



World Education Services

University Rankings

www.wes.org

World Education News and Reviews - August 2006

University Rankings

Introduction

Emloyees at World Education Services are asked on occasion how they view a certain university or college in relation to its peers. Our response, more often than not, is simple: does the institution in question have official recognition or accreditation from the relevant regulatory body in its home country? Beyond this criterion we make no further judgment call. All institutions that do have accreditation or state recognition are deemed to be of equal merit. This is not to say, however, that the quality of instruction from institution to institution in any given country, or even from department to department within the same institution, doesn't vary greatly; it is simply to say that we are not in the business of making that particular assessment.

Those seeking to distinguish different levels of quality within a large national education system or across borders would have to look beyond accreditation or government recognition – a process which, broadly speaking, assures that the institution in question meets certain minimum standards of academic quality, but says very little beyond that. In the United States, for example, accreditation from a regional accrediting organization indicates that an institution meets set minimal standards for its faculty, curriculum, student services and libraries with little extra qualitative information made available. This lack of transparency and dearth of comparable information has made the annual rankings performed by *U.S. News & World Report* (USNWR) a highly (many would say disproportionately) influential tiebreaker for parents and students trying to make a decision on university or college study options.

In the absence of governmental initiatives, and emboldened by the commercial success of the USNWR rankings — first produced in 1983 — a plethora of commercial media publications and research organizations around the world have stepped in to bridge the information gap. Indeed, the last decade has witnessed a veritable explosion in the number of ranking systems available to consumers. Today domestic rankings are produced in more than 20 different countries, while two well-known institutions now produce international rankings. Furthermore, rankings are increasing in sophistication and minutia, with institutions as well as programs receiving performance grades and ranks.

The growth in domestic and international rankings can be attributed to two major factors: the massification of tertiary education and globalization. The current generation has witnessed governmental policies that have opened higher studies to much broader sections of society while increasing levels of personal wealth and liberalized visa regulations have greatly increased transnational student mobility. As a result, today's college-age student is exposed to much greater opportunity and choice, and it has largely been left to the private sector to produce the comparative information needed to inform that choice.

While commercial rankings have proven popular with prospective university students and their parents, they have also been cause for a fair degree of scorn and criticism among university faculty and administration, many of whom point to weak research methodologies.

Beyond Commercial Rankings

Although media remains the dominant force in university rankings, there is a movement among research organizations and professional associations to advance the utility and objectivity of the

ranking exercise. As a result there are now an increasing number of rankings produced by not-for-profit enterprises.

Government agencies can also be seen to have entered the fray, and in a number of countries universities are comparatively assessed for funding purposes, as part of the accreditation process, or as a means of stimulating competition and encouraging institutional quality improvements.

Given the relative influence that university rankings now have over student preferences, it seems fair that the rankings themselves should be subject to critical examination and commentary. One initiative seeking to address some of the common criticisms has brought together ranking experts from both the commercial and non-commercial worlds. Meeting for the third time in May 2006, the International Rankings Expert Group have defined a list of best practices in university rankings that they hope will evolve into something of a self-regulation process (see below).

While far from exhaustive, the following is offered as a regional and country-by-country guide to the world's university ranking systems. We have, where available, provided results from the most recently published rankings often times comparing results across different publications, in addition to outlining methodologies and, where appropriate, criticisms.

What follows is a brief overview of the most common ranking methodologies.

Approaches to Ranking

“WEIGHT AND ADD”

If commercial media outlets are at the forefront of the ranking business, then the “weight-and-add” approach is the lead methodology used by most ranking systems, and the one familiar to most consumers. This methodology, which was first introduced by USNWR and has subsequently been reproduced, tailored and tinkered with by many others since, ranks institutions on an ordinal scale according to an aggregate overall score. This overall score is derived from a number of weighted criteria, which in turn are measured by a range of sub-indicators acting as proxies for academic or institutional quality.

The weightings that are applied to each category and indicator are generally based on the subjective (and, some might say, arbitrary) opinions of researchers as to the relative importance that each indicator is deemed to represent. The weighted scores for each indicator are then added to produce an overall mark or score. This overall score, when compared against other universities, will decide a school's final position on the ranking, or league table. The number of criteria, indicators and their weightings vary from ranking to ranking although the basic model varies little. One slight variation on the theme is the ranking of institutions by department, program or subject area. The more comprehensive rankings produced today tend to include both overall institutional rankings and departmental or subject rankings.

While the weight-and-add approach is undertaken by a majority of rankers and is the approach familiar to most consumers, other methods and approaches exist, and have often been introduced in response to criticism of existing ranking models. The introduction of departmental and subject rankings, for example, came in response to criticism that while whole-of-institution rankings might provide an insight into the overall prestige of an institution, they offer little practical guidance to students wishing to learn about the best available programs in their particular discipline.

Criticisms

The weight-and-add approach and its practitioners have long been subject to a number of criticisms, the most common of which can be summarized as follows:

1. There is an almost infinite number of measures that can be defined to gauge quality in higher education, and even the most intricate ranking system will fail to capture them all. It is therefore the job of the ranker to define a set of criteria and weightings by which to measure quality. Considering there is no widely accepted set of criteria, rankings will be open to criticism based on the indicators employed and the weighting formula applied.
2. Many privately published annual ranking systems, which cynics say are driven by the commercial need to sell more publications, often have wildly fluctuating results from year to year because the weightings and criteria are changed. This means that year-on-year comparisons, if made, can be very misleading
3. Ranking systems often result in a large number of institutions clustering around a median score. Schools that score toward the lower end of the cluster by a statistically insignificant margin from those toward the top end of the cluster receive a ranking that appears to grade them as being of significantly poorer quality. The argument, therefore, is that final rankings tend to exaggerate the differences borne out by the actual statistical findings.
4. Rankings are something of a self-fulfilling prophecy. If reputation is considered a significant factor in overall quality, those institutions that have traditionally fared well on previous rankings will have their reputation positively affected next time out and vice versa.
5. Institutions, in whose interest it is to place well on rankings, provide much of the data collected by certain ranking systems. This has led to the criticism that institutions engage in data manipulation in an effort to impact their final ranking; after all, the smallest statistical changes can often mean the difference of a place or two on a ranking.
6. Many data collecting exercises are driven by the information that is available rather than the information that is necessary to accurately gauge the level to which an institution or department meets a particular quality criteria. Rather than engaging themselves in rigorous data collection exercises, many ranking systems default to accepting whatever information is publicly available, especially if they have to meet annual deadlines.

Responding to Criticism: Building International Consensus on Quality Indicators and Ranking Methodology

16 PRINCIPLES

In an effort to address some of the concerns listed above, and also to inject a greater degree of transparency into the ranking process, an international group of researchers, publishers and higher-education experts gathered in Berlin in May 2006 to build consensus on a set of principles for ranking colleges and universities¹. The meeting was convened by Romania-based UNESCO-CEPES and Washington-based Institute for Higher Education Policy, and conclusions from the meeting were summarized in a document known as the *Berlin Principles on Ranking of Higher Education Institutions*. The Berlin Principles have been offered up as a set of voluntary guidelines for those in the business of producing rankings. It is hoped that the guidelines will serve as the inspiration for a future self-regulation process.

The document outlines four broad areas for consideration in the production of rankings: purpose and goals of the ranking, design and weighting of indicators, collection and processing of data, and presentation of ranking results.

Among the 16 principles are recommendations that rankings should: recognize the diversity of institutions and take into account their different missions and goals; be transparent regarding the

methodology used for creating the rankings, measuring outcomes, such as retention and graduation rates, in preference to inputs, such as scores on admissions examinations, where possible; use audited and verifiable data whenever possible; and offer consumers a choice in how rankings are displayed, such as by allowing them to determine how factors are weighed on interactive web sites.

All 16 principles can be viewed at: www.cepes.ro/hed/meetings/berlin06/Berlin%20Principles.pdf

ALTERNATIVES TO THE WEIGHT-AND-ADD APPROACH

With something of a consensus building in recent years that ordinal weight-and-add rankings are somewhat flawed and can often be misleading, a number of ranking initiatives have sought to offer alternative methods of comparison. One innovation aimed at addressing a number of perceived shortcomings in the weight-and-add approach is the introduction of ranking systems that group or cluster institutions and departments within certain bands, while also allowing consumers the choice of manipulating the weightings of criteria most important to them.

“Group and Cluster”

Based in Germany, the Center for Higher Education Development has developed a ranking system in conjunction with *Die Zeit* and the German Academic Exchange Service that allows the consumer much greater latitude in defining his or her search for academic quality than under traditional all-of-institution rankings.

Instead of combining individual indicator scores to produce an ordinal league table of institutions, CHE assesses individual departments from universities in Germany, Austria and Switzerland with each quality indicator standing independently. Departments are assessed as being in the 'top third,' 'middle third,' and 'final third' as compared to peer departments in that particular discipline and for each particular indicator. Departments are therefore clustered rather than ranked for each particular indicator and considered equal to their cluster peers.

Consumers are then invited to create their own ranking by choosing the discipline and quality indicators that are important to them. Those indicators are then sorted by the website to provide comparative information on the various institutions offering programs in the consumer's field of interest.

The reasoning behind clustering is a belief that "league positions can be dangerous because minimal differences produced by random fluctuations may be misinterpreted as real differences. By contrast, the ranking group approach ensures that the top and bottom group differ to a statistically significant degree from the overall mean value."

By partnering with *Die Zeit* and DAAD, CHE is able to promote its rankings to both a domestic and international audience. *Die Zeit* publishes the rankings annually in a special supplement and also maintains a dedicated website while DAAD publicizes the ranking to an international audience and makes the information available in English through its own dedicated website.

A number of other ranking systems have adopted this tailor-made methodology, most notably in Europe, and many of the methods used by the Germans have been adopted by the International Rankings Expert Group as examples of best practice enshrined in the *Berlin Principles on Ranking of Higher Education Institutions*.

In production since 1998, this ranking innovation has received a fair degree of attention, and CHE *et al* have been contracted by ranking organizations in other countries to produce similar systems.

Latent Variable Analysis²

Researchers from the RAND Corporation devised a model using latent variable analysis as part of a research project aimed at identifying top-quality institutions of higher education for Qatari students on government scholarships. The model seeks to address three critiques commonly leveled against university rankings:

1. How does one properly account for uncertainty in the assignment of ranks to different institutions?
2. When are measured quality differences between institutions statistically significant?
3. How important is each ranking criteria to the overall quality ranking of the institution?

In answering these questions, the model seeks to provide a measure of how certain institutions are similar or dissimilar while also allowing the possibility to differentiate among institutions in a meaningful way. The third question is designed to reveal the factors that make institutions similar or dissimilar.

Using the same data provided by The (London) *Times Good University Guide* and *U.S. News & World Report* rankings, the latent variable technique ranks universities according to their deviation from the overall mean in each category. Where an institution clearly scores higher on all measured criteria, it will be ranked with a high degree of certainty; where an institution has very similar scores to a number of other institutions across a number of the different criteria it will still be ranked, but with a much lower degree of certainty. Using data from the 2005 *Times Good University Guide*, only four universities can be ranked in the top 10 with a high degree of certainty. The schools ranked 6 to 10 have a probability ranking that suggests that they could rank as low as 20.

The latent variable method provides important insights into the heterogeneity that exists between institutions of higher education and also suggests the use of caution in asserting that certain institutions are of a higher quality than others based on entirely subjective quality indicators and weightings. In assigning no preset weighting to each criteria, the model allows testing for statistically significant differences among institutions by placing greater emphasis on those criteria that have greater derivations from the mean for each particular institution.

GOVERNMENT INITIATIVES

In addition to commercially produced rankings, there are also countries around the world that are attempting to inject a degree of comparability into the way they go about accrediting their institutions. In India, for example, the National Assessment and Accreditation Council (NAAC) applies a grade to the outcomes of the institutional evaluations that its inspection teams perform. Although the grading system has been changed on a number of occasions, and the process as a whole has been open to plenty of criticism, it does provide a degree of comparability that can be used by those at home and abroad wishing to distinguish the best Indian institutions from their more mediocre counterparts.

In Britain, research funding is distributed according to a peer-review exercise known as the Research Assessment Exercise conducted approximately once every five years. The results of this exercise, conducted across more than 60 disciplines, provide an insight into research quality at Britain's universities and colleges; however, it says little about teaching at the undergraduate

level. The Tertiary Education Commission conducts a similar exercise, known as Performance-Based Research Funding, in New Zealand.

1 The 2006 meeting built on the findings and papers from an original meeting in June 2002 in Poland, which was the first of its kind to address the need to analyze the strengths and weaknesses of the various approaches to ranking. A follow-up meeting was held in Washington in December 2004. The next meeting of the so-called International Rankings Expert Group is scheduled for the fall of 2007 in Shanghai.

2C. Guarino, G. Ridgeway, M. Chun, and R. Buddin (2005). "A Bayesian latent variable model for institutional ranking," *Higher Education in Europe* 30(2):147-165.

© 2006 World Education Services
Bowling Green Station
P.O. Box 5087
New York, NY 10274-5087
USA