3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

130

Grading Universities Is Better Than Ranking Them

Peter Smith

Mahidol University International College
999 Phutthamonton 4 Road, Salaya,
Nakhonpathom, Thailand, 73170.
icpeter.mahidol@gmail.com

Keywords: Universities, International ranking, Grading proposal.

Abstract: Despite their present popularity I want to argue that the major university ranking systems are both misguided and may generate negative consequences. If we must evaluate on a comparative basis, as sometimes I think we must, then a grading system is a much better system than ranking.

My paper presents the main systems of ranking currently in use, points out what I believe to be the weaknesses of ranking and the often negative consequences it has, and suggests an alternative system based on my preferred system of grading.

Problems of ranking systems include the biases introduced by the initial ranking criteria chosen, the unlikely changes that occur in a particular university's rank from year to year, the unreliability of some of the data used and the very concept of a single ranking score. The negative consequences of ranking include an obsession with short-term and narrowly focused 'production targets' rather than wider and more general concerns with quality and excellence; the need for a vast and time-consuming bureaucratic apparatus of data collection; the practice of 'gaming' the system; false comparisons; and facile competitions for prestige ('ranking envy').

Although criticisms of ranking systems are now widespread, ranking remains enormously popular and is unlikely to end any time soon. Merely pointing out the weaknesses and deleterious consequences of ranking systems is unlikely to persuade participating organizations to end their involvement in the competition for prestige. What is needed is a practical alternative — which I think a simple grading system provides.

1. Introduction

Since the 1980s, various systems of ranking universities have become increasingly popular, first in the United States, with the now ubiquitous *US News and World Report* (1983-), and then internationally, from 2003 onwards, with the development of a whole bevy of ranking systems, most notably the 'Academic Ranking of World Universities' (ARWU) by Shanghai Jiaotong University (2003-); what is now the Quacquarelli Symonds (QS) World University Rankings (2004-); and the *Times Higher Education* (*THE*) World University Rankings (2009-) [1]. Produced on an annual basis, these four ranking systems have come to have an enormous impact on university and college objectives and thinking, whilst the three international systems have also become a major influence on national educational policies in many countries, and the *US News* an enormously influential guide to college choice for millions of American high school students and their parents.

Despite their popularity, I wish to argue that the various ranking regimes are both flawed as a system for evaluating universities and have negative consequences. Nevertheless, I recognize that an

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

131

international system for evaluating and comparing universities has considerable value both for national education ministries and university administrators anxious to improve the quality of education in their countries and institutions, and for would-be students deciding where they should study. The demand for evaluation systems is considerable, and anyone critical of the present rankings is bound, I think, to try to propose alternatives. This I will seek to do in the final part of the paper after a brief review of some of the main criticisms of the present ranking systems.

2. Criticisms

Numerous commentators – mostly serving academics – have pointed out the shortcomings of these various ranking systems, but to little avail: the ranking systems have proven largely immune to criticism [1, 2]. These criticisms include the following:

2.1. **There is no agreement on what basis academic excellence is to be assessed**. This is vividly illustrated by a comparison of the three major international ranking systems, each of which has its own mix of criteria of excellence – and its own way of measuring those criteria. In broad terms, all three focus on (i) the quality of teaching and 'education'; (ii) the quality of faculty; and (iii) research output and citations, but each defines, weights and quantifies these variables in different ways. Unsurprisingly, the resultant rankings vary considerably in their choice of universities that are given high rank: there is no universally agreed hierarchy of universities. In the 2013 rankings, for example, the National University of Singapore is ranked 24th in the QS system, 26th in the *THE*, but is only in the 101-150 band for ARWU, whilst the more specialized London School of Economics ranks 32nd in the *THE* system, but 68th in QS and again is in the 101-150 band in ARWU [3]. None of these three systems is necessarily any better than the others, but such marked divergences indicate clearly how varied the basis for their evaluations are. Every ranking system is making subjective and arbitrary value judgements about what 'quality' is.

2.2. **The very concept of a single set of ranks is problematic**. Universities are complex institutions, often with large numbers of schools, departments and faculty. Is it possible – or even sensible – to try to represent this complexity with a single score? Again, different institutions have different objectives which can not be easily accommodated within a single ranking system. For example, some colleges deliberately seek to cater to an academic or social elite whilst others try to be open to a much broader range of students. Again, some institutions have a narrow subject focus and others offer a wide range of schools and disciplines. In such cases, there are no meaningful overall comparisons, and by trying to devise a single ranking for all colleges, we end up with a system that is ill-suited to at least some – or perhaps many. If higher education is a complex product crafted by its practitioners rather than some mass-produced homogenous commodity, we would expect individual institutions to have their own distinctive strengths which are unlikely to be identified by a single composite score.

2.3. **Universities change slowly and an individual institution is highly unlikely to be significantly better or worse this year than it was last year – or even several years ago**. Yet ranks, and the aggregate scores of variables on which they are based often fluctuate wildly. Which is more likely? That the university has rapidly changed or that the ranking system is faulty? For example, the University of Nottingham in England was ranked 174th by *THE* in 2010-11, then shot up to 120 in 2012-13 before falling to 157 in 2013-14 [4]. If the ranking was accurate, only seismic shifts would produce such a degree of change.

2.4. **Some of the major changes in rank are attributed to changes in the rankings systems themselves, with changes in the weight given to particular variables**. This can be seen as a good

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

132

thing – with greater experience, ranking agencies are able to critique their own work and seek to improve it by developing better measures of excellence, but the massive changes in the ranking of particular institutions resulting from sometimes quite minor adjustments to the ranking criteria suggest that at least some of the variables are highly arbitrary. The fluctuating methodologies and evaluation criteria employed by ranking organizations also make any meaningful comparisons difficult.

2.5. **Setting the agenda**. Each ranking systems is based on a particular set of choices about what 'excellence' in higher education is. These values are not neutral, but reflect the preferences of the particular ranker. They are often implicit and may be unquestioned. By choosing them as criteria of excellence, rating agencies influence universities to value them, and so the university's mission, objectives and activities may be shaped by them. The rankings agencies are doing more than merely reporting data in an objective fashion but are rather taking sides in educational debates. In some instances, these choices also create systematic bias. Several examples may be given:

2.5.1. **What is a university for?** There is no one universally agreed goal of higher education, but clearly the main international ranking systems focus on variables related to research output and reputation status and ignore or place little value on such goals as involvement with its local community or practical problem solving. Again, the actual quality of teaching – as opposed to proxy measures (below) is effectively neglected because it is difficult to measure.

2.5.2. **A Science bias**. One very evident bias occurs in the ARWU ranking from Shanghai with the high value it puts on Science and Mathematics compared to other disciplines with Nobel Prizes, Fields Medals, and science journals all gaining high marks. Prestigious colleges which are focussed on the Social Sciences or Business such as the London School of Economics, the Institut d'Etudes Politiques de Paris (Sciences Po) and the Harvard Business School are disadvantaged in such a schema. Ironically, ARWU otherwise prides itself on the 'objectivity' of its criteria.

2.5.3. A related bias is **the privileging of journal articles over books** – thus ignoring what has traditionally been one of the most highly-regarded academic products in the Humanities and Social Sciences [5]. In part this is presumably because articles are easier to produce on a year-to-year basis and so measure in annual academic audits. The process of journal citation further favours the Sciences because of the greater measurable productivity enabled by the tradition of frequent journal publication and multiple authorship by a research team. A solitary scholar in the Humanities or Social Sciences who labours for several years over a carefully crafted book may be making a major contribution to scholarship but it will receive little attention in a system biased towards journals.

2.5.4. Another implicit value in some of the ranking systems is **the importance attached to financial inputs**. This is particularly obvious in the *US News* rankings which measure and reward institutional wealth and expenditures, including endowments, the number of alumni donors, faculty salaries and expenditure per students. Again, the *THE*'s variables include research income (including from industry and income per academic).

2.5.5. The *US News* is also said to **value 'selectivity' over efficiency**, giving higher scores to colleges that turn away a larger proportion of applicants than to more inclusive colleges which accept a wider range of applicants and are then successful in helping them complete their degrees. A corollary of this emphasis is the neglect of price and cost effectiveness as variables – even though the cost of higher education is likely to be a major concern for most American families [6].

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

133

2.6. **Is there a language bias?** In all three international ranking systems, almost all top-ranked universities are in English-speaking countries. Does that result really reflect reality? It seems unlikely that universities outside of the English-speaking world are really so inferior – some bias in the variables seems more likely. One possibility is that the importance given to English-language journals plays a role here. Serious scholarship in other languages is less likely to appear in internationally accessible journals and so will be under-represented in global comparative data.

2.7. **The measurement of chosen variables is often problematic**. There is no way of directly measuring the educational quality of a college or the teaching and research abilities of its faculty so measurable proxies are used instead. Some of these proxies are of dubious value. The quality of teaching, for example, is unlikely to be directly correlated to such variables as class size, faculty salary, the proportion of professors who have the highest degrees and are full-time employees, or by research productivity. Such characteristics may be relatively easy to measure, but that does not make them necessarily relevant.

2.8. A related problem is **the generation of seemingly precise data on the basis of an inexact methodology**, and the belief that the rankings themselves have some objective reality – an example of what I have termed a 'false facticity', the common belief that there are measuring tools for everything, and that the very act of assigning a number to something gives that number an objective reality. Einstein's alleged adage is relevant here: 'Not everything that can be counted counts, and not everything that counts can be counted' [7].

2.9. There is **controversy regarding the validity of reputational surveys** on which several of the ranking systems rely heavily, with considerable doubt over whether those asked to evaluate rank are sufficiently knowledgeable to judge the quality of numerous institutions, including those in other countries, and recognition of the danger of a feedback mechanism developing in annual rankings whereby raters' evaluations are influenced by their knowledge of the existing ranks.

2.10. More generally, there is **reason to doubt at least some of the data**. Some data is subject to manipulation (below), but there is also a distorting statistical impact when small sets of data are used, as with the number of Nobel Prize winners or measures of citations in academic fields with only a few publications [8]. Random distortions can also occur as with the rankings drop of the University of Malaya from 89th in the *THES*-QS ranking in 2004 to 169th in 2005 after Malaysian Chinese and Indians were no longer classified as international students [9], and the placement of the University of Alexandria as 147th in the *THE* ranking in 2010, a relatively high placing due mainly to including the 320 articles published by one academic in a journal of which he himself was the editor [7].

2.11. Possible **negative consequences** of ranking include the following:
2.11.1. **Work load**. As the ranking exercises have grown more complex, the amount of work required from the universities in providing data has increased. This is likely to increase the importance of administration over teaching and research, and in some cases may impose heavy administrative work-loads on academics and lessen job satisfaction. The work also adds to the universities' administrative costs.

2.11.2. **Privileging**. Rankings may deepen the divide between elite and non-elite institutions where such distinctions exist, and in some countries, governments may decide that the best use of scarce resources

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

134

is by disproportionately expending them on institutions with a proven record of excellence shown by rankings success. Such colleges are likely to already be amongst the wealthier institutions so the divide between resource-rich and resource-poor institutions is increased. Again, in order to enhance their perceived excellence, universities are likely to become more selective in their admissions policies, favouring merit over need [10].

2.11.3. **Prestige threat**. An annual ranking serves to sharply focus attention on differences of rank between universities, perhaps leading both to pride in gaining rank and 'ranking envy' on the part of those institutions of lesser or declining rank. Universities can – and should – strive towards excellence, but worrying about their marginal success in sets of controversial and contested variables imposed by ranking agencies may not be an effective way of achieving excellence.

2.11.4. '**Gaming the system**'. If rewards of prestige and material advantage are to be gained from high rank, then we should not be surprised if some institutions seek to manipulate their scores, and there is a growing body of evidence of such practices. [1, 9, 11, 12]**.**

2.11.5. **'Short-termism'**. The ranking process is likely to lead to an urgent concern with enhancing university performance in areas valued by the ranking process. This reinforces the bias towards easily measured but generally narrowly focussed journal articles (above), and disregards more reflective work such as 'big idea books' that once might have helped to define the discipline but may now be neglected [13].

3. Alternatives

Although extensively criticized, the present ranking regimes are also widely held to have several positive consequences. These may include: (i) universities paying more attention to those factors which are valued in ranking systems and therefore becoming more aware of the facilities or qualities which they need to develop or improve – paying more attention to encouraging research, for example; (ii) the provision of more resources, especially by governments seeking greater educational prestige; and (iii) increased transparency and public accountability. Any alternative systems of evaluation must have similar positive consequences.

Several alternative systems of evaluation already exist at a national level – particularly in the United States. These include the National Survey of Student Engagement (NSSE) that tries to assess the extent and nature of student learning, and the Collegiate Learning Assessment (CLA) that assesses students' abilities in critical thinking, analytic reasoning, written communication and problem solving. Both of these collect data from surveys of freshmen and seniors in participating institutions, and so have an important 'value-added' component in their analysis [2]. There is also the distinctive system used by the *Washington Monthly*, which differentiates universities on how well they promote research, social mobility and an ethic of service to the community [14, 15]. Other approaches could include studies of the career success of graduates; some system of 'Do-it-yourself' rankings in which users are able to arrange ranking data according to their own educational priorities [2]; or merely aggregating the rank scores of the three most widely-used international systems. It would also be possible to give much more attention to existing rankings by individual subject area – which are relatively unproblematic, than to university level rankings.

Another alternative approach is to grade universities in much the same way that academics grade students' essays and examinations – some variant of the A, B, C, D, F system. Such an approach was apparently proposed by the American Bureau of Education in 1911 but was poorly received. It has recently been reproposed by Marguerite Clarke, however, with broad 'quality bands' replacing

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

135

misleadingly precise ranks [2, 16].

My own proposal would mirror the grading system I use with my students' work. In relation to universities, the main characteristics would be as follows:

-1. Establish clear and public criteria for each grade.

-2. Do *not* mark on a curve: if all institutions are excellent, then they are all graded 'excellent' (A); if all are poor, they are all graded accordingly (D).

-3. Discuss the results with all institutions, identifying strengths, weaknesses and areas for possible improvement.

-4. Do *not* reify the grades: institutions – like students – may change over time (both for better or worse). Be prepared for change.

The ultimate purpose of the exercise would not be to define the University of A as being better than the University of B but to facilitate a system in which all universities could improve and excel. Ideally, freed from the pressure of exact annual rankings, universities could more easily focus both on their own definitions of excellence and on their particular strengths and weaknesses.

Such a system of grading could also be used on a subject base. Even with ranking, subject specific comparisons are much more meaningful than evaluating universities as total institutions: individual academics within a particular discipline will be more likely to have a clear and precise knowledge of the best universities for their particular speciality or sub-specialty, and such comparisons have often been used in advising undergraduates which graduate school to attend [2]. However, such advisement normally takes a grading rather than ranking form with advisors giving their students a short list of excellent institutions from which to choose rather than specifying an exact rank order.

Who would implement such a system? The main promoters of university rankings at presently are either ranking organizations (most of which are commercial in nature and purpose) or governments. Whilst universities supply most of the raw data that is used in ranking exercises, they are strangely passive in either determining or evaluating ranking procedures or even in defining what university excellence might be. This is an obvious lacuna which may well contribute to the perceived weaknesses of the present ranking systems. Groups of universities (such as the Russell Group in Britain) or international associations of universities could well take a lead in ranking reform – including consideration of proposals such as grading.

As to the details of any grading scheme, they must include the following: (1) A clear definition of what academic quality is held to be; (2) The devising of measures that directly measure this definition of quality; and (3) The avoidance of any measures that can be subject to perverse manipulation.

References

1. P. Smith, Ranking and the globalization of higher education, *Silpakorn University Journal of Social Sciences, Humanities, and Arts*, vol. 12/2, pp. 35-69, 2012.

2. L. Myers and J. Robe, College Rankings: History, Criticism and Reform, Washington, D.C.: Center for College Affordability and Productivity, March 2009.

3. For the respective rankings of the National University of Singapore and the London School of Economics see Academic Ranking of World Universities 2013 (at http://www.shanghairanking.com/ARWU2013.html); QS World University Rankings 2013 (at http://www.topuniversities.com/university-rankings/world-university-rankings/2013#sorting=rank+region=+country=+faculty=+stars=false+search=); and THE World University Rankings (at http://www.timeshighereducation.co.uk/world-university-

3-5 February 2014- Istanbul, Turkey
Proceedings of INTCESS14- International Conference on Education and Social Sciences

136

rankings/2013-14/world-ranking). All last accessed 2 December 2013.

4. C. Phelps, We'll never be royals, *The Chronicle of Higher Education*, 16 October 2013.

5. See, for example, the complaint of American sociologists against the National Research Council's ratings: S. Jaschik, Sociologists blast doctoral rankings, *Inside Higher Ed*, 21 March 2011.

6. M. Gladwell, The order of things: What college rankings really tell us, *The New Yorker*, 14 February 2011.

7. D. D. Guttenplan, The questionable science behind academic rankings, *New York Times* (Global edition), 15 November 2010, p. 11.

8. P. Achard, Rankings: A case of blurry pictures of the academic landscape? *Inside Higher Ed*, 21 September 2010.

9. S. Marginson, The power of rankings, *University World News*, Issue 0005, 11 November 2007.

10. D. Lederman, You think we're rankings-obsessed? *Inside Higher Ed*, 1 February 2010.

11. D. Lederman, 'Manipulating', er, influencing 'U.S. News', *Inside Higher Ed*, 3 June 2009.

12. C. Diver, Is there life after rankings? *The Atlantic*, November 2005.

13. R. Baggaley. How the RAE is smothering 'big ideas' books, *Times Higher Education*, 25 May 2007.

14. Introduction: A different kind of ranking, *Washington Monthly*, September-October 2013 (at http://www.washingtonmonthly.com/magazine/september_october_2013/features/introduction_a_different_kind046446.php).

15. A note on methodology, A Note on Methodology, *Washington Monthly*, September 2005 (at http://www.washingtonmonthly.com/features/2005/0509.methodology.html).

16. M. Clarke, News or noise? An analysis of *U.S. News and World Report*'s ranking scores, *Educational Measurement: Issues and Practice* 21, no. 4 (Winter 2002), pp. 39-48.