# SEARCHING FOR THE GOLD STANDARD: THE *TIMES HIGHER* EDUCATION WORLD UNIVERSITY RANKINGS, 2010-2014

**Richard Holmes**
*University Ranking Watch*
*E-mail: rjholmes99@yahoo.com*

## ABSTRACT

This paper analyses the global university rankings introduced by Times Higher Education (THE) in partnership with Thomson Reuters in 2010 after the magazine ended its association with its former data provider Quacquarelli Symonds. The distinctive features of the new rankings included a new procedure for determining the choice and weighting of the various indicators, new criteria for inclusion in and exclusion from the rankings, a revised academic reputation survey, the introduction of an indicator that attempted to measure innovation, the addition of a third measure of internationalization, the use of several indicators related to teaching, the bundling of indicators into groups, and most significantly, the employment of a very distinctive measure of research impact with an unprecedentedly large weighting. The rankings met with little enthusiasm in 2010 but by 2014 were regarded with some favour by administrators and policy makers despite the reservations and criticisms of informed observers and the unusual scores produced by the citations indicator. In 2014, THE announced that the partnership would come to an end and that the magazine would collect its own data. There were some changes in 2015 but the basic structure established in 2010 and 2011 remained intact.

**Keywords**: *rankings, universities, Times Higher Education, citations*

## INTRODUCTION [1]

The classification and ordering of colleges and universities are not new. As long ago as 1908 the Carnegie Foundation for the Advancement of Teaching created a list of 14 American institutions that were ranked according to the amount of money spent on instruction and would become the basis for Edwin Slosson's Great American Universities (Ballantyne, 2002). In the 1980s, the ranking of universities took on a new dimension when the US News & World Report (USN) started America's Best Colleges. This met a growing need for guidance for students across the United States who could no longer rely on advice from friends, parents and teachers (Wildavsky, 2010).

Rankings went international, although not yet global, when the Hong Kong based magazine, Asiaweek, published a ranking of Asian universities in 1999 with a second edition in 2000. This was a broad based index that assessed universities according to academic reputation, student selectivity, faculty resources, research performance, income, staff student ratio, number of postgraduate students, citations and internet bandwidth ("Asia's best universities", 1999, 2001) At the end of 2001, however, the magazine ceased publication something that its publisher, Time Incorporated, blamed on a decline in advertising revenue (Schwartz, 2001).

The origins of the first really global rankings, the Academic Ranking of World Universities (ARWU) produced by the Center for World-Class Universities at Shanghai Jiao Tong University, go back to 1998 when the Chinese president declared that the country needed a number of world-class universities and Shanghai Jiao Tong University was one of those selected for this status. Liu Nian Cai, then a professor in the Chemistry and Chemical Engineering department at the university, worked on benchmarking with American universities, and this eventually led to the first ARWU in 2003 based on publicly accessible research data (Liu, 2009).

The Shanghai rankings caused quite a stir. They showed that the world of scientific research was dominated by English-speaking universities, especially those in the USA, while continental European institutions generally did poorly. In 2011 Edouard Husson, a former advisor to the French government, spoke about the consternation when no French institution reached the top fifty of the rankings (Myklebust, 2011).

A year after the first appearance of the Shanghai rankings, two more appeared. The Webometrics rankings used several Internet based indicators and had the virtue of ranking many more universities than ARWU or later rankings but the academic world and the public were not convinced that it was an accurate representation of quality.

In 2004, John O'Leary, then the editor of Times Higher Education Supplement (THES, renamed Times Higher Education (THE) in 2008), and Martin Ince came out with the first edition of the THES-QS World University Rankings. Ince, who was in charge of the rankings project, has said that THES was looking for international metrics that would capture academic quality, graduate employability, research, teaching and world reach (Ince, 2010). The indicators and their weighting were decided by THES but the collection of data was done by Quacquarelli Symonds (QS), a British company that until then had specialized in recruiting students for graduate business programmes.

Almost every aspect of the THES-QS rankings had some sort of problem. Among them was the collection of data from institutions to calculate the faculty student ratio, the percentage of international students and faculty and the number of citations per faculty. In many countries such data are not easy to come by.

It was also claimed with some justification that the citations per faculty indicator was rather crude since it took no account of disciplinary peculiarities with regard to the frequency and timing of citations and that the faculty student ratio was not a good proxy for teaching quality or the resources available for teaching. The latter was also subject to fluctuations as institutions and QS made errors, corrected errors and struggled with unclear definitions and administrative reorganization. The blog University Ranking Watch (Holmes, 2006a, 2006b, 2007) revealed how Duke University and École Polytechnique experienced dramatic changes in their positions as a result of errors in calculating this indicator.

It was, however, the academic opinion survey that aroused the strongest complaints. The response rate to the survey question was very low, probably because the database was full of duplicated names. Furthermore, it came out that the respondents were required to have no

more expertise than the ability to sign up for an online subscription. Over the years since 2004 the academic survey and the employer survey, which had a 10% weighting and was introduced in 2005, have evolved but few observers have much confidence in them.

In 2008, QS introduced a raft of changes that included standardization so that the curves for all indicators had the same shape, using Scopus rather than the Thomson Reuters (TR) databases as the source of citation data, and imposing restrictions on survey respondents voting for their own institutions.

Methodological changes combined with errors and correction of errors, changes in the distribution of survey responses and the entry of new universities into the rankings with concomitant changes in the means from which indicator scores were calculated ensured that the THES-QS rankings were unstable, with some universities rising or falling dozens of places over the course of a single year. It should be noted, however, that the QS world rankings have become more stable over the last few years (Holmes, 2014). There was also concern about the QS Stars system where universities pay QS for an audit that results in the award of one to five stars.

Despite methodological changes, criticism mounted and penetrated into the power structure of the THES, which was evolving from a traditional weekly newspaper into a magazine. After a change of ownership in 2005, the publication was renamed Times Higher Education in 2008. During 2009 the THE editorial team reviewed the rankings and in October announced that it was ending its partnership with QS and would turn to TR to produce a new set of rankings (Baty, 2010a, 2010b, 2010c, 2014).

The new THE rankings were an ambitious project, purporting to be "robust, transparent and sophisticated" and "the most exact and relevant world rankings yet devised" (Baty, 2010d, para. 1). It became clear, however, that there was a contradiction between the need for consistency and stability and the drive for accuracy and validity. While THE did win the approval of the elite universities of the UK and continental Europe, it is debatable whether they had in fact created rankings that could accurately be described as exact and relevant.

## THE FIRST EDITION OF THE TIMES HIGHER EDUCATION WORLD UNIVERSITY RANKINGS 2010

The new THE rankings that finally emerged in the autumn of 2010 had several distinctive features.

### Choice and Weighting of Indicators

Previously, the weighting of the various indicators appeared to be arbitrary or perhaps influenced by the convenience or commercial interests of the publishers. THE, however, went to considerable lengths to consult with a wide range of views and approaches. During the months that followed the separation from QS, they began a series of discussions and consultations about the form that the new rankings would take. In a comment on 1 December 2009 at the University Ranking Watch blog, Phil Baty, THE Rankings editor, wrote that THE wanted to "to start from scratch and develop a new rankings methodology in direct consultation with the international university community" (as cited in Holmes, 2009, n. p.). THE opened an online survey for academics and administrators, started a platform group to consult with major universities and tapped the advice of its editorial board, which included figures such as Philip Altbach, head of the Centre for International Higher Education at Boston College, Drummond Bone, a consultant on international higher education and Bahram Bekhradnia, director of the Higher Education Policy Institute (Holmes, 2009).

The structure of the rankings that eventually emerged in the autumn of 2010 was influenced by the opinions of academics at highly regarded institutions and seemed to mark a shift towards research intensive, industry linked, internationally orientated universities that emphasized doctoral supervision rather than undergraduate teaching or taught master's courses.

THE also considered input from some of those who had been criticizing the rankings. Phil Baty, for example, has gone on record as taking account of the suggestions from the editor of University Ranking Watch that the weighting given to teaching related components be increased and that for international students reduced. He also noted that Ian Diamond of the University of Aberdeen had been a keen supporter of

field normalization of citations data and that the new rankings would take note of his views (Baty, 2010a).

Other factors came into play. The original intention was that the weighting given to reputation surveys should be reduced but it turned out that institutional data was not always reliable and so the reputation survey with two questions about postgraduate teaching and research ended up with 34.5% of the total weighting. Income from industry, which at one point was supposed to have 10% of the weighting, was in the end given only 2.5% because many universities were unable to come up with the relevant data.

At the beginning of 2010, THE published the first draft of its proposed new methodology, which was then sent out to members of the advisory board and the platform group.

After this round of consultation, a revised structure emerged, which and it was announced on 2nd September. The first difference was that weighting of the income from industry indicator was reduced from 10% to 2.5% because it was self-reported data and hence not always reliable. In 2010 and in later editions of the rankings, it was noticeable that several US universities, including in 2014 the University of Minnesota, Indiana University and UCLA were given blank spaces for this indicator and received an adjusted score based on the other indicators.

The 55% allocated to research was now increased to 62.5% divided between 30% for Research: Volume, Income and Reputation, which included five indicators and 32.5% for Citations.

To allocate nearly a third of the total weighting to citations was unprecedented and to choose only one of the many ways of measuring citations was surprising. It is a reasonable inference that the interests of TR, whose InCites system for staff evaluation depended on the calculation of impact factors normalized by field and year, played a disproportionate role in the choice and weighting of indicators.

## Inclusion and Exclusion

Ranking organizations have taken different approaches to the number of institutions that qualify for ranking. At one extreme is Webometrics, which now ranks over 25,000 universities, using publically accessible data. Other rankers have been much more selective.

When the Shanghai ARWU rankings were developed, the Center for World-Class Universities deliberately relied only on publicly available information and took no account of whether any university wished to be in the rankings or not. The Shanghai rankers started with all universities whose staff or alumni had won Nobel prizes or Fields medals, employed a highly cited researcher, published a paper in Nature or Science or had a significant number of papers in the Science Citation Index-Expanded or the Social Science Citation Index.

The THES–QS rankings adopted a somewhat different approach. They started with the top three hundred universities by research output then added a number of universities that were thought worthy of inclusion because they published significant research in languages other than English and those included in the Asiaweek rankings. After that, universities have been considered on a case by case basis but once included are not allowed to withdraw. Should a university decide not to submit data then QS uses old data or data from websites, government agencies or third party sources (Sowter, 2008). The number of ranked universities has risen from 500 in 2004 to 566 in 2007, 834 in 2013 and 907 in 2015 (Topuniversities, 2009, 2013, 2015).

TR and THE decided that universities must submit current data if they wished to be included in the THE rankings and there were no substitutes for self-submitted information. They were taking a risk here since it was precisely this policy that contributed to the end of the Asiaweek rankings and it did seem during the first year of the new rankings that THE and TR heading for trouble.

When the THE rankings came out in 2010, those absent included the Chinese University of Hong Kong, the University of Queensland, Tel Aviv University, the Hebrew University of Jerusalem, the Francophone Catholic

University of Louvain, Fudan University, Rochester, Calgary, all of the Indian Institutes of Technology, Sciences Po Paris, and the University of Texas at Austin. The omission of the two leading Israeli universities was apparently the result of some sort of misunderstanding with the request for data not reaching the right person (Schtull-Traurring, 2010).

A Faculty Council meeting at The University of Texas (UT) at Austin provided insight into why some universities were reluctant to take part in the new rankings and also into how that initial reluctance was eventually overcome. Professor Thomas Palaima said that it was embarrassing that some American public universities had taken part while UT did not. Some had done well in the new rankings even though they had been slipping in the THES-QS rankings. Eventually UT decided to participate (University of Texas, 2011).

It should be noted that the THE rankings, like their predecessors, did not have a fixed membership and that the number of ranked universities could change as institutions opt in or drop out. One consequence of this is that the processed scores could change because mean indicator scores were changing.

## The Reputation Survey

The academic survey was much more systematic and rigorous than QS's. The actual administration was done by the professional pollsters IPSOS Mori. Survey forms were sent to a selected target group, namely those who had published papers included in the Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI) and Arts and Humanities Citation Index (A & HCI). To top up the numbers in the social sciences and the arts and humanities some forms were sent to a sample of academics drawn from the IBIS Worldwide Academic and Library File produced by Mardev.

The number of forms sent to each country was decided according to the number of researchers there as found by UNESCO and provided in the Global Perspective on Research & Development report (UNESCO, 2010). The new survey also asked two specific questions about research and about postgraduate supervision (Global Institutional Profiles Project, 2010).

It was not possible at first to examine the results of the THE survey directly since they were bundled into two clusters with other research and teaching indicators. Later however, THE did issue separate reputation rankings and these seemed to be quite plausible. They did not contain the unexpectedly high scores for some Asian and Latin American universities that have been characteristic of the QS world rankings. It was also evident that after the first half dozen universities, the number of votes, and hence the scores, dropped precipitously so that a university's overall score might be influenced by small changes in the number of people supporting it in the survey.

Alex Usher of Higher Education Strategy Associates has reported research that indicates that the number of votes for universities outside the top 150 is usually likely to be no more than 40 and that small changes in survey responses can have a disproportionate effect on overall scores (Usher, 2014).

THE had originally proposed to drastically reduce the weighting given to the reputation surveys. It appears, however, that TR found that data from institutions was difficult to obtain and was not always reliable. The two reputation surveys, therefore, ended up having a more substantial weighting than had originally been intended. The teaching survey received 15% and the research survey 19.5%.

## Structure of the Rankings

The new rankings dropped the THES-QS employer surveyor and did not attempt to find any other indicator of student quality. They also included a measure of innovation and engagement with industry, research income from industry, with a weighting of 2.5%. Another new indicator was public income as a percentage of total research income with a weighting of 0.75%.

The new rankings were also distinctive in that three of the five groups of indicators contained more than one indicator. International Mix included two indicators: International Faculty and International Students. The category of Teaching – the Learning Environment consisted of five indicators. One of these, staff student ratio had been included in

the THES-QS ranking but was improved by removing full time research staff from the staff side of the equation. Added to this were a reputation survey of postgraduate teaching, income per academic staff, PhD degrees per academic staff, PhD degrees per bachelor degrees, and undergraduate admissions.

The research indicator group included papers per academic and research staff, research income, public research income as a percentage of total research income and a research reputation survey.

The grouping of indicators in clusters meant that it would be very difficult, if not impossible, to figure out the causes of a university rising or falling in the rankings. If a university improved in the Teaching or Research cluster, for example, it could be because the university had done better in one or more of three or five distinct indicators.

## The Citations Indicator

While the citations indicator stood alone without being combined with any others, there was a reduction in transparency here as well. Checking the citations per faculty indicator in the THES –QS and then the QS World University Rankings was fairly simple. The number of citations for each university over the relevant period could be checked with the Scopus database. The number of faculty could usually be found on university web pages, national databases or QS profiles.

This was not possible with the new indicator. TR used normalization by field and year. That meant each paper was compared with the world average for one of 250 possible fields in five years of publication and one, two, three, four or five possible years of citations. Calculating the world average for such a large number of citations per year per field would be impossible for anyone without TR's resources.

Previous world rankings had given a substantial weighting to citations. ARWU had an indicator with a 20% weighting that counted the number of highly cited researchers in the TR lists. QS had counted the number of citations per faculty and given it a 20% weighting. The new THE-TR citations indicator gave much greater prominence to citations (32.5%) than any previous rankings.

As we shall see, the normalization process used in the new THE rankings produced some very strange results that called into question the competence and objectivity of the rankers.

## REACTION TO THE 2010 RANKINGS

The first edition of the new rankings was generally not well received. Its merits were submerged by the extraordinary placing of Alexandria University in Egypt in the top two hundred of the world's universities and in fourth place for research impact.

This said much about the narrow worldview of THE and TR. They seemed to have no idea that scholars and scientists in the Arab and Muslim worlds had a good idea of the general merits of universities in Egypt and its neighbors. THE produced a most unconvincing explanation that noted that a Nobel laureate had studied there decades ago before departing to the US and that there was once a famous library in the city two millennia before the foundation of the university. They had to admit, however, that the real cause was the writings and the citations of precisely one man (Holmes, 2010). This is where the story of the THE rankings intersected with a major scandal of academic publishing.

The man was Mohammed el Naschie, an Egyptian by birth and ancestry who had obtained a doctorate in engineering from University College London. He later started a journal, Chaos, Solitons and Fractals, devoted to the disciplinary borderlands where applied mathematics met theoretical physics. This eventually came under the imprint of Elsevier. El Naschie published several papers every year in the journal and in the International Journal of Nonlinear Sciences and Numerical Simulation which was then published by Freund of Tel Aviv and of which he was a regional editor, sometimes more than one in a single issue. In most of those papers, he cited other papers that he had written, many of them in the same year.

At this point, he was involved in a libel case against Nature that has since been resolved in the journal's favor. During the trial and judgment, it became clear that El Naschie's writings were largely devoid of academic merit (Cressey, 2012).

For the 2010 rankings, TR used the methodology they had developed for their InCites system which had been delivered to many universities around the world. We will look at the system later but essentially what it did was to greatly amplify the number of citations in fields that had few citations, especially if they tended to have few citations in the first one or two years after publication. This was what happened with El Naschie and his publications and citations.

There were many complaints from prominent academics and ranking experts. Philip Altbach of the Center for International Higher Education at Boston College remarked that "[s]ome of the rankings are clearly inaccurate. Why do Bilkent University in Turkey and the Hong Kong Baptist University rank ahead of Michigan State University, the University of Stockholm, or Leiden University in Holland? Why is Alexandria University ranked at all in the top 200? These anomalies, and others, simply do not pass the smell test" (Altbach, 2010, para. 19)

Paul Wouters of the Centre for Science and Technology Studies (CWTS) at Leiden University, commenting on an article by Guttenplan (2010) in the New York Times, said, "the way the THES uses citation analysis does not meet one of the requirements of sound indicators: robustness against simple forms of manipulation" (Wouters, 2010, para. 2).

The new rankings did have some tepid support from leading politicians and administrators. Among them was the UK Minister for Higher Education, David Willetts, who congratulated THE "for reviewing the methodology to produce this new picture of the best in higher education worldwide" (Times Higher Education, 2010, para. 10).

In the same issue of THE, Dirk van Damme, head of the Centre for Educational Research and Innovation at the Organization for Economic Cooperation and Development, hinted that the rankings might be positive on the whole. "Rankings are not perfect. They need to be improved continuously and they sometimes lend themselves to dreadful misuses. But they enhance accountability and transparency, they stimulate comparability and competition, and in so doing they strengthen the global system of scientific research and higher education. This is only the beginning and much more work needs to be done" (Van Damme, 2010, para. 17).

Phil Baty (2014) has claimed that the response to the new rankings "was rich and encouraging" (p. 128). He reported that David Naylor of the University of Toronto had noted that THE had undertaken widespread consultations and that Ian Diamond of the University of Aberdeen referred to "sensible conversations that had led to positive changes" (p. 129). But on balance the criticism of the 2010 rankings was much louder than the praise.

## CHANGES 2010 - 2011

There was little debate and no admission of serious error. THE and TR continued to insist most of the time that there were just a few statistical anomalies and that the basic system was sound. During 2010 and 2011, there was evidently a lot of discussion about changes to the ranking methodology, this time largely in private, which resulted in several significant changes (Baty, 2011a).

One change was that normalization, used in 2010 for the citations indicator, was now applied to the number of papers per academic and research staff, the number of doctoral degrees awarded and research income. This probably helped universities with strengths in the social sciences and reduced the scores of those with a high volume of medical research. Among other things, it contributed to a noticeable improvement for the London School of Economics. The exact effect of these changes is not clear since three indicators were still combined under the Research: Volume, Income and Reputation category.

There was also some redistribution of the weighting of the various components. Within the research cluster, the weighting for research reputation was reduced from 19.5% to 18% while that for research income per academic was increased from 5.25% to 6% and the volume of research per academic and research staff from 4.5% to 6%. The public research income as a percent of total research income indicator, which had accounted for 0.75%, was now deleted.

A new indicator measuring international research collaboration was introduced with a weighting of 2.5 %. The other two internationalization

indicators were realigned. Instead of 3 % for international faculty and 2 % for international students (Baty, 2010b), it was now 2.5 % for each.

TR also introduced a number of changes to the citations indicator. First, the period for which citations were counted was extended from five to six years. In addition, the threshold for ranking was raised from fifty to two hundred papers per year. The weighting for the indicator was reduced from 32.5% to 30% and finally, a "regional modification" was introduced, by which the normalized citation counts for universities were divided by the square root of the counts for the country in which they were located. In effect, universities would receive a substantial boost just for being in a low impact country. This may well have contributed to a few universities in countries like Turkey, Chile, Morocco and Italy getting remarkably high scores for the citations indicator.

## REVIEW OF THE THE RANKINGS AFTER 2010

One distinctive feature of the rankings after 2010 was that uniquely among international rankings, they continued to bundle indicators together in groups with only the group score being given.

The indicator group Teaching – the Learning Environment contained five separate indicators: reputation survey of postgraduate teaching (weighting of 15%), PhD awards per academic staff (6%), undergraduates admitted per academic staff (4.5%), income per academic staff (2.25%), and PhD awards per bachelor awards (2.25%), making a total of 30% (Baty, 2011a). As noted above, a change in a university's score for this group of indicators could result from a change in the score for one or more of the component indicators. Similarly, the Research: Volume, Income and Reputation indicator combined scores for research reputation (18%), research income (6%) and papers per academic and research staff (6%) and, a change in the score for this indicator group could result from a change in one or more of three indicators.

The International Outlook: Staff, Students and Research indicator now had three components, each with a weighting of 2.5%, international students, international faculty and international collaboration. A rise in a

university's score could result from an increase in the number of international students or international faculty, a decline in the total number of students or faculty, an increase in the number of international collaborations, a reduction in the total number of publications, a decline in the mean score among ranked universities or some combination of any of these.

There were two indicators that stood alone, innovation: industry income and citations: Research impact. Each of these was problematical. It turned out that data for the income from industry indicator was not always available or reliable and that several universities were not scored and their total score was decided by the other indicators.

There were also serious problems with the citations indicator which continued to present problems for THE and TR after 2010. A major objective of the changes to this indicator was removing the embarrassment of Alexandria University. In this, THE and TR were apparently successful. The citations score for Alexandria fell from 99.8 to 61.4, which was still far ahead of its realistic and very modest research score of 7.8 but enough to relegate it from the overall top 200 to the 301-350 band (Times Higher Education, 2011).

The citations indicator however, continued to produce implausible if not downright ludicrous results. In 2012, the joint top universities for research impact were Rice University in Texas and Moscow Engineering Physics Institute (MEPHi). Rice as top university for research impact, rather than Caltech, Berkeley, MIT or Harvard, was a little difficult to believe but MEPHI, a single subject research institute although an excellent one by all accounts, was absurd. It turned out that MEPHi, because it taught only a single narrow subject, should not have been there in the first place and it was duly removed in 2013. Its presence in 2012 seemed to have been due to two reviews of particle physics, each with over a thousand citations and over a hundred contributors. MEPHi was credited with all those citations, just like all the other contributing institutions, as though it was the sole affiliation of the authors of the reviews. For most of the other universities involved with the reviews, this did not matter very much because they were producing thousands of papers and tens of thousands of citations. MEPHI in contrast was not producing very many papers so those thousands of citations made a big difference.

Over the next few years, there were more incongruous entries in the THE world rankings with institutions getting scores for citations that were far ahead of their overall scores or those for the research indicators. These included Tokyo Metropolitan University, Royal Holloway University of London, Florida Institute of Technology, St George's University of London, King Mongkut's University of Technology, Thonburi, Scuala Normale Superiore di Pisa, and three universities in Istanbul: Boğaziçi University, Istanbul University and Sabançi University (Holmes, 2012).

Evidently the indicator still had several problems. One was that it failed to use fractional counting of citations, that is dividing all the citations by the number of contributing institutions, a procedure used without difficulty by the CWTS Leiden Ranking. As a result, an institution that had made a contribution, no matter how slight, to a publication with scores, or even hundreds, of "authors" and had received hundreds or thousands of citations would receive a huge and incongruous boost providing that its total number of publications remained low. Such multi-contributor publications were most common in particle physics and included the biennial Review of Particle Physics and papers derived from the Large Hadron Collider project. They could also be found in medicine and genetics.

Next, TR continued to count self-citations, something that QS, to its credit, had stopped doing in 2011. The most obvious example of self-citation was the aggressive self-promotion of El Naschie in the pages of Chaos, Solitons and Fractals and the Journal of Nonlinear Sciences and Numerical Simulation. That became less of a threat after the tweaking in 2011 but there could be no guarantee that self-citation might not give an undue advantage to less than scrupulous research teams in subsequent years.

Another problem, which few experts discussed until 2014, is the "regional modification", the division of the final score of a university for the citation indication by the square root of the score for the country where the university is located. Two universities might start out with identical scores for citations but if one was located in a high scoring country and one in a low scoring country they would end up with very different scores.

The justification for the regional modification was hard to follow. Simon Pratt of TR is reported to have said that "there are significant contrasts in citations behaviour and patterns in different geographic regions. It is argued that not all of these are indicative of underlying research impact. For example, universities in the US are part of a very large research community…which may lead to higher innate citation rates than their peers in developing countries…" (as cited in Baty, 2011b, para.13) and that a "modification to normalise citations by region can help spotlight exceptional institutions in typically low-citing countries. Such a change can also result in a more diverse rankings table that highlights excellence in developing countries" (Baty 2011b, para. 13). It was also claimed that some countries did not provide adequate research funding or encourage networking and that their universities needed a boost to compensate. This is in fact only partly true. Many universities in the Gulf are extremely generous with funding and sponsorship for overseas conferences and so on while there is quite a lot of money around in some Southeast Asian countries although it does not always reach the most qualified researchers.

The result of all of the above was that many universities in South America, Eastern and Mediterranean Europe, and South and East Asia received scores for citations that were out of proportion to other scores especially the research cluster. Why THE should take such a risk with their credibility is not clear but it is possible that it was a financially rewarding proposition for TR to save costs by using procedures identical to their lucrative InCites staff evaluation system.

## PUBLIC PERCEPTIONS AFTER 2010

Despite these difficulties, it seems that by 2014 THE had recovered from the disappointing reception of 2010. As early as 2011, the publication received two awards from the Professional Publishers Association for Business Media Brand of the Year and Weekly Business Magazine of the Year (Professional Publishers Association, 2011). Then, in 2012, a report from the steering group of the British Academy (Foley & Goldstein, 2012) reviewed the THE world rankings but not those published by QS or the Shanghai Center for World-Class Universities.

By 2011, after the second edition of the new rankings, Ferdinand von Prondzynski, then head of Robert Gordon University in Scotland, could say that the THE rankings "over recent years have been accepted as the most authoritative international league table" (von Prondzynski, 2011, para.1) although that could imply that they had also been such during the THE-QS period. By 2012, he was reported as saying that they were "increasingly seen as the gold standard" (as cited in Baty, 2012, para. 14).

Meanwhile, Dirk Van Damme became more positive than he had been in 2010 and enthused that the "THE World University Rankings invite us to look beyond the top of the list. They provide a unique opportunity to witness the dynamics in the sub-top league of aspiring institutions, the global expansion and dispersion of the system, and the varying capacities of national systems to succeed in the global war for academic excellence" (Van Damme, 2012, para. 19).

Further endorsements followed. In 2013, Shashi Tharoor, Indian Minister of State for Human Resources Development, asserted that the THE rankings were the "principal yardstick we should look to" (as cited in Baty, 2013, para 13) and a year later when the Economist was reviewing the quality of Chinese universities it saw fit to mention only performance in the THE rankings and ARWU ("A matter of honor", 2014).

In 2014, an official Norwegian report was very critical of global university rankings for several reasons. It analyzed three rankings, the Shanghai ARWU, the Leiden ranking and the THE World University Rankings but not the QS rankings. (Piro, Hovdhaugen, Elken, Sivertsen, Benner & Stensaker, 2014).

In Saudi Arabia, the Custodian of the Two Holy Mosques offers scholarships at 200 universities from the big four brand names, QS, THE, ARWU and the new Best Global Universities produced by the US News ("Scholarship students to study at the world's best universities", 2014). Similarly, the same four rankings had been mentioned earlier for a Hong Kong government scholarship program (Education Bureau, 2014).

It seems that when the media or government bodies wish to be inclusive, they refer to the big three or, now, the big four, including the

US News Best Global Universities, rankings. Should they need to be parsimonious and refer only to two then those two will usually be the THE and the Shanghai rankings. When only one is referred to, that is very often THE. Thus, in 2014 the Asahi Shimbun referred to the position of Japanese universities in the THE rankings but no other (Takahami, 2014) as did the Wall Street Journal in 2015 (Obe, 2015).

The most notable endorsement was from the President of Peking University, Zhu Shanlu, who described Phil Baty as the "education secretary for the world", a phrase repeated in a presentation sent to a seminar in Moscow in April 2015 and at other events (Siwinski, 2015, para. 1).

The supposed virtues of the new THE rankings were even used to denigrate the U-Multirank rating system, which met with considerable hostility from the British higher education establishment and a few leading research universities on the European continent. The European Union committee of the House of Lords reported that THE "told us that their approach seeks to achieve more objectivity by capturing the full range of a global university's activities – research, teaching, knowledge transfer and internationalization – and allows users to rank institutions (including 178 in Europe) against five separate criteria: teaching (the learning environment rather than quality); international outlook (staff, students and research); industry income (innovation); research (volume income and reputation); and citations (research influence)" (House of Lords European Union Committee 2013, para. 55).

So, by 2014 the THE rankings had apparently won the favour of the higher education administrative elite of the United Kingdom and a swathe of research orientated universities in continental Europe. Also, it appeared that many universities in Latin America, Africa, and the Middle East were attracted by the presentation of the THE rankings as prestigious and by participation at exclusive summits. It is not impossible that many universities with middling reputations were aware that one or two contributors to a multi-contributor publication with a huge number of citations could lead to a high overall score.

There were signs that some informed observers were becoming disillusioned. Perhaps the most significant of these was Simon Marginson

of the University College London Institute of Education who said in 2013 that the THE rankings were fatally flawed outside the top fifty although they were still better than the QS rankings (Hare, 2013). Even so, the world rankings and their regional and specialized spin-offs have continued to have a generally respectful reception from administrators and the media.

## THE END OF THE THE-TR PARTNERSHIP 2010

At the end of 2014, THE announced that it was ending its partnership with TR and would use data derived from Scopus. It became clear that THE intended to exert greater control over the data and was essentially trying to move upstream and monetize the data collection and distribution process (Baty, 2015).

THE claimed that this would inaugurate a new era of openness and accountability. One sign of this was that when the results of its new academic reputation survey became available, THE was now able to provide a country by country breakdown of the respondents to the survey. This was a small step but one that might be the harbinger of more significant changes.

In early 2015, THE indicated that it was aware of the distorting effects of multi – author publications and indicated that it was considering two possible ways of getting around this, simply deleting these with more than a certain number of authors or using fractional counting. In July, THE published an experimental ranking indicator for 30 African universities that was nothing more than field normalized citation per paper. This time, however, there was a difference. THE, now managing the data collection and processing by itself, had used fractional counting. There were a few cases of universities that had been in both the African rankings and the previous year's world rankings. It was interesting that Université Cadi Ayyad Marrakech's citations score was noticeably reduced (Bothwell, 2015).

However, when the world rankings were published THE simply resorted to removing 649 papers with more than a thousand authors, although it seemed that they were considering introducing fractional

counting of citations in later editions of the world rankings. In addition, THE announced that they would divide the citations scores in two, half of them with the regional modification and half without.

The 2015 rankings saw some unprecedented changes. Some universities, such as Twente, Moscow State and University College Dublin did much better than in 2014. At the top end, Oxford and Cambridge rose to overtake Harvard, a shift that appeared to have little to do with Oxford or Cambridge but reflected a large and unexplained fall in Harvard's score for Teaching.

In contrast, there were some dramatic falls in the scores and places of French, Japanese and Korean institutions. Several Turkish universities suffered calamitous tumbles of hundreds of places, because of the withdrawal of the advantages accruing from participation in the Large Hadron Collider project and from the full regional modification (Times Higher Education, 2015).

## CONCLUSION

At the end of 2015, it was still not clear whether THE would proceed with further changes. Nor is it certain whether it can maintain its popularity with the world's leading research universities. There is a clear dilemma here. THE could make further changes, especially to the citations indicator, but that would call into question their reputation for reliability and consistency. But if they do not make changes, it is likely that will be more and more anomalous results. This problem is not confined to THE: QS and ARWU have also been forced to confront such problems over the last two years.

Whatever happens, it is likely that the story of the THE rankings will continue to be an interesting one and that the attempt to create global university rankings that are accurate, reliable and valid has yet to succeed.

1. Note

In this paper, 'THES-QS World University Rankings' and 'THE-QS World University Rankings' refer to the rankings published between 2004

and 2009 by the newspaper *Times Higher Education Supplement* (THES), which became the magazine *Times Higher Education* (THE) in 2008, with data collected by QS Quacquarelli Symonds (QS).

'QS World University Rankings' refers to the rankings published in 2010 and after by QS using the methodology that had been developed for the THES-QS/THE-QS rankings and with data supplied by Scopus.

'THE World University Rankings' refers to the rankings published by Times Higher Education (THE) from 2010 with data provided by Thomson Reuters (TR) from 2010 to 2014 when THE announced that it would process and analyze data collected from Scopus.

## REFERENCES

Academic Ranking of World Universities. (2015). *Methodology*. Retrieved from http://www.shanghairanking.com/ARWU-Methodology-2015.html

Adams, J., & Baker, K. (2010). *Global opinion survey: New outlooks on institutional profiles*. Thomson Reuters. Retrieved from http://ip-science.thomsonreuters.com/m/pdfs/Global_Opinion_Survey.pdf

Altbach, P. (2010, November 10). The state of the rankings. *Inside HigherEd*. Retrieved from http://www.insidehighered.com/views/2010/11/11/altbach

Asia's best universities 1999. (1999). *Asiaweek* Retrieved from http://www-cgi.cnn.com/ASIANOW/asiaweek/universities/schools/index.html

Asia's best universities 2000. (2001). *Asiaweek*. (2001). Retrieved from http://cgi.cnn.com/ASIANOW/asiaweek/features/universities2000/index.html

Ballantyne, P. (2002). Psychology, society, and ability testing (1859-2002): Transformative alternatives to mental Darwinism and interactionism. Retrieved from http://www.cyberus.ca/~pballan/

Baty, P. (2009). New data partner for World University Rankings. *Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk/408881.article

Baty, P. (2010a, August 26). THE World University Rankings. *Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk/413160.article

Baty, P. (2010b, September 16). World University Rankings 2010-2011 methodology. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/world-university-rankings-2010-11-methodology

Baty, P. (2010c, September 2). THE rankings release set for 16 September. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/world-university-rankings/2010-11/world-ranking/analysis/robust-transparent-and-sophisticated

Baty, P. (2010d, September 16). Robust, transparent and sophisticated. *Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk/413249.article

Baty, P. (2011a, October 6). World University Rankings 2011-2012 methodology. *Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk/world-university-rankings/2011-12/world-ranking/methodology

Baty, P. (2011b, February 1). What's normal for a given region. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/news/whats-normal-for-a-given-region/415047.article

Baty, P. (2012, December 22). The fragile and changing nature of university reputation. *University World News*. 19 February 2012. Retrieved from http://www.universityworldnews.com/article.php?story=20120214154324613

Baty, P. (2013, June 5). India must embrace the internationalization of higher education. *Times Higher Education*. Retrieved from http://www.timeshighereducation.co.uk/world-university-rankings/news/india-embrace

Baty, P. (2014). The Times Higher Education World University Rankings, 2004-2012. *Ethics in Science and Environmental Policy 13*, 125-130

Baty, P. (2015, January 30). THE World University Ranking masterclass in Moscow. Retrieved from https://www.youtube.com/watch?v=_KoAao4xfmY. Accessed 15 January 2016

Bothwell, E. (2015, July 9). South African institutions top THE rankings pilot. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/news/south-african-institutions-top-rankings-pilot

Cressey, D. (2012, July 6). Nature publishing group wins long-running libel trial. *Nature*. Retrieved from http://www.nature.com/news/nature-publishing-group-wins-long-running-libel-trial-1.10965

Education Bureau. (2014). Hong Kong scholarship for excellence scheme. Education Bureau: The government of the Hong Kong special administrative region. Retrieved from http://www.edb.gov.hk/en/edu-system/postsecondary/policy-doc/hkses/index.html.

Foley, B. & Goldstein, H. (2012). *Measuring success: League tables in the public sector*. The British Academy: London. Retrieved from http://www.britac.ac.uk/policy/Measuring-success.cfm. Accessed 22 December 2015

Global Institutional Profiles Project. (2010). Academic reputation survey: Stage 2 methodology. Retrieved from http://ip-science.thomsonreuters.com/m/pdfs/GIPP_AcamRep_methodology.pdf. Accessed 22 December 2014

Guttenplan, D. D. (2010, November 14). Questionable science behind academic rankings. *New York Times*. n.p.

Hare, J. (2013, October 16). Ranking of unis is 'bad science', says Simon Marginson. *The Australian*. Retrieved from http://www.theaustralian.com.au/higher-education/ranking-of-unis-is-bad-science-says-simon-marginson/story-e6frgcjx-1226740536595

Holmes, R. (2006a, September 8). More on the rise of Ecole Polytechnique. *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

Holmes, R. (2006b, October 1). Another Duke scandal. *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

Holmes, R. (2007, January 27). More about Duke. *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

Holmes, R. (2009, December 1). Whither the Times higher ranking. *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

Holmes, R. (2010). Debate, anyone? *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

Holmes, R. (2011). The THE-QS World University Rankings, 2004-2009. *Asian Journal of University Education*, 6 (1), 91-113.

Holmes, R. (2012, November 15). Ranking's research impact indicator is skewed. *University World News*. Retrieved from http://www.universityworldnews.com/article.php?story=20121109142209937

Holmes, R. (2014, September 18). QS World University Rankings. *University Ranking Watch*. Retrieved from http://rankingwatch.blogspot.my/

House of Lords European Union Committee. (2013). Modernisation of higher education in Europe: The EU contribution. Oral and written evidence. Retrieved from http://www.parliament.uk/documents/lords-committees/eu-sub-com-g/WrittenandOralHEevidence.pdf

Ince, M. (2010). International workshop on performance management: Kent Business School Presentation. Retrieved from http://www.slideshare.net/KentBusinessSchool/04-martin-ince-kent-presentation-july-2010

Liu, N. C. (2009). The story of academic rankings. *Shanghai Express*. Retrieved from http://shangaiexpress.blogspot.com/2009/10/67-story-of-academic-rankings-nian-cai.html.

A matter of honor. (2014). *Economist.* Retrieved from http://www.economist.com/news/china/21633865-china-trying-reverse-its-brain-drain-matter-honours

Mykelbust, J. P. (2011, May 29). France: Rankings spurred university reforms. *University World News*. Retrieved from http://www.universityworldnews.com/article.php?story=20110527211248644

Obe, M. (2015, August 3). Japan rethinks higher education in skills push. *The Wall Street Journal*. Retrieved from http://www.wsj.com/articles/japan-rethinks-higher-education-in-skills-push-1438571119

Piro, F. N, Hovdhaugen, E, Elken M., Sivertsen G., Benner M., & Stensaker B. (2014). *Nordiske universiteter og internasjonale universitetsrangeringer*. Nordic Institute for Studies in Innovation, Research and Education. Retrieved from http://www.nifu.no/files/2014/09/NIFU-Rapport-25_2014.pdf

Professional Publishers Association. (2011). 2011 PPA Awards winners revealed. Retrieved from http://www.ppa.co.uk/events/news/2011-ppa-awards-winners-revealed-at-star-tinged-ceremony/

Scholarship students to study at the world's best universities. (2014, December 3). *Arab News*. Retrieved from http://www.arabnews.com/saudi-arabia/news/668986

Schtull-Traurring, A. (2010, September 19). Israeli universities dropped from list of top 200 in world. *Haaretz*. Retrieved from http://www.

haaretz.com/print-edition/news/israeli-universities-dropped-from-list-of-top-200-in-world-1.314520

Schwartz, S. (2001, November 30). Asiaweek closes with 80 jobs lost. *South China Morning Post*. Retrieved from http://www.scmp.com/news/hong-kong

Sowter, B. (2008). THES-QS World University Rankings: New Developments in Ranking Methodology 4 April 2008. Retrieved from http://www.sc.mahidol.ac.th/e_learning/lecture.htm#

Siwinski, W. (2015, October 17). The renaissance of national rankings. *University World News*. Retrieved from http://www.universityworldnews.com/article.php?story=20151015173737446

Takahami, Y. (2014, September 27). Cash injection to help select universities become elite global players. *The Asahi Shimbu*. Retrieved from http://ajw.asahi.com/article/behind_news/social_affairs/AJ201409270024

Thomson Reuters. (2012). Global Institutional Profiles Project: Frequently asked questions. Retrieved from http://ip-science.thomsonreuters.com/m/pdfs/DataCollection_FAQ.pdf

*Times Higher Education.* (2010, September 16). UK universities global path in world rankings. Retrieved from http://www.timeshighereducation.co.uk/world-university-rankings/2010-11/world-ranking/analysis/uk-education

*Times Higher Education.* (2011). Alexandria University. Retrieved from https://www.timeshighereducation.com/world-university-rankings/alexandria-university?ranking-dataset=1053

*Times Higher Education.* (2015). World University Rankings 2015-2016. Retrieved from https://www.timeshighereducation.com/world-university-rankings/2016/world-ranking#!/page/0/length/25

Topuniversities. (2009). QS World Universities Rankings 2009. Retrieved from http://www.topuniversities.com/university-rankings/world-university-rankings/2009?destination=node/10798

Topuniversities. (2013). QS World Universities Rankings 2013. Retrieved from http://www.topuniversities.com/university-rankings/world-university-rankings/2013

Topuniversities. (2015). QS World Universities Rankings 2015. Retrieved from http://www.topuniversities.com/university-rankings/world-university-rankings/2015

UNESCO. (2010). A Global Perspective on Science, Technology and Innovation. Retrieved from http://www.uis.unesco.org/ScienceTechnology/Pages/default.aspx#top

University of Texas. (2011). Minutes of the regular faculty council meeting, March 21 2011. Retrieved from http://www.utexas.edu/faculty/council/2010-2011/minutes/min032111/III.A.html

Usher, A. (2014). The problem with global reputation rankings. Higher Education Strategy Associates. Retrieved from http://higheredstrategy.com/the-problem-with-global-reputation-rankings/

Van Damme, D. (2010, September 16). Mapping the plateaux as well as the peaks. *Times Higher Education*. Retrieved from https://www.timeshighereducation.co.uk/world-university-rankings/2010-11/world-ranking/analysis/rankings-globalisation

Van Damme, D. (2012, October 4). Elite at the crossroads as aspirants chart fresh course. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/world-university-rankings/2012-13/world-ranking/analysis/world-education-rankings

Von Prondynsky, F. (2011). *Times Higher Rankings*. A University Blog. Retrieved from https://universitydiary.wordpress.com/2011/10/06/times-higher-rankings/

Wouters, P. (2010, November 18). Questionable science behind academic rankings. *New York Times*. Retrieved from http://www.nytimes.com/2010/11/15/education/15iht-educLede15.html