by a fall in the University's position). Institutions advertise favourable ranking positions in marketing and promotional materials (for example: Swinburne University of Technology's display advertisements in December 2005 referred to Swinburne's fifth position on the Learning and Teaching Performance Fund 'ranking'; Wollongong's display of its five-star rating in *GUG* on the University's campus sign). Some institutional heads have viewed media ranking as a useful comparison on the

premise that the media rankings could be extremely good for universities at the top in terms of pressing governments for funding.

5. Criticisms of media rankings

The general criticism about the media ranking relates to the methodology that addresses quality in a superficial way but projects a complex image. Some categorise institutions and publish rankings for each category. Some assign weights to indicators and use the weighted scores to rank the institutions. The indicators considered for rankings and the methods followed to arrive at the scores on the indicators vary.

The variety of methodologies, and thus of criteria used, suggest that any single objective ranking could not exist. (Rocki 2005)

The public tend to think that because we can have football league tables, we can do the same for institutions. However, football league tables describe a limited form of activity, and only after each team has competed against every other one at this one activity.

Many academics, including those from highly ranked institutions, question both the data and the processes used by some of the ranking services. The standard market research techniques followed in data collection have raised the question: Can we assume that what works for pet food, perfume and pesticide will also work for education? Most rankings rely on two types of data — data given by institutions that is accepted, often without a reliable validation process, and data obtained from opinion polls in the name of ‘expert opinion’. With both components on shaky grounds, the use by the media groups of complex formulae with weights and indicators only helps to project a pseudo-image of being ‘scientific’ to outcomes that may be statistically irrelevant. In particular, the following areas are of concern.

Choice of indicators

Academics believe that the rankings do not take into account the important qualities of an educational institution that cannot be measured with weightings and numbers:

Seven of the 10 [rankings] do not include an indicator for teaching quality. ... Obtaining independent, objective measures of teaching quality is difficult, expensive and time-consuming. (Van Dyke 2005)

In other words, rankings are largely based on what can be measured rather than what is relevant and important. Conversely, reputation is used too often as a measure of academic quality in media ranking. Some media rankings give as

much as 50% weight to reputation which is a very subjective and inaccurate measure to denote quality.

Arbitrariness in giving weightings

The arbitrariness in the way weights are allotted is not explained in the media rankings: *UNSWR* gives 16% weight to graduation rate, while *Macleans* gives 2%. To give another example, *THES* allotted 5% for the proportion of academic staff recruited from overseas and a 5% score for proportion of international students, with no clear rationale. Meaning is further undermined by the arbitrary allocation of a score of 49 out of 100 for all institutions (e.g. Australian institutions) where academic staff recruited from overseas are not systematically reported. This supports criticisms that *THES* ranking is not a good indicator of the quality of Australian universities. Leslie Wagner, then Vice-Chancellor of Leeds Metropolitan University, was highly critical of the British University league tables:

The Times by its choice of measures provides a particular view of the role of universities and by categorising those who conform to that role as good and those who do not as less good, it acts as a corrosive force in the development of higher education. The tables reinforce an elitist view of the purposes of higher education and weaken attempts at diversity. (CVCP News 1998, quoted in Provan & Abercromby 2000)

Marginson (2006) observes that *THES* ranking ‘is a shabby survey and perceived as such.’

The general argument is that choice of weights is subjective and arbitrary, with little or no theoretical or empirical basis. The difficulty, of course, is how to report results without assigning weights, since the various measures cannot then be combined into any overall ranking or clustering. (Van Dyke 2005)

The honest and rational conclusion from this would be that any overall ranking or clustering is meaningless.

Formula change

Critics argue that yearly formula changes indicate a dangerous tendency to exploit the public. The *US News* produced its first national rankings in 1983, basing its results on the composite scores of an opinion survey of college and university presidents. Over the next two decades, the magazine continuously refined and made more complex its methodology, adding quantifiable variables and shifting the weights assigned to items. Every year, the resulting methodology, by its very construct, resulted in wide swings in the rankings

even without a significant change in the quality of the institutions. It became almost impossible to interpret shifts in an institution's rank in terms of change in its relative academic quality. A college that is ranked fourth one year and seventh the next may have had no change in its performance relative to other institutions, yet may still have moved because of changes in the ranking method.

Changing the method and giving new rankings are seen as market ploys by the media to sell the publications. Jumbling the rankings in this way leads to increased sales of the ratings publication, compared to a listing that remains relatively static from year to year. For example, *India Today* in its 2004 ranking declared that a host of institutions made a comeback on the 2004 national list pushing aside previous leaders. It considered that the moral, if there was one, was that the level of competition among these institutions was so high that there was no room for ordinary performances, even for one year. However, it was not clear how expert opinion could have captured institutional developments within such a short period. According to *India Today*, staying on top, as the eight years of polling has shown, is never an easy process. For those who did not figure in that year's roll of honour, it is a comfort to know that there is always next year.

Reliance on opinion polls

The dependence on opinion polls and surveys invites strong criticisms from the thinking community. For example, the World University Rankings of *THES* published in 2004 considered five factors and for one of the factors — a survey of 1,300 academics in 88 countries — it allotted 50% weight. *THES* asked academics to name the top institutions in the areas and subjects on which they felt able to make an informed judgement. It does not say what 'top' might be. *THES* may claim that experts name top institutions in their field only, but we cannot be sure how many of those experts really knew anything about the quality of the institutions at that point in time.

'Expert opinions' suffer from two major flaws. Firstly, they suffer from distortion in perception and the halo-effect. One department's reputation that the expert is familiar with may indiscriminately influence the rating of the whole institution. Marginson (2006) reports that, 'One US survey found the Princeton Law School was ranked seventh in the country. But Princeton did not have a law school. It was created by the halo effect!'

The other flaw is more subtle and occurs due to divergence and rapid expansion of knowledge. Experts may be biased or uninformed about all the institutions they are rating. Today a scholar no longer knows enough about the work of scholars in other disciplines at other institutions to rate their quality accurately. An expert's opinion on the current quality of an institution may not be anywhere close to reality. The assumption that the experts selected are the ones who know enough about the quality of the institutions at the national and

international level is a disaster. How many institutions beyond our own do we know well enough to rate? Even in a relatively small field, where experts are more likely to be familiar with one another's work, the expert opinion might not reflect reality. An added risk is the professional 'cannibalism' of equally well-established academicians not accepting each other's view and being highly critical of each other's work. Also, there is a question over the seriousness with which respondents are likely to treat an opinion poll.

This all makes the reliability, validity and objectivity of such methods questionable, and reliance on expert opinion (not professional judgement) is a flaw, especially in the light of the fact that expert opinions could be grossly inappropriate. Little scientific justification exists for continuing to rely on expert opinion.

Statistical validity

There is a considerable academic literature on league tables. The consensus is that the great majority fail the normal tests of reliability and validity, including statistical validity, that one would expect of any serious social science enterprise. (Brown, forthcoming)

For example, differences between the scores of two institutions may be statistically insignificant, but the methodology places them in an arbitrary scale, which exaggerates the differences. In other words, there may be little difference in quality between the institutions ranked first and eleventh (where eleventh does not even find a place in the top-ten list), but the uninformed public naturally thinks otherwise. Gerhard Casper, President of Stanford University, sent a personal letter to the editor at *USNWR* indicating his displeasure with the university ranking system:

Could there not, though, at least be a move toward greater honesty with, and service to, your readers by moving away from the false precision? Could you not do away with rank ordering and overall scores, thus admitting that the method is not nearly that precise and that the difference between #1 and #2 — indeed, between #1 and #10 — may be statistically insignificant? (Casper 1997, quoted in Provan & Abercromby 2000)

When considering multiple ranking systems, it becomes questionable whether meaningful differences between institutions exist ... statistically significant differences generally do not emerge between ... universities that find themselves at a distance from one another [in] the rankings. (Guarino et al. 2005)

Guarino attempts to avoid the ‘arbitrary weighting’ problem by using a Bayesian latent variable model ‘to estimate a university’s “quality” based on a set of observable features’. The hypothesis appears to be that the observed parameters are related in some linear way to the desired but unmeasurable ‘quality’ feature. Some of the coefficients are estimated, and the rest are then based on some random samples drawn from a population with the right distribution. The authors claim that this technique: ‘(a) determines the relative importance of different university features using information embedded in the data rather than subjective opinion; and (b) simultaneously determines the degree of uncertainty that surrounds the ranks.’

The technique is applied to the *THES* ranking of UK universities, and the *US News & World Report* ranking of some of the US universities. The resulting orderings show a significant shuffling, and also show the much-reduced extent to which meaningful differences can be claimed. For example, University College London (UCL) is listed tenth in the *THES* rankings, so the reader is invited, explicitly or implicitly, to understand that UCL is worse than the ninth and better than the eleventh listed universities. The Bayesian analysis suggests that it is only possible to reliably state that (in terms of the parameters under consideration) UCL is ‘worse’ than the first two entries on the list, and ‘better’ than those from twenty-second onwards.

The model being followed for the Australian *GUG* is slightly different from the other rankings. Its use of a five-star system to group institutions solves the problem of a statistically insignificant ranking system. For example, even though the difference in scores between two universities is statistically insignificant, the other ranking systems would allot different ranks to the two universities. But in *GUG*, such universities would probably be in the same category grouped with other universities of similar quality. This comes close to what some of the quality assurance agencies do. For example, the National Assessment and Accreditation Council (NAAC) of India accredits universities and colleges. In addition to the accreditation decision, it grades the institution on a nine-point scale, using letters and pluses, with C at the bottom of the scale and A++ at the top end of the scale (C, C+, C++, B, B+, B++, A, A+, A++). The national quality assurance agency of Indonesia grades the programs it accredits on a four-point letter scale. Critics of the Australian model complain that the public like things ranked numerically and any ranking that strays from this is not what people want to see, no matter how misleading a numerically ranked table may be.

Pseudo-scientific outcome

Similar to the use of complex formulae with indicators and weightings, media efforts also use the quality assurance outcome as one of the inputs, to give a misleading image of being ‘scientific’. In the UK, quality assurance outcomes

are used by the media-ranking groups. In addition to the Research Assessment Exercise (RAE) and Teaching Quality Assessment (TQA) outcomes, the quantitative institutional information collected by the Higher Education Statistics Agency (HESA) is readily available on topics such as entry requirements, campus accommodation, social facilities, exit qualifications and subsequent employment of students. All of these data are used to produce a range of ranking of institutions in terms of specific or overall performance. *THES* has used the RAE results while the *Daily Telegraph* offered a ranking based on the TQA outcome. The *Sunday Times* included both the RAE and TQA results. What is missing is the overall soundness of the method, and using one or two pieces of validated reliable data does not change the subjectivity of the media ranking.

6. Rankings and ratings by academic groups

While criticising the media groups and their ranking efforts, one should not forget that the system of higher education has all along been using ranking or methods close to ranking.

The ranking of students for their academic achievement needs no elaboration. Although there is a move towards letter grades, ranking of student achievement also exists in many countries. There is a heavy reliance on ranking in the admission process of some institutions and academia and the public have accepted it. Governments have done evaluations that come close to ranking, choosing the ‘key point institutions’ by the Chinese government is an example. Selecting key institutions with respect to various criteria has been practised in China since 1954.

The Top American Research Universities listed by the Lombardi Program on Measuring University Performance at the University of Florida’s Center for Studies in the Humanities and Social Sciences offers the assessment of the best public universities based on: total research and development; federally sponsored research and development; national academy members; Guggenheim and Fulbright awards; PhDs awarded; postdoctoral students; and National Merit and National Achievement Scholars (Garter 2002).

The National Research Council (NRC) in the United States did a very comprehensive study of the research doctorate programs in selected fields in 1995 and published a ranking of which there are criticisms about data quality and interpretation. Academia strongly expressed the view that the purpose of the NRC study should not be to fuel warfare over professorial status; instead, it should be to provide useful information to various stakeholders (including political and business leaders, foundations and professional associations, scholars and administrators, and students) about which programs are the most productive in creating new knowledge.

The *Educational Rankings Annual* of the Gale Research Inc. provides information and statistics taken from numerous sources that are broadly recognised as authoritative. Over 3,600 institutions are ranked and listed. An extensive index includes subject terms as well as institutions being ranked and it is perceived as a source providing a wealth of information.

In 2002, the Swiss Centre for Science and Technology Studies published its ‘Champions league’ of research institutions which ranked universities by their performance in research journal publications during the period 1994–1999. The next ranking was for the period 1998–2002. The centre ranks the top 683 institutions by four measures:

- total research journal publications
- number of publications in subfields with a substantial number of publications
- publications in qualified subfields as a percentage of all publications
- citations per research publication.

The centre publishes ranks by each of its four measures.

Shanghai Jiao Tong University’s institute of higher education (SJTU) first published its world academic ranking of universities in 2003, considering 6 factors:

- alumni who won a Nobel prize or a Field medal for mathematics (10%)
- staff who won a Nobel prize or a Field medal for mathematics (20%)
- highly cited researchers in 21 broad subject categories (20%)
- articles published in *Nature and Science*(20%)
- articles in *Science Citation Index-expanded* and *Social Science Citation Index* (20%)
- correction for institutional size (10%).

Williams and Van Dyke’s ‘International standing of Australian universities’ (‘the Melbourne Institute Index’) is based on 19 weighted factors. Williams and Van Dyke asked vice-chancellors of 172 overseas highly ranked universities to compare the international standing of each Australian university with the standing of universities in their continent and they asked Australian and Aotearoa New Zealand deans to compare the international standing of each Australian university with US universities. Williams and Van Dyke’s respondents placed most emphasis on the ‘quality/international standing of staff’, which is based on research performance.

7. Criticisms of academic ranking

Firstly, ranking implies placing every institution in a particular position of merit and it is difficult to evolve an instrument and method that can dependably distinguish between institutions of close standing. Secondly, quality assurance requires that institutions be evaluated without losing sight of their own sets of

goals and objectives. Any attempt to compile ranking tables according to generic criteria is contrary to the principle of quality assurance. It is also harmful to institutional diversity.

No clear or universally-agreed measure of quality in HE exists. Absent any consensus, each ranking combines an available set of observable indicators — each of which serves as a rough proxy for some factor notionally tied to quality — in a formula that can be questioned both on the basis of its contributing elements and the manner in which the elements are combined. (Guarino et al. 2005)

Thirdly, even if we assume that there is a dependable methodology to rank HEIs, it would be feasible only in small systems. As ranking would also imply that the whole system has to be covered within a time frame, it would be futile to attempt ranking in a large and complex system. At the most it can be done only at a superficial level, akin to the methodology followed by the media. Consequently, lack of validation of self-reported data, inconsistency in terminologies, lack of peer review, inability to consider institutional diversities, etc., would become unavoidable, thus rendering the outcome of the whole process useless.

For these reasons, quality assurance agencies have a leaning towards narrative reports and ratings, and ranking efforts by them have been very few.

Further, institutional improvement is the focus of most of the quality assurance efforts. Direct ranking, even if we assume that we do it within a short time, will put institutions into very fine classifications, which may make improvement difficult for the institutions at the lower rung. What may be useful is a middle ground, putting a large number of institutions into a few groupings, which would allow differentiation but preserve a wide range of institutional types with scope for improvement. This line of thinking has made many quality assurance agencies opt for reports and grades.

For example, the outcome of the quality assurance procedure in Indonesia includes a grade on a four-point scale, grade A to grade D. Grade A indicates that the course of study conforms to international standards, grade B indicates that the course is of good quality, grade C indicates that the course fulfills minimal requirements and grade D means not accredited. The NAAC in India classifies an institution on a nine-point scale.

These examples show that quality assurance agencies and HEIs also engage in rating with respect to institutional quality, but criticisms abound here also. Especially in the quality assurance mechanisms that result in a multipoint classification or grade, apprehensions are often raised about: the reliability of the instrument used to make discriminations; the benchmarks applied for different quality levels; the inter-team variances due to the human element involved; the ambiguity in the evaluative language used to record assessment

outcomes; the issues of institutional diversity; and the conflict with respect to ‘contextualisation vs consistency in the assessment framework’.

Rankings — even if done by academic groups — rely on quantification, indicators and weights. They force the multidimensional quality aspect into a linear scale. In this process, certain aspects of the institution that cannot be measured with weightings and numbers get distorted. The ‘health warning’ in the introduction to the UK’s ‘University management statistics and performance indicators’ includes such comments as ‘Comparisons should be made with extreme care’; ‘Careful attention must be paid to the caveats which precede each table’; and ‘Unfortunately [combined] rankings do attract a great deal of attention and mislead more than they enlighten’. The authors of the SJTU ranking have themselves published extensive criticisms/limitations of their own ranking (Nian & Cheng 2005) though Nian and Cheng conclude with the fatalistic ‘clearly they [ranking systems] are here to stay’.

Choice of indicators

The rankings by the SJTU are heavily weighted towards the sciences, and the research publications measures overlap. In Williams and Van Dyke (2005) the weighting given to laboratory disciplines is more than the other disciplines. Also, not including equity of undergraduate as a separate factor, some of the aspects that are criticised are calculating the revenue per student in a misleading manner for dual-sector institutions; the compressed scale for the undergraduate students; and the overlap in the research publication factors.

The researchers acknowledge that ... [The Melbourne Institute Index] offers few insights into which institution has the best teachers or who provides the most value-added experience for students ... [and] ignore the local contribution of universities. ... Both rankings [i.e. the New Zealand research ranking also] show a strong link between organisational mission and the final rank. (Clarke 2005)

In other words, they suffer from exactly the defect that purpose-related, peer-based quality audit is designed to avoid.

Arbitrariness in giving weightings

The manner in which weightings are allotted is also criticised. In the ranking by the Swiss Centre, the weightings given to the citations indices vary. For the arts and humanities citation index the weighting is 5%, for social sciences the citation index gets 11%; and for science it is 84%. The centre gives research journal publications in the arts, humanities and the social sciences less weight. The centre counts research publications in journals only, and therefore does not count books and other research publications. This understates research

publications in the arts, humanities and social sciences which publish more books and works of art than the empirical fields.

The comments on the SJTU and Melbourne Institute indexes referred to in the previous paragraphs also indicate arbitrariness.

Inconsistencies

Inconsistency among different rankings results in misleading outcomes. Jiao Tong ranks Harvard University first and Stanford University second while the Swiss centre ranked Stanford University eighth. Jiao Tong ranks the University of California, Berkeley fourth, which is placed ninth by the Swiss centre.

To cite another example, in the SJTU listing Australian National University (ANU) is ranked fifty-third and the University of Melbourne twenty-second. In the *THES* World University Rankings, ANU is ranked sixteenth and the University of Melbourne twenty-second. Williams and Van Dyke (2004) rank both ANU and the University of Melbourne at rank one. Although this is a national ranking about the international standing of Australian universities, it is interesting to note that two institutions of the same rank at the national level were ranked very differently in two different ranking systems at the international level — the difference in the ranks being sixth in one case and twenty-ninth in the other.

Statistical validity

Substantial analysis of the rankings' statistical validity has been carried out in the US, Canada and the UK. Most of the analyses completed in these countries find that the rankings are statistically irrelevant. In 1997, the National Opinion Research Center (NORC), a non-profit corporation affiliated with the University of Chicago, analysed the ranking methodology. It identified five main weaknesses, two of which relate directly to statistical validity. First, assigning weights to different indicators was found to be too arbitrary and rankings were sensitive to relatively small changes in the weightings used. Small changes in the weightings of indicators altered the results from year to year without any tangible change between institutions. Second, some measures were redundant; some variables were measured directly and they also had influence as indirect measures on other variables. When correction for the interaction between variables is not done, it results in overemphasis of those measures. It might lead to institutions attending to certain aspects of institutional quality, the ones that would fetch them a higher ranking, at the expense of some other vital aspects.

Unhealthy competition

There are also concerns that rankings will add to unhealthy competitiveness in the higher education system. Institutions competing for top rankings may forego cooperation with other institutions, which can be detrimental to the student and the institution as well as higher education in general.

It depresses me to witness our President negotiating with an editorial board to change their measure so the public will think we are better than others. (Professor Reuben Kaufman of the University of Alberta, in Provan & Abercromby 2000)

The question about the role of rankings in maintaining the divide between public and elite institutions and in the corporatisation of higher education has also been raised.

In general the criticisms about rankings whether produced by media or academic groups indicate that both of these suffer from similar limitations. But the danger in the academic rankings is that, in spite of its limitations, it gradually gains some amount of acceptance. At the beginning, academics were indifferent to media rankings and it took them quite some time to express their concerns. Contrary to this lukewarm response, rankings by academic groups have received almost immediate attention and they usually come with inherent acceptance. All criticisms are seen as just different academic viewpoints, which is more harmful to the higher education system.

8. Who uses the rankings?

If rankings are so inaccurate and irrelevant, why do academics still discuss university rankings? Is it because stakeholders find them useful? In fact, the usefulness of ranking to any group of stakeholders is not very promising and providing such a ranking has not been the agenda of academic groups.

Students

Students' choice of institution may depend on a variety of factors such as locale, type of institution, fee structure and even the 'freedom' on campus. In most cases, the students shop around: looking for an institution that will meet their expectations and criteria, both academic and otherwise. The evaluation is individualistic, depending on the various considerations students care to exercise. It is reasonable to expect that students may use many sources to make a choice of the institution, rather than just relying on ranking.

Studies conducted in India (and other countries) have shown the factors that influence the choice of institution are:

- proximity to home
- reputation of the college
- availability of the course
- fee structure
- placement
- sports facility
- company of friends.

It may be tempting to believe that the factor listed by students as ‘reputation’ is due to media ranking. It is true that rankings contribute to reputation and vice-versa, but the students’ responses for the question on ‘sources of data’ that influence the choice of the institution do not mention media ranking in a significant way. This means that the respondents have not judged the reputation of the college by the rankings, but by data from other sources. Around 60% of the students indicated that the most influential data-source in their choice of institution is ‘Advice from parents’. Around 80% indicated advice from parents, advice from friends, views of current students and advice from others under the first two rankings. Around 90% of the respondents indicated ‘Advice from parents’ among one of the top four priorities; following this is Web information. Rankings are placed last. This was the case when the survey was conducted in 2004, which covered all first-year students in four colleges of a major city in India that found a place in the top ten rankings of *India Today*. The figures in other colleges could be lower still (Stella 2005a).

Studies have been conducted on the information the stakeholders seek from external sources and the use to which the information is put. A Chilean study (Lemaitre 2002) found that quality and standards are not the main criteria influencing the decisions on choosing an institution of higher education; neither applicants nor their parents, teachers or vocational guides have an idea of what ‘quality’ is. General institutional reputation and prestige seemed to be enough for their choices. The study concludes that the public has to be educated about the use of information sources.

A study carried out in 1999 by the Higher Education Funding Council of England (HEFCE) found that there was a low level of awareness of quality-related information, despite the availability of the Quality Assurance Agency’s assessment reports on its web site (and extensive mailing of reports to various groups of stakeholders). A telephone survey conducted among the public by the Council on Higher Education Accreditation (CHEA) in 1999 found that most of the respondents did not know who performed accreditation in the United States. The same survey revealed that a good number of respondents did not have a clear idea about accreditation and were ready to take educational decisions not based on the accreditation status.

Studies have shown similar findings about the league tables and rankings published by the media. There has been some limited research completed in the United Kingdom and Canada on student university choice. Studies have been conducted about who uses the rankings and league tables and with what impact. Patricia McDonough and other researchers from the Graduate School of Education at the University of California studied the data drawn from a 1995 Freshmen Survey that was conducted on a national scale, surveying 221,897 full-time freshmen at 432 four-year colleges and universities in the United States. Sixty per cent of students surveyed found the ranking to be of no importance, 30% somewhat important and 11% rated the ranking as very important in making their university choice (McDonough 1998). A 1997 survey of college freshmen released by the Higher Education Research Institute of the United States reported that, of the 251,232 freshmen surveyed, only 8.6% considered college ranking in national magazines to be 'very important' when selecting a college. The sources of information that were most frequently consulted (in rank order) include: parents, friends and classmates, students attending or soon to be attending colleges of interest, college brochures, and materials in high-school guidance and career centers. Only students from the upper middle-class and upper-class families tend to use these guides.

Governments and funding agencies

Governments and funding agencies rarely consider media rankings, paying more attention to academic efforts. One exception was in the Philippines, where government officials approached *Asiaweek* for additional information on higher education in order to make budget decisions. More generally, quality assurance outcomes inform and influence policy-making either directly or indirectly. This is understandable as most of the quality assurance agencies have been either government initiatives or established to serve government functions. Even if the quality assurance agency is not a government initiative, and even if the funding bodies and governments have their own mechanisms to monitor higher education institutions, there is usually a close link between the government's decision and quality assurance outcomes.

The funding links in the United Kingdom and the United States are two good examples to be noted. In the United States, although accreditation is still considered voluntary, the student aid funds (scholarships) from the federal and state governments are linked with the accreditation status of the institutions. Many other federal funds and billions of dollars in state funds are also dependent on the accredited status of educational institutions.

Employers

While governments and funding agencies might consider quality-related data provided by quality assurance agencies, employers are less likely to actively seek information on institutional quality. In general, the most important criterion that guides the recruitment choices is the set of personal qualities of the candidate, such as aptitude for team work, creativity, capacity to solve problems, openness to further learning and communication skills. These qualities may be assessed directly by the employer through interviews and tests. Specific information on institutional quality and standards tends to be secondary, although general institutional prestige is often acknowledged and taken into account. Media rankings and league tables could possibly play a role in informing institutional prestige.

9. Conclusions

With improving family finances, increasing social aspirations and greater awareness that investment in education will lead to manifold personal benefits, parents and students today do not hesitate to invest resources for access to quality education. As the investment in terms of finance, efforts and time increases, an increasing percentage of students and parents will start looking for quality institutions and begin to consult the wide variety of available sources. For this population, popular rankings might have an appeal. For students bound for higher education and looking for a quick easy way to compare the many (in some countries, thousands) of institutions available to them, ratings services, though less influential than personal advice, could be an attractive way to reduce the choices to a manageable number.

In an age when people love lists and orderings, institutional rankings tap into this contemporary fad. No matter how irresponsible and meaningless the institutional rankings are, unless stakeholders are oriented to use meaningful quality-related data they are likely to continue. It is the responsibility of the academic community to bring to the attention of stakeholders the flaws in the subjective methodology followed by the media and their 'profit and publicity' driven motives.

We already know some of the right things to do, but don't always do them or promote them. Professional judgement is always needed to interpret all indicators, even apparently straightforward ones such as student–staff ratios.

Peer review seems the most valid method for measuring the quality of academic research outputs. However, this perceived validity translates into high costs and complex procedures. (Clarke 2005)

It is a sad commentary that all, including the universities themselves, seem to prefer rankings to the more nuanced narrative output of an audit, even though rankings are coarse, misleading and add no value, but appear readily ‘intelligible’.

In fact, it is too generous to the rankings simply to observe that they add no value — they are actively harmful.

Ranking systems are not neutral to the institutions they study but affect them. Many ranking systems heavily weight indicators of reputation ... Given that the rankings themselves play a prominent role in affecting reputation, the circular nature of these endeavours makes them a particularly strong self-perpetuating force. Thus, it is important to ask whether this force acts in the best interests of students and society. If universities are to be judged by the standards set by ranking systems and have strong incentives to conform to them, does moving in this direction take us closer to or further from true educational quality? (Guarino et al. 2005)

A study conducted in the United States gives evidence to this concern and reports that ‘universities were offering financial aid to attract students who could improve their rankings’ in the *USNWR* and they ‘have been very frank about this, and said they are essentially buying students’ (*THES* 10 February 2006). *The Chronicle of Higher Education* reports that:

In an effort to raise their *USNWR* rankings ... law-school admission officers are over-emphasising standardised test scores and thereby discriminating against African-American applicants. (Marcus 2006)

Sharing this concern and as a move towards doing the right things, the following section lists a few right things that can be done by the various interested parties, preferably in collaboration, to counter the effects of inappropriate, inaccurate and irrelevant rankings through constructive ways to promote meaningful use of quality-related data.

10. Actions

Institutions

Higher education institutions should make as much institutional data available to the public as possible, structured so as to be relevant to making informed decisions. The institutional web site is a good place to begin to make the data available. Higher education institutions should also ensure that the data is authentic, unambiguous to the users and free from discrepancy and inconsistency.

Quality assurance agencies

Quality agencies should reflect on the information they provide and the form in which it is provided, in order to meet the needs of users who currently turn to rankings. The agencies might re-examine how to incorporate some of the desired elements into their reporting strategy. Providing institutional profiles and making assessment reports public are two ways in which quality assurance agencies have tried to address the need for more public information.

Partnerships

A responsible partnership between higher education institutions, the quality assurance agencies and the media can also be a way forward. This would ensure there is more meaning in the data made available to the public — not a far-fetched idea, if these groups learn to build on the commonalities and each other's strengths. In the various media rankings, five areas of emphasis are often used to evaluate or rank HEIs: academic staff, resources, student quality, student support, and curriculum requirements. These areas are also important to quality assurance agencies. They are also the areas in which the HEIs make claims in their prospectuses and information brochures. This commonality implies that there is enough common ground for the HEIs, quality assurance agencies and media groups to become partners. Instead of relying on reputational scores and other easily measurable indicators in these five core areas, the media should use more reliable data by working with HEIs and quality assurance agencies.

The major criticisms about the media effort lie in the subjective methodology and the apparent 'profit and publicity' driven motive. To correct these two flaws, academics can offer a reliable methodology and the merits of both the approaches should be combined. The media should make a commitment to do away with mere opinion polls and self-reported data, which often reflect only casual opinions, gossip and exaggerated data, in favour of open, factual, holistic and systematic 'quality-related data' about the institutions. Using the quality assurance outcome as a major input is one way to go forward. Peer reviewed and validated data of the academic groups have already been used as a basis for rankings by some media groups.

An example of this approach is the partnership between the Centre for Higher Education Development (CHE) and *Stern* magazine in Germany which provides a personalised university ranking system of 242 nationally recognised universities and professional schools.

Stakeholder-oriented comparative data

Increasingly, users of report cards can, via the Internet, choose the indicators of greatest interest to them and ascribe their own weights or priorities. Such an approach provides a solution to two of the most persistent criticisms of report cards — that weights are subjective and arbitrary, and that the evaluations address a narrow group of users. (Van Dyke 2005)

In the United States, the members of the Common Data Set Initiative (CDS), which is a collaborative effort among data providers in the higher education community and publishers, present the institutional profile in the website on a structured form. One of the goals of this collaboration is to improve the quality and accuracy of information provided to all involved in a student's transition into higher education. There are similar examples where groups of institutions have joined together to provide institutional data on common data structures.

The personalised ranking effort of CHE is a joint undertaking of the media and academia. Unlike the other fixed media rankings, CHE's efforts have resulted in a stakeholder-oriented flexible ranking system. The user can determine the ranking of any university according to the criteria of his/her choice in many dimensions. Groups of similar HEIs may care to explore this effort jointly. When alternate sources of reliable institutional data are available to the public, the harmful effects of inappropriate uses of rankings should decrease.

Publicise what rankings are 'measuring'

The meaning of the word 'quality' has shifted over the last 40 years. The word that best captured its meaning 40 years ago was 'reputation', and it was applied to such things as Rolls Royce cars and Oxford and Harvard Universities. Since then, the word 'quality' has become 'democratised', and the phrase that best captures its current meaning is 'fitness for purpose' (where the various purposes may be manifold, complex, and even contradictory).

Rankings have emerged on this scene and purport to measure quality, whereas in fact they are mostly measuring (long-standing) reputation. In commenting on the *THES* and *SJTU* rankings, Marginson (2006) observes that they produce very different results 'below the very top'. Usher and Savino (2006) report that:

Regardless of the ranking scheme employed, "top universities" are almost always going to come out as top universities. The variation between rankings occurs lower down the scale; there, even small changes in methodology can change rankings significantly.

Usher & Savino also say:

... institutional ranking systems don't measure what the authors think they are measuring. ... What our results here show is that most indicators are probably epiphenomena of some underlying ... feature that is not being measured. ... some ... "X factor" ... Our guess is that age of institution, faculty size and per-student expenditure are probably excellent candidates to be these "X factors".

Therefore, one main problem with rankings is that they are measuring one thing, but most of their readers think they are measuring something else. If it is true, as several respected writers have already said, that 'rankings are here to stay', it is essential that academia:

- does not give them spurious validity by cheering whenever one's own university scores highly
- actively publishes well-argued criticism when their producers make inflated or misleading claims
- helps the public to distinguish between the information they often say they want and the information that their enquiries show they actually need, and
- provides easily accessible information of the sort that the public needs.

People are already saying 'rankings are here to stay', often in the same breath as they complain about the rankings. Merely to accept them as a 'given' is analogous to observing that 'nuclear weapons are here to stay', but then use this as a reason not to work for their containment and limitation. A large number of campaigns and reach-out strategies at various levels may be necessary for different stakeholders to help them use quality-related information in the most appropriate and helpful way. In both secondary and higher education institutions, quality literacy should become a part of academic and personal guidance and counselling. This would facilitate the emergence of a quality-literate society — a society that can take decisions informed by quality-related data. This is a priority area for many quality assurance agencies the world over. Intergovernmental bodies such as UNESCO and the OECD have recognised this as one of the strategies under the capacity development framework for member countries.

REFERENCES AND FURTHER READING

- Brown, R. forthcoming, 'League tables: Do We Have to Live with Them?'
- CHE (Centre for Higher Education Development), www.che.de
- CHEA (Council on Higher Education Accreditation), www.chea.org
- Clarke, M. 2004, 'Weighing Things Up: A Closer Look at the US News and World Report's Ranking Formulas', *College and University Journal*, vol.79, no.3, pp.3–9.
- Clarke, M. 2005, 'Quality Assessment Lessons from Australia and New Zealand', *Higher Education in Europe*, vol. 30, no. 2.
- Cohen, D. 1999, 'Magazines Rankings of Asian Universities are Popular With Readers, Not Academics', *The Chronicle of Higher Education*, vol. 45, no. 36, p.A51.
- CDS (Common Data Set Initiative), www.commondataset.org
- Crissey, M. 1997, 'Changes in Annual College Guides Fail to Quell Criticisms on Their Validity', *The Chronicle of Higher Education*, vol. 44, no. 2, p.A67.
- Davis, D. 1996, 'The Real World of Performance Indicators: A Review of Their Use in Selected Commonwealth Countries', London, UK, Commonwealth Higher Education Management Service.
- Dill, D. & Soo, M. 2004, 'Is There a Global Definition of Academic Quality? A Cross-national Analysis of University Ranking Systems', *Public Policy for Academic Quality*, University of North Carolina at Chapel Hill, January 2004.
- Garter, D. S. 2002, 'A Review of Measures Used in US News & World Report's America's Best Colleges', An Occasional Paper, The Lombardi Program on Measuring University Performance, Summer 2002.
- Good Universities Guide* 2006 ed., Hobsons Australia Pty Ltd, Melbourne.
- Gourman, Jack, & Princeton Review 1997, *The Gourman Report: a Rating of Graduate and Professional Programs in American and International Universities*, 8th ed., rev., Random House, New York.
- Gourman, Jack c. 1989, *The Gourman Report: a Rating of Undergraduate Programs in American and International Universities*, 7th ed., National Education Standards, Los Angeles, California.
- Graham, H. D. & Diamond, N. 1999, 'Academic Departments and the Ratings Game', *The Chronicle of Higher Education*, vol.45, no. 41.
- Guarino, C. et al. (2005), 'Latent Variable Analysis: a New Approach to University Ranking', *Higher Education in Europe*, vol. 30, no. 2.

- Holub, Tamara (2002), 'College Rankings', ERIC Digest (ED 468728), Report: EDO-HE-2002-03, ERIC Clearinghouse on Higher Education, Washington DC.
- Hunter, B. 1995, 'College Guidebooks: Background and Development', *New Directions for Institutional Research*, 88, EJ 518243, ERIC, Washington DC.
- India Today*, 13 June 2005.
- India Today*, 17 May 2004.
- Kersten, G. 2000, 'Grading on the Curve: College Ratings and Rankings', *Points of Reference*, January 2000.
- Kirk, S. A. & Corcoran, K. 1995, 'School Rankings: Mindless Narcissism or Do They Tell Us Something?', *Journal of Social Work Education*, vol. 31 no. 3, pp.408–14.
- Lemaitre, M. J. & Raczynski, D. 2002, 'Stakeholders in Higher Education: How to Deal with Complexity', in *Quality Assurance in Higher Education: Quality, Standards and Recognition*, Gnanam A. and Antony Stella (Eds.), Allied Publishers, New Delhi.
- McDonough, P. M., Antonio, A. L., Walpole, M. B. & Perez, L. X. 1998, 'College Rankings: Democratized College Knowledge for Whom?', *Research in Higher Education*, vol. 39, no. 5, pp.513–37.
- McGuire, M. D. 1995, 'Validity Issues for Reputational Studies', *New Directions for Institutional Research*, no. 88, pp.45–60.
- Mangan, Katherine S. 2006, 'Law Schools' Emphasis on LSAT Scores Hurts Black Applicants, Report Says,' *The Chronicle of Higher Education: Today's News*, 21 February 2006.
- Marcus, Jon 2006, 'Aid in US Favours Wealthier Freshers', *Times Higher Education Supplement*, 10 February 2006.
- Marginson, S. 2006, 'Australian Universities in the Global Context', *Campus Review*, 22 March 2006, pp.8–9.
- Merisotis, J. & Sadlak, J. 2005, 'Higher Education Rankings: Evolution, Acceptance and Dialogue', *Higher Education in Europe*, vol. 30, no. 2.
- Monks, J. & Ehrenberg, R.G. 1999, 'U.S. News & World Report's College Rankings: Why They Do Matter', *Change*, vol. 31, no. 6, pp.42–51.
- Morrison, H. G., Magennis, S. P. & Carey L. J. 1995, 'Performance Indicators and League Tables', *Higher Education Quarterly*, vol. 49, no. 2.
- Nian, C. I. & Cheng, Y. 2005, 'The Academic Ranking of World Universities', *Higher Education in Europe*, vol. 30, no. 2.

- OECD (Organisation for Economic Co-operation and Development),
www.oecd.org
- Provan, David & Abercromby, Karen 2000, 'University League Tables and Rankings: a Critical Analysis', Commonwealth Higher Education Management Services (CHEMS), paper no. 30, London.
<http://www.acu.ac.uk/chems/onlinepublications/976798333.pdf>.
- Rocki, M. 2005, 'Statistical and Mathematical Aspects of Ranking: Lessons from Poland', *Higher Education in Europe*, vol. 30, no. 2.
- Stella, A. 2005a, 'Measuring Institutional Quality in Higher Education: Rankings, Ratings and Reports by Academics and Media', *University News*, Association of Indian Universities, New Delhi, vol. 43, no. 22, pp.1–6.
- 2005b, *University News*, no. 23, pp.18–23.
- 2005c, *University News*, no. 25, pp.1–5.
- Stuart, D. L. 1995, 'Reputational Rankings: Background and Development', *New Directions for Institutional Research*, no.88, pp.13–20.
- Tight, M. 2000, 'Do League Tables Contribute to the Development of a Quality Culture? Football and Higher Education Compared', *Higher Education Quarterly*, vol. 54, no. 1, pp.22–42.
- UNESCO (United Nations Educational, Scientific and Cultural Organization),
www.unesco.org
- Usher, A. & Savino, M. 2006, *A World of Difference: a Global Survey of University League Tables*, ON: Educational Policy Institute, Canadian Education Report Series, Toronto.
- Van Dyke, N. 2005, 'Twenty Years of University Report Cards', *Higher Education in Europe*, vol. 30, no. 2.
- Williams, R. & Van Dyke, N. 2004, 'The International Standing of Australian Universities', *Melbourne Institute Report No. 4*, The University of Melbourne, 24 November 2004.
- Yerbury, D. 2006, 'Spreading Universities' Foreign Risks', *The Age*, 12 January 2006.