

Chapter 13

The Future of University Rankings

Ulrich Teichler

13.1 The Amazing ‘Popularity’ of University Rankings

In analysing the abundant meta-evaluative literature on university rankings, we note an impressive range of arguments and analyses of the concepts, methods, results, perceptions and possible impact of such kinds of activities to put individual institutions somehow ‘on a map’. In various areas of research, we often note that a certain approach draws substantial attention because it is viewed as very ambitious and promising and as deserving further enhancement through the involvement of the brightest scholars in the respective field. In the case of university rankings, however, most experts would agree that the great attention paid to this domain by many experts is not an indication of respect for high quality analysis. Rather, rankings draw attention as a consequence of a seemingly paradoxical mixture of conceptual and methodological weakness on the one hand and political power on the other to influence the views of the ‘map’ of higher education and to elicit activities aimed at changing the existing ‘map’.

The expert literature obviously does not spend much time and energy on defining rankings. ‘Rankings’ exist, and efforts to clarify their definition would be futile, given the lack of precision of what is meant by that term. However, we note that ranking studies are usually described as data presentations with three general features:

- Ranking studies are activities of vertical sorting. Rank lists or scales are established according to ‘very good’, ‘high-quality’, ‘excellent’, ‘world class’, ‘renowned’ or whatever the positive end in higher education might be called.
- Ranking studies carry out an inter-institutional comparison. Higher education institutions or their sub-units (departments, etc.) are compared, as a rule, across all higher education institutions within a country, region or worldwide.
- Ranking studies provide information with the help of relatively short lists of quantitative measures for ranking and rating the units to be compared.

U. Teichler (✉)

International Centre for Higher Education Research Kassel (INCHER-Kassel)
University of Kassel, Kassel, Germany
e-mail: teichler@incher.uni-kassel.de

Even the exercise of defining rankings in a most simple way shows how highly normatively loaded the activities are. Are vertical differences so important compared to horizontal differences (e.g., the ‘profiles’ of individual institutions or units) that attention is appropriately concentrated on the vertical dimension of the higher education ‘map’? Are rankings in fact instruments that disregard or even undermine the importance of horizontal diversity? Are institutional aggregates such as those of higher education institutions as a whole or their sub-units really the key carriers of quality? Are data on the quality of institutions more or less artificial aggregates of the very heterogeneous quality of academic work of individual scholars or groups of scholars within each institution which is only marginally influenced by its local environment?

Experts agree that ranking institutions is not a recent phenomenon; there is a long history of rankings. Higher education in the United States in the twentieth century has been viewed as highly successful and has served as a role model for higher education in many other countries. As well, quantitative educational measurement has been more popular in the USA than in most other countries for a long time, so it does not come as a surprise to find that many consider the USA the breeding ground of ranking studies. In fact, the first national ranking studies in the USA can be traced back to the 1920s and the first worldwide popular ranking study of universities was published in the USA in the 1980s by US News & World Report. However, there are other countries with a long tradition of national rankings. For example, rankings have played a more important role in Japan than in the USA, where competition for entry into specific Japanese universities is fiercer and where the diversity of higher education is more strongly viewed as vertically shaped with only a limited role of horizontal diversity. Consequently, there have been many ranking studies carried out in Japan over many decades. Since 1995, Asahi Shinbunsha, the publisher of the most highly regarded national newspaper, has been regularly publishing a synopsis of all available university rankings in Japan. In the 2005 edition of ‘Daigaku ranking’, the results of 717 ranking studies were presented.

In efforts to identify the driving forces for the emergence and spread of rankings, it is often pointed out that pervasive secular trends in higher education have been the major ‘push’ factors for the greater attention paid to university rankings since the 1990s. Three major trends are most often referenced:

- Massification of higher education
- Increased competition
- Internationalisation of higher education

Certainly, mass higher education, competition and international interaction must be considered in this framework. Mass higher education is viewed widely as a push factor for the stratification of higher education. When, for example, 25% of an age group graduate from higher education, the distinction between the top quintile and the second highest quintile of higher education might be functionally equivalent to the distinction between higher education and non-higher education in the past, when only 5% of students graduated from higher education. Also, the fiercer the

competition, the more attention given to the issue of whether a university is number 75 or number 80 in a ranking list. Finally, worldwide rankings are only of interest if higher education systems are not nationally segmented.

There are good reasons, however, to challenge the emphasis on secular trends when the popularity of rankings ought to be explained. In some countries, higher education had already been highly stratified before it had reached the stage of expansion commonly named 'mass higher education'. Fierce competition between higher education institutions, between students and possibly between other stakeholders had also existed in some countries before concepts of managerialism and entrepreneurialism as basic features of higher education governance at individual higher education institutions spread globally. Even with regard to 'internationalisation', there is a need to be more precise in identifying the dimensions according to which 'world-class universities' had been highly international in the past and the dimensions according to which we note an increasing internationalisation in the recent two decades.

There is another arena of discussion about the potential driving forces for the increasing attention paid to university rankings. While the arena named above is characterised by an historical analysis of the functional change of higher education, the other is shaped by psychological observations and political reflections about the paradox of rankings. It raises the question of how and why are such vague and simple measures employed to rate or rank the quality of academic work which is possibly the most complex and sophisticated feature to be assessed? Is there a hidden or overt 'virtue' in the primitiveness of information systems in higher education? How widely are the explicit or implicit ideologies of the producers of university rankings shared by the actors in higher education or its environment? Are academics so attracted by the 'excellence versus mediocrity' perspective that they consider horizontal diversity as marginally relevant at best? Do politicians believe that the quality of academic work and its relevance for society will improve if greater pressure is put on academics to follow the main stream? Does society believe in an elite knowledge society with a high concentration of academic expertise in a few locations, or does it believe in a knowledge society characterised by a spread of knowledge?

13.2 Towards a Typology of Meta-Evaluative Views of University Rankings

The discourse on the strengths and weaknesses of university rankings can be characterised as a complex interaction of methodological and functional arguments. Nobody can claim therefore that a single classification of the various types of meta-evaluative arguments clearly surpasses the quality of any other classification. In choosing a classification of arguments here, the guiding principle is to identify possible implications for the future.

The first type of argument is that the 'success story' of rankings is fundamentally based on their primitiveness. It is 'sexy' to get a quasi-objective confirmation of rumours such as conventional wisdom, surprises, new gossips, etc. Feelings of superiority and inferiority, heroism, virtue versus sin or shame, condemnation and absolution, 'excellence' and triviality are aroused. The less one knows and the less one cares about the quality of university rankings, the more one can enjoy the ranking 'games'. Of course, one has to trust blindly that there is a certain minimum authority behind the rankings. Journalistic evidence rather than academic evidence seems to suffice.

The second type of argument is based around a pragmatic discourse on the normal state of the 'quality' of indicators. Macro-societal indicators similarly defined and employed worldwide can be viewed as powerful instruments of 'transparency' and 'comparability'. We easily rely on indicators such as 'growth domestic product', 'unemployment' and 'rates of educational attainment' of the adult population. We know that there is a discrepancy between measurement with the help of a 'proxy' and the real character of the phenomenon mirrored by the indicator. We tend to accept the 'proxies' pragmatically because otherwise we would be confined to ignorance or guess work. But even if one takes such a pragmatic point of view in accepting complexity-reducing simple indicators as best possible proxies and as the 'least bad way' of measuring reality, one has to 'admit' that the rankings of 'world class universities' have not achieved the status of such general pragmatic acceptance. The expert discourse on university rankings suggests that even the simplest indicator of quality in higher education cannot be based on a single measure, but rather has to be an aggregate of various measures. Moreover, the expert discourse shows that no minimum consensus has emerged as regards a small list of key indicators to be included in an aggregate measure. However, we note a considerable readiness in higher education policy and practice as well as among higher education experts to accept the pragmatism of indicators. Citations of articles published in peer-reviewed journals are often referred to in the public discourse as a good indicator for quality of research. This often is done without any conceptual and methodological caveats that this can be interpreted as an indication that there is a readiness to accept relatively simple indicators, even in the intellectual ambitious and complexity-conscious academic environment.

The third type of argument concentrates on the assumed positive impact of rankings on higher education. 'Transparency', 'healthy competition', 'creative concentration of talents' and similar arguments are put forward.

The fourth type of argument concentrates purely on data improvement. For example: How can we reduce the problem of missing data?; How can we increase the response rate in reputation surveys?; How can we reduce institutional misallocations of authors in citation?; How can we ensure a more or less identical definition of foreign academic staff and students?; Are the definitions of staff and students similar in the statistics of the various countries and institutions?; etc. Even the argument that ranking lists should be replaced by a vertical classification of grades of quality can be viewed as a purely methodological argument. If the vertical differences between individual ranks are so small that they justify a ranking order,

a classification of 'outstanding', 'very good', 'good', etc. would be understood as purely methodological improvement.

Most advocates of ranking studies do not limit their critique of the current state of rankings to purely methodological weaknesses. Rather, the fifth type of argument focuses on the 'validity' of rankings. The term 'validity' is employed in the ranking discourse if one accepts the prevailing philosophy of university rankings while calling for new or improved indicators closer to the reality to be indicated. Experts who call for a better 'validity' of rankings often believe in or at least accept the presumed virtue of a vertical sorting of institutional aggregates in higher education with a small list of indicators as creating desirable 'transparency' and contributing to 'healthy competition'. The following question addresses the need for better 'validity' without challenging the philosophy of rankings. How could indicators be operationalised so that they are not biased against certain disciplines, that they do not discriminate against small institutions, that they take care to strike a balance of the core functions of higher education (teaching, research and possibly service) and that they do not disregard different national conditions of higher education (for example, defining research quality not only by measuring the quality of the texts published in the English language)?

The sixth type of argument focuses on deficiencies of the prevailing ranking studies which are unlikely to be redressed in the framework of the prevailing ranking philosophies and ranking practices. The following are examples of critique that go beyond the intentions and the potentials of the prevailing ranking milieu. Rankings provide information on assumed quality differences, whereby their causes and the possible improvement remain a 'black box'. Rankings do not take into account the 'value-add' achieved by the higher education institutions. Rankings claim that input, processes and output are closely linked, or that achievements in teaching and research are closely linked without taking into consideration the actual extent of linkage or dissociation. Rankings neglect horizontal diversity and are useful if an institution strives for 'fitness of purpose' which does not represent the main stream. Rankings claim to serve the 'transparency' for varied purposes, although different kind of information is needed for varied purposes: government might need information for 'accountability', the university management for priorities of research promotion or for strategic choices, or for the improvement of the organisational effectiveness. Students, as well, need other types of information and again other types of information are needed by possible partners of research and technology transfer. This kind of critique does not call into question the potential value of empirical data on higher education as feedback. Often, it is based on even higher expectations as regards the utility of a good data base. Rather, they consider the selection and presentation of data in the customary ranking studies as a distortion or under-utilisation of the potential of information.

The seventh type of argument might be characterised as a fundamental critique of rankings. In addition to the critique of possible biases and distortions vis-à-vis the reality of the higher education system, critique is most frequently voiced in the context of the possible or actual adverse impact ranking systems. Such impact could include undermining the extent of horizontal diversity, a discouragement of

unconventional approaches in research and teaching and an ‘over-competition’ which destroys potential and discourages the losers. Other consequences could be an ‘over-concentration’ of high quality resources in certain places that may lead to only small gains through concentration and to serious losses everywhere else. Moreover, rankings can undermine meritocratic reward by stifling the advantages of the historically privileged institutions and of the winners due to symbolic advantages and successful short-term tactics. Finally, rankings might mobilise the above-average institutions while having a zero-effect or even discourage below-average institutions which need appropriate feedback in order to improve. Whatever the distortion of the data, they are likely to elicit even higher distortions as a vicious circle of mal-information and adaptive behaviour. As a rule, experts voicing such a fundamental critique of rankings leave it open as to whether they believe in possible improvements of systematic empirical information as feedback for higher education, or whether they consider such efforts as futile because rankings are bound to produce distorted information.

13.3 The Possible Futures of Rankings and Beyond

We have experienced many rapid changes in higher education over the last few decades, and many changes were not predicted beforehand. As a result we do not feel confident in predicting the future as far as university rankings are concerned. But we can suggest some likely scenarios.

First, we may experience a situation best described as the inertia scenario. If an element of higher education has been present for quite a while or has emerged in recent years, it is ‘here to stay’, as many advocates of university rankings point out. Interest in vertical lists of universities and a belief in their quality and virtues will be too stable to challenge them, and there is no evidence that this feature will be not protected by the widespread system inertia in general.

The second type of possible scenario could be the trend scenario. In several countries, there is a long tradition of rankings. In recent years, interest in rankings has spread to other countries. As ranking construction increases and greater attention is paid to rankings worldwide, this trend will be reinforced by other trends such as massification, increasing competition and internationalisation. As a result, a further spread of rankings can be expected.

Third, we suggest a ‘politics works’ scenario. Rankings may change higher education that the way the ideologists of the ranking movement hope and primary critics of higher education rankings fear. For example, horizontal diversity may become more or less irrelevant for higher education, and the competition for the highest possible rank according to relatively homogeneous criteria might become even more pervasive.

Fourth, we can imagine an emerging problem and emerging problem awareness scenario. For example, if main stream rankings are biased towards the research functions, there is potential for serious problems in the quality of teaching and learning.

This in turn may lead to major steps being taken to redress this deficiency, including a reform of an information system which tends to elicit undesirable adaptations.

Fifth, it is also worth considering alternative scenarios where a more desirable higher education system emerges. The recent spread of ‘diversity management’ is one example of newly emerging paradigms which could challenge the basic concepts underlying rankings. In this scenario, inter-institutional (vertical) diversity concurrent with a relatively high intra-institutional homogeneity would become the most desirable and productive and the idea might spread that intra-institutional diversity will be a matter of fact and will be the most productive future of higher education.

Sixth, there is no reason to exclude the potential for a ‘turn towards high quality rankings’ scenario. The methodological optimists may turn out to be right in saying that those who are involved in the production of rankings and those who are involved in the funding of rankings or of other information systems on which rankings are based, are willing to strive for a higher complexity of rankings, as well as for a broader concept of ‘validity’ than those which are now in place.

Seventh, at least for the sake of logical completeness, we suggest the ‘increasing complexity of balanced information’ scenario. Open, thorough and unbiased feedbacks, as well as evidenced-based strategic action may become so highly appreciated in a knowledge society that biased systems of information gathering can no longer overshadow other relevant information systems. What role would remain for rankings if a scenario of impressive transparency and rationale actors became a reality?