

# Basic explanation of key enhancements in methodology for 2007

---

Ben Sowter, Head of Research, QS – 22 October 2007

## Background

The THES – QS World University Rankings began in 2004 and have attracted a great deal of comment, reaction and feedback since first publication in October of that year. Since that time the project has assimilated a great deal of new ideas and had evolved into a stronger, more robust measure of comparative international university quality. Inclusion of the employer review in 2005 and the increased response to the Peer Review questionnaire are examples of these enhancements. The process continues... aside from improved response rates in both survey elements there are four key developments in 2007 that have had an impact on results. In as simple terms as possible, this document outlines those changes.

### **DEVELOPMENT: Peer reviewers prevented from promoting their own university.**

Since the inception of the rankings the Peer Review has been the centrepiece of the ranking, thus even the smallest alteration to its compilation can have a major effect on the overall performance of institutions. In the first three years of the rankings, no restrictions have been placed on universities identified as excellent by peer reviewers meaning that universities could potentially encourage their own academics to sign up and complete the questionnaire in their favour. Whilst there has been no evidence to suggest a deliberate assault by any single institution in this respect, as the awareness of both the ranking and the science behind it has become more widespread, it has become necessary to eliminate a reviewer's own university from the list they are presented with in the questionnaire.

### **EFFECT**

The effect of this development will be most profound where the peer review has received a particularly impressive volume of response in a country in comparison to other countries in the same region. The effect is likely to be further exaggerated if that country has a small number of institutions in the original list. An academic from the University of Arkansas is perhaps less likely to select their own institution (one of over 60 in the US) than an academic from Nanyang Technological University in Singapore (one of just two in Singapore).

### **DEVELOPMENT: Switch to Scopus from ESI (Thomson) for citation data**

For the 2006 results the time period for citation counts was slashed from 10 years to 5 years in response to feedback suggesting that the rankings ought to be a more contemporary measure of university strength. In 2004, when the rankings began, the only reputable source of citation data was Thomson's Web of Science – the ESI is an associated, simplified product that provides an indication of research strength by university and seemed the most appropriate basis for our citation indicators at the time.

Coincidentally, Scopus was also born in 2004 and has rapidly evolved since that time. In 2007, Scopus has been able to answer many of the questions left unanswered in three years of working with ESI – we have been able to find data for many institutions that have not been represented in this indicator in the past and we have also been able to query the entire Scopus database rather than simply the slices of Web of Science represented by ESI. The general consensus in published reviews of both systems (e.g. [Fingerman 2006](#)) seems to be that they both have their merits and can be used to complement one another. The vast majority of any criticism for Scopus seems to relate to its tracking of research and, in particular, citations from before 1996 but – since we

are only concerned with the most recent complete 5 year window – any weaknesses in this respect have no bearing.

#### **EFFECT**

1. The Scopus database has a less pronounced bias towards the US, resulting in a reduced advantage in their favour in this indicator
2. Scopus covers a larger number of papers and journals overall leading to greater representation from lesser known universities and institutions from academic systems with less emphasis on publication
3. Scopus covers more sources in languages other than English resulting in better numbers for institutions with large volumes of high quality research in their own language

#### **DEVELOPMENT: Consistent usage of Full-Time Equivalent (FTE) data for all personnel related data**

When requesting “Number of Faculty” from a university, the diversity of response can be surprising. Indeed some institutions respond having read the question as “Number of Faculties”. Each year, the QS research team has attempted to eliminate ambiguity by making the definitions sent out in the data request increasingly precise. In 2007, for the first time in an attempt to minimise confusion institutions were asked to provide both a Headcount and an FTE figure, where the FTE figure was to be utilised for the ranking. Where an FTE number has not been supplied, one has been extrapolated based on the relationship between Headcount and FTE numbers at other institutions in the same country or region. This ought to provide us with a much more accurate picture of the scale of institutions and provide a stronger logic to our selection of indicators.

#### **EFFECT**

This will have its most profound effect on institutions with large numbers of part-time faculty or students that may have been either over or under represented in key indicators in previous years.

#### **DEVELOPMENT: Z-score aggregation of indicators to generate overall scores**

In previous years, scores for each indicator have simply been scaled against the top performer in that indicator – the top institution awarded 100 and subsequent institution’s scores scaled against that maximum. There are some problems with that approach:

- The smallest of errors in the top institution’s data can cause dramatic “ripple” effects through the entire data for that indicator
- The curve for each indicator is exceptionally steep, leading to an emphasis on excellence in any given indicator that supersedes the weighting for that indicator. For example, the difference in score represented between coming first and second in “International Students” could represent the equivalent of coming 150<sup>th</sup> as opposed to 200<sup>th</sup> in the Peer Review, despite the fact that the weighting is only 5% as opposed to 40%.

A Z-score, also called a normal score or standard score, is a common and accepted method of standardising or normalising statistics and has been utilised for many years in certain domestic rankings of universities. There is a good technical explanation on the method on Wikipedia ([http://en.wikipedia.org/wiki/Standard\\_score](http://en.wikipedia.org/wiki/Standard_score)).

#### **EFFECT**

- Institutions demonstrating excellence in indicators with a low weighting will have less advantage (e.g. London School of Economics for International Students)
- Small errors, if present, in top performing institutions’ data may influence their own position, but will have negligible influence on the whole dataset
- The curve for each indicator will be smoother resulting in a more functional application of the weightings

- Extreme results will not undermine the weighting for any given indicator (Caltech scored 100 to Harvard's 55 for Citations per Faculty in 2006 essentially halving the influence of the indicator for all remaining institutions)
- Significant changes to an institution's performance in any one indicator will now only have a proportionate and predictable influence on their overall position, thus... THE RANKINGS WILL BECOME MORE STABLE year on year.