Ranking Universities: Criteria and Consequences

How to cite:

For guidance on citations see FAQs.
Introduction

This Supplement brings together a number of brief items addressing the issue of university rankings – or league tables, as we’ve come to know and love (or hate) them in the UK. The first piece explores some of the background to the debate: the impact of rankings, the main criticisms and broader issues of policy and principle. It arose from an initial review of the literature for current research on university league tables and their impact on institutional behaviour being undertaken by CHERI and Hobsons Research, commissioned by the Higher Education Funding Council for England (HEFCE). A brief description of the various strands of this research project follows. One of the weightier volumes on university ranking that has emerged recently is editors Sadlak and Lui’s The World-Class University and Ranking: Aiming Beyond, and a summary of this collection of papers is also featured in this supplement. Many of this book’s authors are members of the International Rankings Expert Group (IREG) and one of the Group’s key contributions has been to draw up a set of Principles on Ranking of Higher Education Institutions, with the explicit aim of evaluating and improving ranking practice. The Principles are included here as one succinct statement of good practice. But many critics question the very principle of creating hierarchies of institutions, and Ulrich Teichler provides a personal reflection on this subject to conclude this Supplement.

William Locke

Contents

League Tables: Impacts, Criticisms and Issues of Policy and Principle (William Locke) 2
HEFCE Commissioned Research on League Tables and Their Impact 3
The World-Class University and Ranking: Aiming Beyond Status (Sadlak and Liu) (William Locke) 4
The Berlin Principles on Ranking of Higher Education Institutions 5
Rankings: Do We Need an Anti-Meritocratic, Collectivistic and Anti-Dynamic Higher Education System? (Ulrich Teichler) 7

William Locke, Assistant Director, Centre for Higher Education Research and Information, Open University.
League Tables: Impacts, Criticisms and Issues of Policy and Principle

Over the last ten to fifteen years, league tables have become an increasingly established part of HE sectors around the world, and the UK is no exception. Often produced by commercial companies, the rankings have been identified as potential information sources for prospective university students, as strategic (and marketing) tools for institutions, as ‘selection guides’ for graduate recruiters, and as benchmarking measures for the sector as a whole. However, despite some evidence that they are being utilised by universities, students, companies and governments – and perhaps because of this – they remain highly controversial measurement tools.

The debate often centres around the robustness and inconsistency (as different rankings are based on different methodologies), and a lack of transparency in the rationale, accuracy and overall value of league tables. Within the sector, it may be commonly perceived that they reflect the value judgements of the compilers, but they still appear to sell large numbers of newspapers and guidebooks, whether or not these value judgements are made explicit. In a highly competitive environment, a higher education institution is seldom likely to avoid the temptation of citing a league table if it is ranked highly in it.

Some commentators have suggested that the mixture of input, process and output measures used by all the tables to provide overall scores means that, to a large extent, they are a measure of institutional prosperity. The majority of the quantifiable variables available for use in league tables are inputs: the qualifications of entrants, student-staff ratios, expenditure per student, spending on libraries and information and communications technology, the qualifications of academic staff and research income, for example. One researcher found that most of the variance in institutional scores in one high profile table could be explained by the research variance thus, potentially, rendering all other measures largely redundant. Other values may be more or less out of date, such as Teaching Quality Assessment/Subject Review scores and Research Assessment Exercise gradings. Even degree results and employment rates are subject to so many different factors – subject differences, the age of graduates, trends in the graduate labour market, etc. – that they are difficult for the lay person to interpret. Such outcome measures may not allow for input differences, and therefore cannot begin to assess ‘value added’, if this were felt to be desirable. Indeed, a recent analysis of the academic effectiveness and efficiencies of post-1992 and Russell Group universities in teaching and widening participation produced results that were the complete opposite of the usual positions of universities in published league tables. The researchers acknowledge, though, that altering the weighting given to the individual factors (or including research activity) could result in a different interpretation of the data.

A few critics have even suggested that the great majority of league tables fail the normal tests of reliability and validity, including statistical validity, required by social science. Input measures, process measures and output measures are often mixed together, without due consideration of the dynamics of quality processes. The measures used may not be robust or comparable between different contexts within an increasingly diverse higher education system. Institutions may have limited control over some of the variables included, and yet their performance is being evaluated by them.

There is limited evidence of students using league tables in making choices about higher education institutions and courses. However, there is evidence that institutions respond to league tables in order to improve their position, especially in international rankings. Whether these efforts actually lead to improvements in the quality or effectiveness of their provision is open to question. In some cases, they may actually detract from this. The HEFCE-commissioned research project described in this supplement will explore many of these issues.

Finally, there are broader issues of policy and principle. Are there public interests that are not represented in the league tables, and should public policy have a role in the development and dissemination of the ranking systems and, if so, what should this role be? Certainly, it is important that those producing league tables should be accountable for the quality of their data collection, methodology and representation, and good practice guidelines such as the Berlin Principles on Ranking of Higher Education Institutions, also included in these pages, are one means of achieving this. But, perhaps, as Ulrich Teichler argues in his very personal contribution to this supplement, ranking institutions promotes the idea that higher education systems should be vertically differentiated – i.e. organised on a hierarchical basis – in order to better contribute to the development of ‘the knowledge society’. And this may be a very problematic idea indeed.

William Locke

---

7 Brown, R (2006) ‘League tables – do we have to live with them?’, Perspectives, 10(2), 33-38.
HEFCE Commissioned Research on University League Tables and Their Impact on Institutional Behaviour

The Higher Education Funding Council for England (HEFCE) has commissioned research into newspaper league tables of higher education institutions (HEIs), how they are compiled and the impact they have on institutions. HEFCE has appointed Hobsons Research and CHERI to undertake together two major strands of research.

The first is an investigation of the methods and underlying data used in those rankings of most interest to English higher education institutions (HEIs): specifically the national rankings produced by the Guardian, Sunday Times and the Times and the world rankings produced by THES and Shanghai Jiao Tong University. This includes:

- semi-structured interviews with representatives from each of the compilers;
- further identification of source data, methodologies, calculations, weightings and approach to sensitivity analysis, and the assumptions and rationale for these; and
- exploration of conclusions drawn from the rankings and comparison with other, competing conclusions.

The second strand consists of small-scale case study research looking at how six higher education institutions are responding to league tables and whether or not they are taking steps to climb the rankings. This includes semi-structured interviews with key personnel and small group discussions with a selection of academic and other staff in each case study institution. Concurrently, an online survey of all English HEIs is gathering top-line data on the impact of league tables on institutions and their views on this.

Both strands of research are informed by the burgeoning literature on league tables and the emerging evidence of their impact on institutions. The findings from the research will be discussed with representatives from sector agencies (e.g. HESA, UCAS) and representative bodies (e.g. UUK and the individual interest groups and GuildHE). The project is due to be completed by March 2008 and the results will be announced at the HEFCE annual conference and published on the Funding Council’s web site in April 2008.
The World-Class University and Ranking: Aiming Beyond Status, Jan Sadlak and Lui Nian Cai (eds), 2007, Bucharest: UNESCO-CEPES, Shanghai Jiao Tong University, Cluj University Press

This volume consists of 25 essays on the emergence of the notion of ‘world-class’ universities and their ranking. Most of them originated from papers presented at the First International Conference on World-Class Universities organised by the Institute of Higher Education, Shanghai Jiao Tong University, held in June 2005. Some papers, particularly those dealing with academic ranking, were originally presented at the second meeting of the International Rankings Expert Group (IReg) on Methodology and Quality Standards of Rankings held in Berlin in May 2006. The main outcome of that meeting was the Berlin Principles on Ranking of Higher Education Institutions, which are included in this supplement and reproduced in the book.

The volume is divided into three parts: ‘Characteristics and Criteria of the World-Class University’, ‘Evaluation and Ranking of World-Class Universities’ and ‘Building a World-Class University’. This provides a broad set of perspectives on the conceptual framework and characteristics of ‘world-class’ universities and their ranking. Collectively, the authors evaluate a number of ranking methodologies, including bibliometric techniques, in order to arrive at more scientifically justified and verifiable outcomes. According to the editors, “…it is essential to inform and educate academia, policy-makers, the corporate sector and the public at large about the various ranking approaches and outcomes, in order to reduce the subjectivity and clichés with regard to the performance of universities and other higher education institutions” (p19).

However, the book is not just about ranking studies and their methodologies, it is also about the nature of universities deemed ‘world-class’ and the global competitive pressures and challenges of aspiring to such a status. Several of the chapters acknowledge that the concept of ‘world-class’ is not clear-cut, and that there are a different models of highly successful universities worldwide. They also emphasise that such a status is usually the result of intensive efforts sustained over a long period in a number of different fields arising from the consistent application of institutional values and approaches. Key to ‘world-class’, argues John Niland in Part One of the book, are reputation and perception (p61), and he outlines a series of qualities and strategies that one might expect to find in a university acclaimed as such: excellent faculty, research reputation, talented undergraduates, an international presence, extensive resources, strong alliances and networks, technological astuteness and good management.

One of the editors, Jan Sadlak, introduces Part Two on the evaluation and ranking of ‘world-class’ universities and argues that competitiveness and ranking are imperatives of the ‘knowledge society’ with higher education increasingly offering the creativity and innovation that is key to knowledge-driven economic success. Consequently, ‘world-class’ universities are in a global competition for the best talent among academics and students, and the recent emergence of world rankings reflect this. Anthony van Raan takes up the baton in his analysis of the Shanghai Jiao Tong University and Times Higher Educational Supplement rankings, identifying problems with their procedures, methods and data sources and offering an alternative approach based on more sophisticated bibliometric evaluation. Several authors agree that a limited number of universities are ever likely to gain entry to the top 200 institutions in the world on the measures employed by these two well-known rankings, favouring it seems large, multi-disciplinary, research-intensive institutions that are strong in science. Müller-Böling and Federkeil offer an alternative methodology for grouping (rather than ranking) German, Swiss and Austrian universities that focuses on 36 academic subjects offered by a substantial number of institutions, and providing a profile of indicators that applicants can choose and weigh according to their own preferences.

The third part of the book, on building a world-class university, starts with Ian Chubb’s observation that “What matters is often the hardest to define, to observe and to compare and, invariably, the most defiant of measurement” (p254). There follows accounts from the Australian National University, the University of Zurich, Uppsala University, Humbold University, the University of Bonn and the École Normale Supérieure of how quality can be defined and evaluated at institutional level and ‘world-class’ status maintained. Akiyoshi Yonezawa takes a broader view and evaluates initiatives by the Japanese government to strengthen the international presence of Japanese national universities, particularly in light of their poor showing in the THES World Rankings, with its emphasis on ‘peer review’. Zhao Chun-Mei analyses the relationship between university ranking and classification, using the Carnegie Classification, and identifies some unintended consequences for institutional policy and practice and the academic system as a whole. Finally, Philip Altbach introduces a note of caution for those aspiring to world status: it will probably require more than US$500m, clever leadership and much good luck. Most – even large and relatively wealthy – countries, he warns, can only sustain one or two such institutions which, in any case, will not be the best at everything they do.

“In sum…” Altbach argues, “…no one has figured out how to rank universities internationally, or even within countries in ways that are acceptable to the academic community or that can withstand serious critiques” (p367). Nevertheless, the authors of the chapters in this book have made a serious contribution to the debates around ranking, both in practical and critical terms. No doubt, these ranking studies will be developed further – and the critiques along with them.

William Locke
Rankings and league tables of higher education institutions (HEIs) and programs are a global phenomenon. They serve many purposes: they respond to demands from consumers for easily interpretable information on the standing of higher education institutions; they stimulate competition among them; they provide some of the rationale for allocation of funds; and they help differentiate among different types of institutions and different programs and disciplines. In addition, when correctly understood and interpreted, they contribute to the definition of “quality” of higher education institutions within a particular country, complementing the rigorous work conducted in the context of quality assessment and review performed by public and independent accrediting agencies. This is why rankings of HEIs have become part of the framework of national accountability and quality assurance processes, and why more nations are likely to see the development of rankings in the future. Given this trend, it is important that those producing rankings and league tables hold themselves accountable for quality in their own data collection, methodology, and dissemination.

In view of the above, the International Ranking Expert Group (IREG) was founded in 2004 by the UNESCO European Centre for Higher Education (UNESCO-CEPES) in Bucharest and the Institute for Higher Education Policy in Washington, DC. It is upon this initiative that IREG’s second meeting (Berlin, 18 to 20 May, 2006) has been convened to consider a set of principles of quality and good practice in HEI rankings – the Berlin Principles on Ranking of Higher Education Institutions.

It is expected that this initiative has set a framework for the elaboration and dissemination of rankings – whether they are national, regional, or global in scope – that ultimately will lead to a system of continuous improvement and refinement of the methodologies used to conduct these rankings. Given the heterogeneity of methodologies of rankings, these principles for good ranking practice will be useful for the improvement and evaluation of ranking.

Rankings and league tables should:

A) Purposes and goals of rankings

1 Be one of a number of diverse approaches to the assessment of higher education inputs, processes, and outputs. Rankings can provide comparative information and improved understanding of higher education, but should not be the main method for assessing what higher education is and does. Rankings provide a market-based perspective that can complement the work of government, accrediting authorities, and independent review agencies.

2 Be clear about their purpose and their target groups. Rankings have to be designed with due regard to their purpose. Indicators designed to meet a particular objective or to inform one target group may not be adequate for different purposes or target groups.

3 Recognize the diversity of institutions and take the different missions and goals of institutions into account. Quality measures for research-oriented institutions, for example, are quite different from those that are appropriate for institutions that provide broad access to underserved communities. Institutions that are being ranked and the experts that inform the ranking process should be consulted often.

4 Provide clarity about the range of information sources for rankings and the messages each source generates. The relevance of ranking results depends on the audiences receiving the information and the sources of that information (such as databases, students, professors, employers). Good practice would be to combine the different perspectives provided by those sources in order to get a more complete view of each higher education institution included in the ranking.

5 Specify the linguistic, cultural, economic, and historical contexts of the educational systems being ranked. International rankings in particular should be aware of possible biases and be precise about their objective. Not all nations or systems share the same values and beliefs about what constitutes “quality” in tertiary institutions, and ranking systems should not be devised to force such comparisons.

B) Design and weighting of indicators

6 Be transparent regarding the methodology used for creating the rankings. The choice of methods used to prepare rankings should be clear and unambiguous. This transparency should include the calculation of indicators as well as the origin of data.

7 Choose indicators according to their relevance and validity. The choice of data should be grounded in recognition of the ability of each measure to represent quality and academic and institutional strengths, and not availability of data. Be clear about why measures were included and what they are meant to represent.

8 Measure outcomes in preference to inputs whenever possible. Data on inputs are relevant as they reflect the general condition of a given establishment and are more frequently available. Measures of outcomes provide a more accurate assessment of the standing and/or quality of a given institution or program, and compilers of rankings should ensure that an appropriate balance is achieved.

9 Make the weights assigned to different indicators if used prominent and limit changes to them. Changes in weights make it difficult for consumers to discern whether an institution’s or program’s status changed in the rankings due to an inherent difference or due to a methodological change.

C) Collection and processing of data

10 Pay due attention to ethical standards and the good practice recommendations articulated in these Principles. In order to assure the credibility of each ranking, those responsible for collecting and using data and undertaking on-site visits should be as objective and impartial as possible.

11 Use audited and verifiable data whenever possible. Such data have several advantages, including the fact that they have been accepted by institutions and that they are comparable and compatible across institutions.

12 Include data that are collected with proper procedures for scientific data collection. Data collected from an unrepresentative or skewed subset of students, faculty, or other parties may not accurately represent an institution or program and should be excluded.

13 Apply measures of quality assurance to ranking processes themselves. These processes should take note of the expertise that is being applied to evaluate institutions and use this knowledge to evaluate the ranking itself. Rankings should be learning systems continuously utilizing this expertise to develop methodology.

D) Presentation of ranking results

14 Apply organizational measures that enhance the credibility of rankings. These measures could include advisory or even supervisory bodies, preferably with some international participation.

15 Provide consumers with a clear understanding of all of the factors used to develop a ranking, and offer them a choice in how rankings are displayed. This way, the users of rankings would have a better understanding of the indicators that are used to rank institutions or programs. In addition, they should have some opportunity to make their own decisions about how these indicators should be weighted.

16 Be compiled in a way that eliminates or reduces errors in original data, and be organized and published in a way that errors and faults can be corrected. Institutions and the public should be informed about errors that have occurred.

Berlin, 20 May 2006
I was pleased to be invited to a conference at the end of last year that brought together a group of world class experts in search of “world-class universities” with the help of rankings studies. Both the OECD and the Rectors’ Conference in Germany (HRK) had invited me to join these experts at a conference held in Bonn in December 2006 and to give a dinner speech, i.e. an occasion with no need to conform to the rules of academic discourse. In reproducing and slightly supplementing my presentation here, I take the liberty of formulating some issues more personally and more provocatively than we are accustomed to do in our daily academic affairs.

When I recently heard that members of those agencies in Germany that have been most active in producing rankings studies for almost a decade were claiming that they had started rankings in Germany, I was reminded that a considerable number of ranking studies had already been undertaken in the 1970s and 1980s. In fact, the first study of that kind which really seized public attention was published in 1978 shortly after I was appointed a professor. It produced a ranked list of 50 German universities (i.e. almost all except for small and highly specialised ones). It turned out that all three universities that were significant for me – the university from which I graduated, the university which awarded me a doctoral degree and the university employing me as a professor – were among the bottom five. In retrospect, I might claim that I have the single most disastrous CV of any German university professor.

In the mean time, we are all interested in rankings. The German Federal and Länder governments even decided recently to provide money for an “excellence initiative” in which ten top universities will be given preferential financial treatment for a period of five years in order to move even further ahead of the crowd. Both the university from which I graduated and the university which awarded me a doctoral degree were pre-selected among the top ten. At the same time, Sociology – my field – at the university where I am a professor was selected by the German agency undertaking ranking studies among the top ten. I do not hesitate to say that, regardless of this, our International Centre for Higher Education Research – my home turf – is viewed by our peers as world class. Based on that experience, I certainly can claim that considerable academic potential may not to be realised if resource allocation decisions are made according to the Matthew Principle.

In 1978 I found myself in a wonderful situation - I received two “calls” for professorial positions at the same time: from the University of Chicago (US), the most famous university in the world in my area of expertise (social-science orientated educational research), and from the University of Kassel (Germany), a newly established university hardly known anywhere at that time. I opted for the latter. In retrospect, one could argue that not only did I have the most disastrous CV, but I also made the most irrational career choice ever made. But I do not believe that it was an irrational decision under the conditions pertaining at the time. The University of Kassel provided exceptional support in allowing me to be instrumental in establishing a “centre of excellence”.

What were the conditions at that time in Germany? A relatively new university had at least sufficient means to support the conditions required for top-class research in selected areas. The presumed academic reputation hierarchy among universities was sufficiently flat: nobody had to believe that academic success was only possible with the help of the historical reputation of their predecessors. Agencies promoting research and other research sponsors wanted to provide financial resources, primarily, to excellent scholars or research teams wherever they were located. Last but not least, we were optimistic that, as individual scholars and teams of researchers, we were the carriers of academic quality and, if we were successful, we would lend our reputation to the institution.

If the conditions prevailing at that time were those that are today taken for granted or even promoted by the experts who undertake university ranking studies or who are advocates of their gospel, my career option in favour of a new university without established reputation would certainly have been suicidal. If I had to decide again, I should be happy to gain from the anti-meritocratic pre-occupation with the established fame of universities. We know the game of “credentialism” and “labelling” - to be at the right place is the most important thing. Nowadays, in allocating funds the agencies responsible are expected to be adherents of collectivism; to favour me if I am at a good place with good colleagues, and to penalise me for my not so excellent environment and colleagues. Under currently prevailing ideologies, and possibly under future conditions, I would have little chance of contributing to a dynamic rise in research quality without the strait jacket of the current hierachy among universities was sufficiently flat: nobody had to believe that academic success was only possible with the help of the historical reputation of their predecessors. Agencies promoting research and other research sponsors wanted to provide financial resources, primarily, to excellent scholars or research teams wherever they were located. Last but not least, we were optimistic that, as individual scholars and teams of researchers, we were the carriers of academic quality and, if we were successful, we would lend our reputation to the institution.

To complete my personal story: When my discipline at my current university was ranked among the top ten in Germany, I noted that it would have been rated substantially better – certainly among the top three – had significant numbers of sociologists at our university not been incorporated within interdisciplinary centres. Seemingly, the university could enhance its rank immediately if we had only disciplinary units, although we believe that our interdisciplinary approach was crucial for our world-wide recognition. But such adaptive strategies for success in rankings rather than for improved quality are truly abundant: We know that more German universities would rank among the “world-class” if the neighbouring Max Planck, Fraunhofer, Leibniz und Helmholtz institutes were incorporated. Some universities advocate such mergers: their motives might be genuinely in favour of quality, but suspicion is also aroused that ranking is seen to matter more.

Ranking studies vary substantially in so far as their conceptual basis and the methods and data employed are concerned. Of course, there are sound reasons to claim that research is so expensive in some areas that a high degree of physical concentration is indispensable. We also know that some students learn better in homogeneous...
than in varied environments. We know that the enormous expansion of higher education reinforces the public belief that quality can only flourish if protected in exclusive pockets. Last but not least, the advocates of ranking can argue that these studies are better than the inevitable myths and rumours. But I would not myself - nor would I recommend my colleagues to - embark on ranking studies, because I do not see any chance of these studies freeing themselves from the underlying anti-meritocratic, collectivistic and anti-dynamic paradigms.

I am not only lucky to know a vast range of ranking studies, but also to be personally acquainted with most of the key scholars and individuals involved. I respect most of them as honest scholars not only striving for high quality in ranking studies but also believing that they contribute ultimately to improving the quality of higher education itself. It is worth reading their own publications and overviews of the methodological issues arising from these studies. But I would advocate a reallocation of resources to studies helping to improve the state of our knowledge and understanding. Some of the huge amounts spent on ranking studies should instead be used to shed more light on two related issues. The first has already been mentioned: research on the (beneficial and detrimental) impact of the spread of ranking studies. The second issue may be even more important: rankings also seem to take for granted that the future of the knowledge society depends critically on the top universities, on the cutting-edge production and dissemination of knowledge, and that a steeply stratified higher education system is desirable. I plea for more research on the actual configuration of national higher education systems and on the various strengths and weaknesses of different degrees of vertical differentiation. And I would be pleased if we could enhance our forward thinking, i.e. our knowledge-based brainstorming about the character of the knowledge society of the future: are we moving towards, and do we need, an "elite knowledge society" or are we moving towards, and need – in addition or instead – the wisdom of the many?

Ulrich Teichler