

# UNIVERSITY RANKINGS: DIVERSITY, EXCELLENCE AND THE EUROPEAN INITIATIVE

GEOFFREY BOULTON

## LEAGUE OF EUROPEAN RESEARCH UNIVERSITIES

Universiteit van Amsterdam - Universitat de Barcelona - University of Cambridge - University of Edinburgh - Albert-Ludwigs-Universität Freiburg - Université de Genève - Ruprecht-Karls-Universität Heidelberg - Helsingin yliopisto (University of Helsinki) - Universiteit Leiden - Katholieke Universiteit Leuven - Imperial College London - University College London - Lunds universitet - Università degli Studi di Milano - Ludwig-Maximilians-Universität München - University of Oxford - Université Pierre et Marie Curie, Paris - Université Paris-Sud 11 - Karolinska Institutet, Stockholm - Université de Strasbourg - Universiteit Utrecht - Universität Zürich

## Author

The paper was written for LERU by Professor Geoffrey Boulton, Edinburgh University, [G.Boulton@ed.ac.uk](mailto:G.Boulton@ed.ac.uk). The paper has benefited from discussions in the Rectors' Assembly and comments from the members of the Research Policy Committee. The work has been supported by Dr. Katrien Maes (chief policy officer) and Mr Joost van Asten (senior policy officer).

## Summary

- International rankings of universities influence the perceptions and priorities of governments, of businesses and students. Rectors and university councils see the achievement of high ranking as a strategic imperative.
- However, their value and benefit is questionable. The fundamental problems are two-fold:
  - Most seek to capture characteristics that cannot be measured directly, and require indirect proxies. How good are the proxies?
  - Different universities fulfill different roles, which a single monotonic scale cannot capture. How can different roles be compared in meaningful ways?
- None of the current ranking systems have the validity, rigour or meaning to be of real value, except those based on citations to evaluate research, and even here, they fall short in assessing research in the humanities and the social sciences.
- Institutions tend to target a high score irrespective of whether the metrics are good proxies for the underlying value of the institution. Rankings will at best be irrelevant to those values or, at worst, undermine them. They encourage convergence towards a research-dominated model, reducing system diversity and undermining the potential to contribute to society in other ways.
- But rankings have such a hold on the public imagination that they are likely to be permanent features of the landscape. Can they be improved? Two approaches have been funded by the European Commission.
- **U-Map** is an attempt at classification describing the diversity of universities by mapping activities, not quality: its purpose being transparency for stakeholders.
- **U-Multirank** is an attempt at ranking evaluating quality in dimensions analogous to those of U-Map: its purpose assessing how well universities perform their different roles, rather than holding all to research-dominated criteria.
- Both have serious defects. They suffer from imprecise proxies and the profound difficulty of finding comparable data between countries. The temptations will be to:
  - require ever more burdensome detail in the hope of penetrating to the heart of the matter,
  - formalise the distinctions that mapping reveals,
  - promote further the idea of the university as merely a source of modular products currently in vogue.
- LERU applauds the attempt to create U-Map as a description of diversity, but is less enthusiastic about U-Multirank, because of problems of data comparability, the potential for game playing when reputations for excellence are at stake, and for the encouragement it gives to target proxies rather than underlying reality.
- Although the U-Multirank approach at least provides a means of exemplifying diverse forms of excellence as an antidote to single monotonic lists, it is inevitable that its individual dimensions will be combined by others to create a single monotonic table of excellence, regardless of the strictures of its authors.
- However, given the likely persistence of “league tables”, LERU supports the Commission’s initiative to develop U-Multirank as a pilot project and as a means for exploring its potential to mitigate the problems of other systems. In this spirit, a number of LERU universities have agreed to collaborate in the project with the intention of improving it as far as is possible.

## The emergence of international rankings of universities

1. The advent of “league tables” of university excellence, first produced in 2003 by Shanghai Jiaotong University, was perhaps the inevitable consequence of the convergence during the 1990s of liberalisation of international markets, enabled by new communications technologies, and the shift of the global economy towards one based on information and knowledge. In this setting, universities, as sources of innovative ideas and highly skilled manpower, have come to be seen as vital agents in maintaining national competitiveness. A metric that purports to show the stock and presumed rate of creation of new intellectual capital in those institutions, universities, whose role it is to conserve and create it, becomes the equivalent of a stock market quotation, and also an index of national intellectual prestige.

## Impacts of rankings

2. Rankings have had a dramatic impact on perceptions of university excellence. They have been dominated by US institutions, with the latest (2009) Shanghai tables showing eight out of the top 10 and 37 out of the top 50 ranked universities from the USA<sup>1</sup>. Europe’s universities, similar in number to those of the USA, perform relatively poorly, with only two in the top 10 and ten in the top 50.

3. Amongst governments, the rankings indicated the extent to which their universities were achieving the excellence presumed to be needed to drive and support national economic prowess. Germany, for example, launched the Excellence Initiative, and the French Presidency of the EU advocated the need for a European system of university ranking that would be more sympathetic to the European university ethos and character. Amongst universities,

the intention to target a high standing in rankings has been included in many institutional mission statements and been integral to many of their strategies. Prospective students, particularly those that move abroad for study, are increasingly using rankings as a guide to choice, and international businesses are increasingly aware of them in seeking out university partners<sup>2</sup>.

4. The Shanghai ranking, which is largely based on achievements in research, stimulated a plethora of imitators. The Times Higher Education journal started to produce rankings from 2004, which purported to rank universities based on attributes wider than research, and has announced the intention to invest heavily in developing its system further. Leiden University has produced a citation-based ranking strictly limited to research performance<sup>3</sup>. The Lisbon Council has produced a ranking of university systems<sup>4</sup> based on criteria such as inclusiveness, access, effectiveness and responsiveness, in which Australia is ranked most highly, UK second, Denmark third and the USA fifth. And the European Union has funded a project called U-Map to characterise and map the diversity of institutions<sup>5</sup> and a project called U-Multirank to rank excellence within these diverse categories<sup>6</sup>; an effort to which we will return later in the paper.

5. If rankings could be created that accurately reflected the diverse values of universities to their societies, they would in theory be valuable in several ways: to university managers in benchmarking their universities against an international scale and identifying where improvements should be sought; to students and academics in matching their choices of where to study or work to their aspirations; to public and private bodies seeking links to universities that would further their objectives; and to governments in helping to align their policies for universities to national needs. The delivery of such benefits is conditional however on the capacity of rankings to measure the values of universities to their societies. If they are poor measures, and yet remain powerful drivers of behaviour, there is a

<sup>1</sup> Academic Ranking of World Universities. 2009. <http://www.arwu.org/ARWU2009.jsp>

<sup>2</sup> Hazelkorn, E. 2007. *The Impact of Global Rankings on Higher Education Research and the Production of Knowledge*. Higher Education Management and Policy, 19(2).

<sup>3</sup> Leiden World Ranking. 2010. <http://www.socialsciences.leiden.edu/cwts/news/the-leiden-ranking.html>

<sup>4</sup> The Lisbon Council. 2008. *University Systems Ranking: Citizens and Society in the Age of Knowledge*. <http://www.lisboncouncil.net>

<sup>5</sup> Centre for Higher Education Policy Studies. 2010. *U-Map: the European Classification of Higher Education Institutions*. <http://www.u-map.eu/>

<sup>6</sup> U-Multirank. <http://www.u-multirank.eu>

serious risk of their delivering more damage than benefit. In this they face two major problems:

- that many of the features they seek to measure cannot be measured directly, but depend for their evaluation on indirect proxies; leaving the question of how good are the proxies?
- that universities now vary greatly in the diverse functions that they are called upon to perform in society, and how therefore can a single, monotonic scale be an accurate measure of institutions that have different roles?

## The problem of inappropriate proxies

6. The two dominant international rankings, those of Shanghai Jiaotong University and the Times Higher, have been subject to strong technical criticism. The fundamental problems are two-fold: how to identify measurable proxies for university activities, for example in education, in outreach, in innovation etc.; and how to combine different categories of activity, for example for education, research and knowledge exchange in ways that have any meaning? We are not persuaded that any of the current ranking systems have sufficient validity, rigour or meaning to be of value, except those that restrict themselves to the use of citations in evaluating research, and even here, they fall short in assessing research in the humanities and to some extent in the social sciences. Measurement error is either large or indeterminate, with the concordance between the 2006 rankings by Shanghai and the Times Higher being modest at best, with only 133 universities shared in their top 200 lists.
7. An analysis of the Shanghai rankings by Billaut et al.<sup>7</sup> using the tools of multiple criteria decision making, suggested that the criteria that are used are not relevant, the aggregation methodology is

flawed, and the overall analysis suffers from an insufficient attention to fundamental structuring issues. They concluded “that the Shanghai ranking, in spite of the media coverage it receives, does not qualify as a useful and pertinent tool to discuss the ‘quality’ of academic institutions, let alone to guide the choice of students and families, or to promote reforms of higher education systems.”

8. A commercial publisher, Thomson Reuters, is to collaborate with the Times Higher Education<sup>8</sup> journal to further develop the Times’ rankings. In addition to research citations and largely research biased data, these have included a worldwide survey of academic reputations. Whereas this might be argued to be a means of creating a more holistic view of universities, it lacks credibility. As internationally experienced academics, although we may know individual departments that are strong in our individual area of research, we are unlikely to know, unless having spent time there, about the educational and intellectual environment. We are very unlikely to have a rigorous sense of that university as a whole, and exceedingly unlikely to be able to make comparisons between 10, 50 or even 100 universities. Such approaches are most likely merely to reinforce existing, conventional stereotypes. We will be interested to see how reputation surveys are weighted against other criteria in the new Times’ rankings.
9. Rankings that purport to measure the overall excellence of a university are inevitably skewed to benefit particular types of institution, largely because of the accessibility of data from citations and related sources about performance in natural science, engineering and medicine. The absence of robust, numerical and discriminatory proxies for education, for the humanities, to some extent for the social sciences, and for other activities such as societal impact, inevitably undervalues these activities. High ranking is most heavily dependent upon research performance in science, particularly in life sciences and medicine, on large size, on income and to some extent institutional age. If these are the

<sup>7</sup> Billaut, C-D., Bouyso, D. and Vincke, P. 2009. *Should you believe in the Shanghai ranking? An MCDM view*. HAL-Articles on line.

<sup>8</sup> Times Higher Education. 29 April, 2010. *Data details unveiled for the THE world rankings*. 3 June, 2010. *THE unveils broad, rigorous new rankings methodology*.

biases, the question arises whether they are or should be the dominant indices in measuring the excellence of a university and its contribution to society.

## The dilemma of diversity

10. Focus on a single, monotonic ranking implies one, largely research-based, criterion for excellence, which masks the excellence of many institutions in generating benefit for society. The idea that such rankings identify the “best” universities is a travesty. They allocate high rankings to those universities that score highly in research in the natural sciences, life sciences and medicine, though not in the humanities and social sciences, and ignore the achievements of those for which internationally competitive basic research is not a first priority. They create a pattern of esteem that has two perverse effects: it encourages students to apply to highly ranked universities rather than ones that might better suit their needs; and it encourages institutions to converge towards a single, research-dominated model, thereby reducing the diversity of a university system and undermining the potential of many to contribute to society in other ways.

## Rankings and university behaviour

11. Where a metric of institutional behaviour is developed, and where a high score is important for the reputation or material benefit of the institution, a rational institution that does not have a strong sense of purpose, will almost invariably target a high metric score irrespective of whether the metrics are good proxies for the underlying value of the institution for society. Universities have proved to be extremely adept at playing this game.

12. A consequence of the reputation race stimulated by rankings is illustrated by a recent survey conducted by the European Commission which showed that across Europe, at least 980 universities proposed, in their mission statements, to achieve a high level of

international excellence in research<sup>9</sup>. It reflects both an unrealisable aspiration and a lost potential for many other areas where universities bring benefit to their communities. The diversity of higher education institutions in Europe is a strength, not a weakness. They are a response to a great and growing diversity of societal demands for knowledge, understanding and technical competence that is beyond the capacity of any one institution to satisfy. Pressures that diminish that functional diversity of institutions, or narrow the focus of even research-intensive universities to science research, drive them inexorably away from their true role in society. If ranking proxies are poor measures of the underlying value to society of universities, rankings will at best be irrelevant to the achievement of those values, at worst, they will undermine it.

13. If research results, particularly in science and its applications, were the dominant benefits that universities offered to society, then rankings that largely reflect this would be sensible ways to stimulate their research efforts. However, LERU has argued strongly<sup>10</sup> that this is a myopic view, which isolates research, and particularly research in science, from the totality of the university enterprise, yielding a fundamentally flawed perspective of the benefits that universities have for society and how they are delivered. Although research is a critical contribution to the core enterprise of the university, to elevate it, and particularly to elevate the performance in certain areas of science, to be the principal determinant of university reputation, as rankings tend to, is to create pressure on universities and their heads to make them a university’s principal priority. The consequence has been that research may now be emerging as the enemy of higher education rather than its complement. In these circumstances, it is important that universities have a stronger sense of themselves rather than being driven by external perceptions or rankings that implicitly embed a flawed perception of their role.

<sup>9</sup> Statement from the European Commission’s Education Directorate-General. Dean Conference, Edinburgh, September 2006.

<sup>10</sup> Boulton, G.S. and Lucas, C. *What are universities for?* League of European Research Universities. September 2008.

## Living with rankings

14. We are not so naive however as to believe that it would be easy for university heads to resist the pressures exerted by rankings in the face of their influence on popular perceptions of reputation, or pressures from university councils or governments. We accept that the popularity of a simple, comprehensible monotonic ranking scale is such that it is likely to maintain a hold on the popular imagination, to the detriment of more discerning analyses that dig more deeply into institutional character.
15. The desires to avoid the perversities of monotonic rankings, to resist the pressure to converge on a single research-obsessed model of higher education and to advertise the strengths of a diverse system, led the European Commission to fund two major feasibility programmes. These are a European Classification of Higher Education Institutions (U-Map), designed to map the diversity of European universities, and U-Multirank, designed to produce international ranking of university performance in dimensions analogous to those used to map diversity.

## A European initiative to categorise diversity - U-Map project

16. U-Map starts from the premise that the large variety of modern demands on higher education systems require a diversity of institutions to satisfy them and that these will have different attributes that need to be recognised by students who seek institutions adapted to their needs, public and private bodies that wish to engage with universities, and governments that need to understand patterns of differentiation of universities for policy purposes.
17. The completed U-Map project proposes a system in which institutional profiles are built up from metrics that measure activities in six dimensions: the teaching and learning profile (e.g. levels and orientations of degrees, subject range), the student profile (e.g. mature, distant, part-time), research activity, knowl-

edge exchange, international orientation and regional engagement. They are then displayed so that the relative magnitudes of these dimensions are apparent for individual institutions (although we have been unable to determine how relative magnitudes are determined). The system also permits institutions that match a given profile to be identified. In order to minimise the bureaucratic burden of compilation on HEIs, it is suggested that data compilations should initially be made from national statistics, which are then verified or amended by institutions. It is proposed this system should be implemented for Europe through an independent body funded by public or private organisations, or by contributions from institutions that figure in the U-Map system.

18. In principle, the objectives of this scheme are admirable, particularly for those, either students or businesses, seeking information about distant institutions, which they are unlikely to know well, although it would be surprising if national governments were not aware of the diversity of their own institutions and the issues that they pose for the structure of their national system. There are however a number of major issues that need to be recognised:

- Is there a need, or is it another expensive tentacle of the audit culture? Is there evidence that there is a lack of “transparency” about HEIs in Europe that inhibits either potential students or potential collaborators in making sensible choices that is sufficient to justify creation of a costly and time-consuming enterprise? The European Commission clearly believes that there is<sup>11</sup>, although it provides no evidence to support that view.
- Such a mapping system could help heads of institutions and their boards or councils to focus on their strengths and distinctive roles within their national systems and at European level. Is it designed with their needs in mind?
- If U-Map is implemented it is likely to have a strong influence on policy. This needs to be considered carefully. The temptation for governments and policymakers will be to institutionalise the distinctions that mapping reveals and to create

<sup>11</sup> European Commission. 2010. *Assessing Europe's University-Based Research*. EUR 24187 EN.



functional and funding systems that restrict further evolution. Whilst implementation of U-Map could combat one of the current dangers in Europe, that of system convergence towards a research-driven model, institutionalisation of diversity could lead to fossilisation.

- Although U-Map addresses the dilemma of diversity as we have described above, the problem of inappropriate metrics remains, with the temptation to seek ever more detailed proxies, with the process they seek to measure remaining elusive.
19. There are also important issues that need to be addressed about the nature of the data in U-Map and its collection:
- The U-Map project has effectively asked a group of universities to determine the indicators that they wish to have used to characterise them. If U-Map is to be rolled out to a wider group of institutions, it must be ready for the “data war” that is likely to develop. For example, the draft report rejected such research-related indicators as the proportion of staff with doctoral degrees and the number of graduate degrees as a proportion of degrees awarded, whilst there are no explicit indicators of doctoral education (expenditure, number of graduates, graduate schools) and many other research-related indicators on the grounds that these reflect a quality dimension. Whilst in narrow sense this might be true, (although the semantic problems of distinguishing activity from quality are great, probably impossible in many areas) other areas of activity/quality overlap are included in U-Map, such as the number of spin-out companies or the number of non-European students, which happen to be important indicators for post-graduate students, researchers and companies who may be seeking “transparency” about university characteristics.
  - There is a major issue of the comparability of data from country to country. If U-Map is to be used on a significant scale by those making genuine international inquiries about choices of institutions, data incompatibility will seriously distort their choices. Moreover, if HEIs are permitted to amend their data, marketing departments are likely to pick and choose the data they wish to be judged on. The problems multiply further if there is an attempt to extend U-Map beyond Europe.

20. Whereas the U-Map scheme does correct one of the

pathologies of monotonic rankings, their blindness to diversity, it carries another danger with it. It must not be used to imply that a university is simply a series of separate functions that can be viewed, chosen or evaluated separately. This is a perception or a reality that sees universities merely as instruments in satisfying a number of market demands. It is a perception that diminishes them. At its worst “mapping” can be a populist trick that makes universities subject to “food labeling”, in which they are seen to deliver highly defined consumer products, rather than understanding for a lifetime. It underestimates the extent to which universities can be a construct of their staff and students and not a purveyor of products. LERU has argued strongly that universities at their best are an interactive whole, and that much of their value to those who work or study in them, or those who engage with them, comes from those interactions.

## A European project to recognise and rank diversity - U-Multirank

21. Whereas clarity about the focus of activity in an HEI may be important in providing information that potential students or staff, and public or private bodies may need, the critical question they ask about an HEI is not so much “what does it do”, but “how good is it at what it does?” Governments and policy makers tend to ask similar questions about international standing for the reasons given in paragraph 3, whilst university heads and their boards or councils are concerned with the university’s standing and market position as reflected in international rankings.
22. In view of the criticisms of the validity of monotonic ranking schemes and their potentially damaging consequences (paragraphs 6-9, 11-13), including their blindness to institutional diversity, the European Commission has funded a project designed to create global rankings for the range of dimensions embedded in U-Map (education, research, knowledge transfer, internationalisation and community engagement). The indicators for each dimension are designed to act as proxies for excellence in that dimension. For example, for the dimension of education, the proxies are: expenditure on teaching, time to degree, graduation rate, graduate unemployment, graduate earnings, staff qualifications, student-staff ratio, external experience of



academic staff, gender balance, diversity policy, internet access, student satisfaction (in computing, libraries, rooms, teacher support, course content, programme organisation), work experience, employability issues, graduate satisfaction and labour market relevance of qualifications.

23. However, just as quality assurance in education is not a proxy for the excellence of education, but rather the excellence of process, neither are the indicators in 22. Their danger, as we have pointed out above, is that these proxies become the objectives of education rather than how learning and true education take place. We are similarly sceptical that the indicators for knowledge transfer and community engagement come close to reaching the reality of these processes. This is not to be critical of the attempt to measure the underlying reality, but to recognise the difficulty of doing so.

24. The possible **benefits** of international rankings that are more discriminatory than current monotonic schemes lie in their potential:

- to stimulate improvements in performance in those areas where the university is of real benefit to society;
- to assist potential students and those wishing to engage with a university in choosing an institution that is excellent in the domain and in the geographic area of interest;
- to provide governments with tools for assessing national effectiveness in higher education.

25. The **difficulties** of achieving these rational outcomes are:

- the fact that the proxies used to assess excellence rarely measure the reality of excellence or plumb the true benefit of universities to society, except perhaps in the area of research;
- that statistics from country to country are not comparable, with the potential to contribute gross interpretation errors;
- that the potential for game playing by institutions is great;
- that individual dimensions are open to the same criticisms that can be made of monotonic rankings.

26. The **dangers** inherent in the approach are:

- that it encourages universities to target the proxies rather than the underlying reality, and depends fundamentally on how effectively the proxies represent that reality (paragraphs 6-9);
- that the large data demands of the system may create a major enhancement of the accounting culture in universities, adding to the detailed regulations, memoranda, instructions, guidance and lists of good practice that flood into institutions, replacing the creative freedom and flexibility that is required for good education and research with the obsession to measure and monitor everything;
- that it further promotes the idea of the university as a supermarket selling modular products that happen currently to be in vogue (paragraphs 12 & 20).

## Conclusions

27. It would be easy to caricature the perspectives of this paper as those of the ivory tower, detached from the exigencies of modernity and reacting against a brave new world of efficient, realistic and evidence-based management. It is rather that whilst we accept that both the U-Map and the U-Multirank projects have potential value, there are considerable difficulties and dangers which might outweigh the value of the projects. We have fewer concerns in relation to U-Map. It recognises the dilemma of diversity, though it must still struggle with the unresolved problems of proxies. The latter is mitigated however by the fact that it is about *what* is done, rather than *how well* it is done. Overall we believe that it could be a useful tool provided that it can be compiled without an excessive burden on contributing institutions, and that it can access data that is reasonably robust and avoid excessive control of data by institutions.
28. Our response to U-Multirank is different. We are sceptical of the utility of ranking tables and of the unintended impacts that they may have. We are also deeply concerned about the massive data gathering exercise of a system that attempts to develop more precise proxies for many diverse functions, and will depend for its discriminatory power on the comparability of inter-country statistics. However, the U-Multirank approach at least provides a means of exemplifying diverse forms of excellence as an antidote to single monotonic lists. Given that the obsession with “league tables” will not go away, LERU supports the Commission’s initiative to develop U-Multirank as a pilot project and as a means to better understand the potential value of international rankings of universities. It is appropriate that it is further developed, notwithstanding its difficulties and dangers. In this spirit, a number of LERU universities have agreed to collaborate in the pilot project with the intention of improving the system as far as possible. Realistically, it is of course inevitable that its individual dimensions will be combined by others to create a single monotonic table of excellence, regardless of the strictures of its authors.

## About LERU

LERU was founded in 2002 as an association of research-intensive universities sharing the values of high-quality teaching in an environment of internationally competitive research. The League is committed to: education through an awareness of the frontiers of human understanding; the creation of new knowledge through basic research, which is the ultimate source of innovation in society; the promotion of research across a broad front, which creates a unique capacity to reconfigure activities in response to new opportunities and problems. The purpose of the League is to advocate these values, to influence policy in Europe and to develop best practice through mutual exchange of experience.

## LERU publications

LERU publishes its views on research and higher education in position papers and advice papers.

Position papers make high-level policy statements on a wide range of research and higher education issues. Looking across the horizon, they provide sharp and thought-provoking analyses on matters that are of interest not only to universities, but also to policy makers, governments, businesses and to society at large.

Advice papers provide targeted, practical and detailed analyses of research and higher education matters. They anticipate developing or respond to ongoing issues of concern across a broad area of policy matters or research topics. Advice papers usually provide concrete recommendations for action to certain stakeholders at European, national or other levels.

LERU position and advice papers are freely available in print and online at [www.leru.org](http://www.leru.org).

Universiteit van Amsterdam  
Universitat de Barcelona  
University of Cambridge  
University of Edinburgh  
Albert-Ludwigs-Universität Freiburg  
Université de Genève  
Ruprecht-Karls-Universität Heidelberg  
Helsingin yliopisto (University of Helsinki)  
Universiteit Leiden  
Katholieke Universiteit Leuven  
Imperial College London  
University College London  
Lunds universitet  
Università degli Studi di Milano  
Ludwig-Maximilians-Universität München  
University of Oxford  
Université Pierre et Marie Curie, Paris  
Université Paris-Sud 11  
Karolinska Institutet, Stockholm  
Université de Strasbourg  
Universiteit Utrecht  
Universität Zürich



#### LERU Office

Huis Bethlehem      tel +32 16 32 99 71      www.leru.org  
Schapenstraat 34      fax +32 16 32 99 68      info@leru.org  
B-3000 Leuven  
Belgium