Journal of Library and Information Studies 9:1 (June 2011) p.1-25

A Comparison of Three Major Academic Rankings for World Universities: From a Research Evaluation Perspective

Mu-Hsuan Huang

Abstract

This paper introduces three current major university ranking systems. The *Performance Ranking of Scientific Papers for World Universities* by Higher Education Evaluation and Accreditation Council of Taiwan (HEEACT Ranking) emphasizes both the quality and quantity of research and current research performance. The *Academic Ranking of World Universities* by Shanghai Jiao Tung University (ARWU) focuses on outstanding performance of universities with indicators such as Nobel Prize winners. The *QS World University Ranking* (2004-2009) by Times Higher Education (THE-QS) emphasizes on peer review with high weighting in evaluation. This paper compares the 2009 ranking results from the three ranking systems. Differences exist in the top 20 universities in three ranking systems except the Harvard University, which scored top one in all of the three rankings. Comparisons also revealed that the THE-QS favored UK universities. Further, obvious differences can be observed between THE-QS and the other two rankings when ranking results of some European countries (Germany, UK, Netherlands, & Switzerland) and Chinese speaking regions were compared.

Keywords: World Universities; Performance Ranking; Scientific Papers; Research Evaluation

1. Introduction

Research evaluation identifies the strength and weakness of a university and provides information for improving academic research. Evaluation results set the benchmark for national investment on higher education (Huang, 2005). Resources and budgets are tightening in higher education. Effective resources allocation is essential for producing remarkable research. Research evaluation promotes informed policy decisions in higher education and offers invaluable information to university administrators, e.g., subject areas meriting investment, faculty recruiting, grant support, etc. Many universities and research institutions are actively conducting internal or external evaluation in order to stay competitive.

This paper describes three well-known global-scaled university evaluation programs whose ranking results are internationally visible and often serve as important external evaluation for world universities. The programs

Professor, Department of Library and Information Science, National Taiwan University, Taiwan.
E-mail: mhhuang@ntu.edu.tw

are the *Performance Ranking of Scientific Papers for World University* directed by Higher Education Evaluation and Accreditation Council of Taiwan (the HEEACT ranking, 2007-present), the *Academic Ranking of World Universities* by Shanghai Jiao Tung University (the ARWU ranking, 2003-present) and the *QS World University Rankings* by Times Higher Education, (the THE-QS ranking, 2004-2009), which split into two independent ranking programs in 2010 (QS World University Rankings and THE World University Rankings).

Given the similarity of the programs in scale and purpose, the ranking systems of the three programs vary in their methodologies which demonstrate significantly different focus, emphases, and evaluation strategies. HEEACT focuses on the scientific research performance of universities and takes into account both recent research performance and the research output accumulated over time. ARWU emphasizes on highly extraordinary research achievement and may fail to differentiate the performance of most universities that constitute the majority of the population. THE-QS relies heavily on peer review and tends to favor the famous and historically established universities.

The three ranking systems also differ in subject scope and subject categorization. All of them provide subject field based ranking results, but the categorization of subject fields is different in each system. The HEEACT and ARWU programs cover only the fields of sciences and engineering and social sciences. THE-QS is the only program that includes arts and humanities.

2. An Overview of University Evaluation

2.1. Types & Levels of University Evaluation

Before discussing on the three university ranking systems, it is necessary to distinguish different types of evaluation. University evaluation is an umbrella term encompassing academic evaluation and research evaluation (see Figure 1). Targets of assessment may include research achievements, university administration, education quality, etc. (Hong, 2009). Some OECD countries have begun the evaluation of their higher education institutions in order to fully understand their performance and service quality (Staropoli, 1987). Depending on the purposes of evaluation, an evaluation program may use dramatically different criteria and indicators.

Existing literature also often fail to clearly define and differentiate the levels of evaluation. Academic evaluation encompasses the assessment of scholarly activities, achievement, outcome of research investment, etc. (Daniel & Fisch, 1990). Research evaluation is even more specific than academic evaluation. Conceptually, university evaluation is the

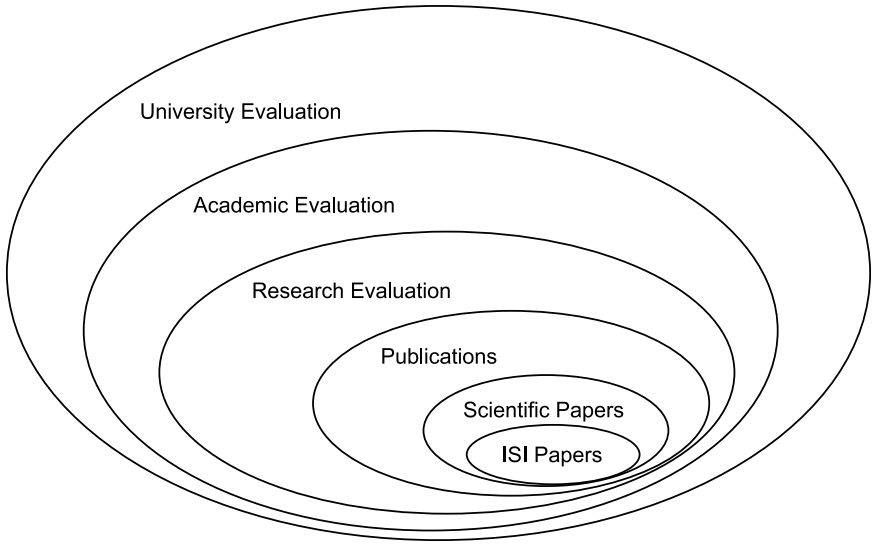

Figure 1 University Evaluation, Academic Evaluation and Various Levels of Research Evaluation

broadest, including not only research but also teaching, services, and overall administration. University evaluation encompasses both academic performance (often discipline-based) and administrative performance (Huang, 2003). The general public may confuse the three levels of evaluation, but they should be clearly differentiated because they essentially assess different levels and aspects of university performance.

Scholarly publication is one of the major indicators used in research evaluation. Specifically, scientific papers indexed in the ISI citation index databases are a well accepted indication of research performance because the databases selectively index academic journals or serial publications that are of higher quality.

2.2. The Subjective and Objective Approaches in University Evaluation

Two major approaches exist in research evaluation – peer review evaluation and bibliometric evaluation. Peer review evaluation serves to identify and improve existing problems or deficiencies via expert opinions (Kruytbosch, 1989). It is widely used in evaluating research grant proposals, publication manuscripts, and tenure granting (Liu, 1998). Many people consider peer review to be a major means of quality judgment and it can overcome certain difficulties in academic evaluation (Campbell, 2002). If designed properly and executed carefully, peer review evaluation can offer invaluable information.

However, peer review evaluation has been

criticized for its subjectivity, e.g., how is "peer" defined and chosen and whether the process of review may result in fair and sound verdicts (Aguillo, Bar-Ilan, Levene, & Ortega, 2010; Bookstein, Seidler, Fieder, & Winckler, 2010; Buela-Casal, Gutiérrez-Martínez, Bermúdez-Sánchez, & Vadillo-Muñoz, 2007). Moreover, in today's research world characterized by inter- and/or multidisciplinarity, individual peer reviewers now have less command of the highly complicated knowledge to fully evaluate a piece of research. The intensity of scholarly communication further weakens each reviewer's ability because one can hardly have full command of the constantly updating research literatures. Peer review now may not be the best quality assessment method as it was supposed to be (Thomas & Watkins, 1998). On the other hand, there is also a growing demand for objective and quantified evaluation. The funding agencies and universities are actively pursuing quantifiable indices for research assessment. Bibliometric evaluation therefore becomes a popular tool in supplementing peer review evaluation.

Bibliometrics can be used to investigate the development, dissemination and status quo of a knowledge field by using statistic analyses on bibliographic data. Bibliometrics is widely used in research evaluation for its objectivity and operatability, even though some people question its conceptual assumptions, procedural validity (van Raan, 1996, 2005), and biases in language, countries, etc. (Kokko & Sutherland, 1999; Leimu & Koricheva, 2005; Liu, Cheng, & Liu, 2005; Van Leeuwen, Moed, & Reedijk, 1999; Wong & Kokko, 2005).

Two reasons support the objectivity of bibliometric evaluation. First, results from bibliometric evaluation can be scientifically verified in replication; it is free from possible reviewer prejudice and bias. Second, the publications and citations based bibliometric evaluation may be viewed as a form of peer review. For example, research papers are reviewed before they are accepted for publication in journals. Journals are reviewed and selected for inclusion in citation databases. A paper cited by the other articles is read by the citer and the action of citing arguably affirmed the cited paper's contribution. In the other word, bibliometric evaluation can be viewed as the totality of multi-layered and bottom-up indirect peer reviews. One can still argue about the existence of reviewer biases embedded in bibliometric data. But with the large amount of data, the impact of individualistic biases is less significant, and thus the result is more objective than that of direct peer review.

Some empirical studies have suggested a good level of consistency between the results from peer review evaluation and bibliometric

evaluation (Norris & Oppenheim, 2003; So, 1998). A few studies further indicated the higher correlations between peer review and bibliometric evaluation in assessing the basic sciences research as opposed to the applied sciences research (Meho & Sonnenwald, 2000; Rinia, van Leeuwen, van Vuren, & van Raan, 1998). Although some studies (Aksnes & Taxt 2004; Makino, 1998) found low correlations and inconsistency, researchers (e.g., Makino, 1998; Weingart, 2005) still agreed with the applicability of bibliometrics in research evaluation in part for its strength supplementing peer review, which can be subjectively biased.

2.3. University Evaluation vs. University Ranking

Evaluation and ranking are two different but related concepts. They differ in their purposes and outcome. Evaluation is not equal to ranking. Evaluation sets a benchmark against which a university performance in certain aspects can be assessed. The goal is to determine if a university passes the assessment, meaning it has achieved at or surpassed a basic level of requirements. Evaluation results do not have to be quantitative. Descriptive evaluation suffices in some evaluation contexts, and some evaluation results indicate simply final decisions such as pass or fail to pass.

Ranking, on the other hand, sorts a group of universities by numerical indicators. Ranking shows a university's relative strength and weakness as compared to its peer institutions in the areas represented by the indicators. It clearly indicates a university's relative location at a scale representing its strength in the measured aspect. The numerical nature of ranking also simplifies comparisons. Ranking is an efficient, convenient, and easily understandable evaluation method, even though some have argued about the fairness of quantitative comparisons of universities where each university is unique and differs to the others in some aspects. Ranking employing measures composed of multiple indicators may to certain extent overcome the possible fairness problems.

Ranking has several advantages. First, it makes it easy for viewers to compare and contrast the performances of the universities being evaluated. Second, ranking indicates each university's relative achievement in certain aspects and thus helps a university to diagnose problems and/or suggests directions of development. Third, research funding agencies and the general public require open information about universities' performances. Ranking fulfills the need for a clear and objective indication about a university's performance. Finally, a carefully designed quantitative data based ranking offers objective information for policymaking.

Ranking is not without controversy. Reliability of the ranking methodology and the validity of indicators used in a ranking system are two major issues causing debates. Reliability requires a ranking system to generate consistent results in replication. Validity concerns how well the indicators represent the evaluation criteria and whether the evaluation is properly conducted. The issues are further complicated by the major ranking systems' use of compound criteria and multiple indicators.

The scale of a ranking project also affects the feasibility of a ranking methodology. In a large scale project comparing performances of universities worldwide, difficulties lie in accessing certain university data and ensuring validity of comparisons. Not all university data is open to the public, and some universities may reject evaluator requests for information. Even when data are available, comparing university performances of different countries or regions can be problematic. University performance is affected by the larger sociocultural and politico-economic context. Whether ranking indicators are fair for all universities is open to question. For example, reputation based evaluation can be highly biased toward famous universities in the Western world or those universities in reviewers' home countries.

3. Methodologies of the Three Ranking Systems

3.1. The HEEACT Ranking

The aim of the HEEACT Ranking is to identify the top 500 universities in the world that have performed well in scientific research. According to its official Web site (Higher Education Evaluation & Accreditation Council of Taiwan, 2009a), the ranking program employs multiple weighted indicators to evaluate the university performance (see Table 1).

Slight changes have been made to the indicators used in the annual evaluations since the program launched in 2007. For instance, the 2007 indicators included a measure called the "Number of subject fields where the university demonstrates excellence." This measure was removed in 2008, and the weighting was allocated to other indicators. In addition, when the program began, it offered only the overall ranking. In 2008, it started to provide field-based ranking for six fields: Agriculture (AGE), Medicine (MED), Engineering (ENG), Life Science (LIFE), Science (SCI), and Social Science (SOC) (Higher Education Evaluation & Accreditation Council of Taiwan, 2008).

The HEEACT Ranking demonstrates the following features in design. First of all, it emphasizes the quality of research; the indicators assessing research quality (*research impact and research excellence*) accounts for

Table 1. The Criteria and Indicators of the 2009 HEEACT Ranking

Criteria	Indicator	Weight	
Research productivity	Number of articles in the last 11 years (1998-2008)	10%	20%
	Number of articles in the current year (2008)	10%	
Research impact	Number of citations in the last 11 years (1998-2008)	10%	30%
	Number of citations in the last 2 years (2007-2008)	10%	
	Average number of citations in the last 11 years (1998-2008)	10%	
Research excellence	h-index of the last 2 years (2007-2008)	20%	50%
	Number of Highly Cited Papers (1998-2008)	15%	
	Number of articles in high-impact journals in the current year (2008)	15%	

Note. From"Performance Ranking of Scientific Papers for World Universities 2009,"by Higher Education Evaluation & Accreditation Council of Taiwan, Retrieved September 2, 2010 from http://ranking.heeact.edu.tw/zh-tw/2009/Page/Methodology

80% of the performance score. The research impact and research excellence criteria address the quality of a university's research output. The calculation of each university's score is based on the number of citations to its published articles, h-index of the last two years, number of its Highly Cited Papers, number of papers published in top journals, and the number of subject fields in which the university demonstrates excellent performance.

Secondly, it avoids biases caused by university size or faculty numbers. Traditionally the size of a university affects its ranking when the number of articles is used as a sole measure for research output. Because the number of articles is closely tied to the number of faculty members, rankings employing numbers of articles often favor larger universities. HEEACT Ranking corrects the flaw by using the average number of citations, the number of subject fields where the university demonstrates excellent performance, and the *h-index*. The inclusion of the three measures which accounts for 40% of the total scores balances the assessments of quality and quantity and provides a fairer representation of a university's performance regardless of its size.

Third, it takes into account a university's short-term research performance (constituting 50% of the score), thus ensures a fairer comparison between universities of varied lengths of history. The HEEACT Ranking indicators

seek to represent both the long-term and short-term research performances of a university. The inclusion of indicators assessing short-term performances corrects the flaws resulted from undifferentiating indicators that favor universities of longer histories. These short-term performance indicators include: the number of articles in the current year, the number of citations in two years, the *h-index* of the last two years, and the number of articles in high-impact journals in the current years.

The HEEACT Ranking's emphasis on the recent research performance makes the ranking a fairer one than those using measures such as THE-QS's use of university reputation and ARWU's use of Nobel Prize winners, which tend to favor universities with longer histories and or in developed countries.

3.2. The ARWU Ranking

Since 2003 the university began to annually publish the world universities ranking results. The goal of the ranking is to objectively identify the top 500 best universities through the use of quantitative data. Table 2 lists the criteria and indicators used in the ARWU ranking for 2009 (Shanghai Ranking Consultancy, 2009a).

Table 2. The Criteria & Indicators of ARWU Ranking for 2009

Criteria	Indicator	Description	Weight
Quality of Education	Alumni	The total number of the alumni of an institution winning Nobel Prizes and Fields Medals.	10%
Quality of Faculty	Award	The total number of the staff of an institution winning Nobel prizes in physics, chemistry, medicine and economics and Fields Medal in Mathematics.	20%
	HiCi	The number of highly cited researchers in broad subject categories in life sciences, medicine, physical sciences, engineering and social sciences.	20%
Research	N&S	The number of articles published in Nature and Science from 2004 to 2008.	20%
Output	SCI	Total number of articles indexed in Science Citation Index-expanded and Social Science Citation Index in 2008.	20%
Size of Institution	Size	The weighted scores of the above five indicators divided by the number of full-time equivalent academic staff.	10%

Note. From "Ranking Methodology," by Shanghai Ranking Consultancy, Retrieved September 4, 2010 from http://www.arwu.org/ARWUMethodology2009.jsp

Obviously, the ARWU criteria and indicators emphasize university performance in research. For example, its assessment of quality of education and faculty emphasizes alumni and faculty's achievements in scientific research. However, its indicators are not without problems. First, it uses the numbers of Nobel Prize and Fields Medal winners as sole indicators for those evaluated aspects. However, the two awards acknowledge only traditional academic disciplines such as physics, chemistry, biology, medicine, mathematics, and economics; they under represent the highly diverse and expanding academic fields. Second, the indicators (winning Nobel or Fields Medal) strongly favor extremely outstanding achievement and under represent the wider range of scholarly achievement. That is, the methodology may effectively single out a few extremely outstanding universities but may fail to distinguish the performances of regular universities, which are the majority of the world's university population. Third, whether having prize winners in its faculty indicates a university's research performance is arguable. A university can recruit a winner through head hunting and immediately gets advantaged in ranking, but it may indeed have no direct contribution to that winner's research achievement.

Also, ARWU uses SCI (Science Citation Index)/SSCI (Social Science Citation Index) papers and papers published in *Nature* and *Science* as indicators of research output. However, the SCI/SSCI paper indicator over emphasizes the quantity of output (numbers of published papers) and fails to measure output quality (the citations/uses to those papers). The *Nature*/*Science* indicator has the same problems with the prize winner indicators; it over emphasizes extremely outstanding research and biases toward certain subject disciplines. Finally, the size of an institution is a questionable criterion. Insufficient or lack of university data may erroneously affect the judgment of a university's size. The various definitions of academic staff in different universities can distort the measurement relating to institution size and cause comparison validity problems in the resulted ranking.

In 2007, it started to provide field-based ranking for five subject fields: Natural Science and Mathematics (SCI), Engineering/ Technology and Computer Science (ENG), Life and Agriculture Science (LIFE), Clinical Medicine and Pharmacy (MED), and Social Science (SOC). Moreover, since 2009, institutions are ranked in 5 subjects including Mathematics, Physics, Chemistry, Computer Science, and Economics/ Business (Shanghai Ranking Consultancy, 2009b, 2009c)

In conclusion, the major feature and

perhaps the biggest problem of ARWU is its over-emphasis on extremely outstanding research. These indicators cannot differentiate the wider range of research performance wherein most regular universities lie. In the other word, it fails to representatively assess and rank the majority of the world universities.

3.3. The THE-QS Ranking

From 2004, the Times Higher Education (THE) began to publish the world universities ranking using annual data collected and analyzed by Quacquarelli Symonds (QS) Company. In addition to the global ranking, it also published the Asian University Ranking in 2009. The two companies have ceased cooperation in 2010, and the THE-QS ranking split into two independent ranking programs, the QS and the THE rankings. The former continued to use the indicators of the previous THE-QS ranking. The latter began to cooperate with Thomson Reuters to develop new criteria and indicators (Times Higher Education, 2010). The former THE-QS Ranking used both qualitative and quantitative indicators, each accounting for 50% of the final score (see Table 3).

Of the six indicators, the scores of the academic peer review and employer review were obtained from adding up the field ranking score. The field ranking which totally depended on academic peer review and employer

Table 3. The Criteria and Indicators of the THE-QS Ranking for 2009

Criteria	Indicator	Description	Weight
Research Quality	Academic Peer Review	Composite score drawn from peer review (which is divided into five subject areas). 9,386 responses.	40%
	Citations per Faculty	Score based on 2004-2008 research performance searched in Scopus factored against the size of the research body.	20%
Graduate Employability	Employer Review	Score based on responses to recruiter survey. 3,281 responses.	10%
Teaching Quality	Faculty Student Ratio	Score based on student/faculty ratio.	20%
Internationalization	Int'l Faculty	Score based on proportion of international faculty.	5%
	Int'l Student	Score based on proportion of international students.	5%

Note. "Rankings 09: Talking points," by Times Higher Education, Retrieved August 26, 2010 from http://www.timeshighereducation.co.uk/story.asp?storycode=408562

review was separated into five fields: Arts & Humanities, Life Sciences & Biomedicine, Social Sciences, Natural Sciences, and Technology.

The THE-QS Ranking had the following problems. First, peer review accounted for 50% of the criteria. The high percentage of peer review can easily bias the ranking toward universities of international visibility. Second, the questionnaire response rate was too low that may cause validity problem. For example, the response rate was less than 0.1% in 2006 investigation. Also, the sample of questionnaire had bias in the investigation. In 2008, the questionnaire sample of peer review mainly from U.S., United Kingdom, and Australia; countries in The British Commonwealth accounting for 32% in academic peer review, 34.5% in employer review; U.S. 10% in academic peer review, 15% in employer review; Asia countries, including India, Indonesia, Philippines, Malaysia, Singapore, China, Hong Kong, Japan, South Korea, Thailand, Taiwan, respectively, 22% and 17%. Third, its evaluation of research performance relied only on the average number of citations per faculty member. While the citation numbers were objective data, using only average citations numbers can favor universities producing only a small body of papers within which a few were more often cited.

4. Comparing the 2009 Results of the Three Rankings

This section compares the 2009 results of the three rankings (Higher Education Evaluation & Accreditation Council of Taiwan, 2009b; Shanghai Ranking Consultancy, 2009d; Times Higher Education, 2009). Specifically, this paper examines the top 20 universities identified by each ranking, the rankings of the universities in four Chinese speaking regions (Taiwan, China, Hong Kong, & Singapore), and the rankings of the European universities.

4.1. The Top 20 Universities in the Three Rankings

Table 4 lists the top 20 universities identified by the three ranking systems. All of the three rankings indicate the superiority of the U.S. universities in scientific research. 15 of the top 20 universities in the 2009 HEEACT ranking were U.S. universities; 17 of 20 in ARWU, and 13 of 20 in THE-QS. All the three rankings unequivocally considered Harvard University the best university in the world. The leading status of the U.S. academia in the world seems uncontroversial from this comparison. One noteworthy difference is that while ARWU and HEEACT both found the University of California at San Francisco as one of the top 20 universities, the THE-QS ranking did not even include it in the top 500 universities.

Table 4. The Top 20 Universities in the Three Rankings

University Name	HEEACT	ARWU	THE-QS
Harvard University	1	1	1
The Johns Hopkins University	2	19	13
Stanford University	3	2	16
University of Washington, Seattle	4	16	-
University of California, Los Angeles	5	13	-
University of Michigan - Ann Arbor	6	-	19
Massachusetts Institute of Technology (MIT)	7	5	9
University of California, Berkeley	8	3	-
University of Pennsylvania	9	15	12
Columbia University	10	7	11
University of Toronto	11	-	-
University of California, San Francisco	12	18	-
University of California, San Diego	13	14	-
The University of Tokyo	14	20	-
University of Cambridge	15	4	2
Yale University	16	11	3
University of Oxford	17	10	5
Duke University	18	-	14
Cornell University	19	12	15
University College London	20	-	4
California Institute of Technology	-	6	10
Princeton University	-	8	8
University of Chicago	-	9	7
University of Wisconsin, Madison	-	17	-
Imperial College London	-	-	5
Australian National University	-	-	17
McGill University	-	-	18
ETH Zurich (Swiss Federal Institute of Technology)	-	-	20
University of Edinburgh	-	-	20

Note. "Performance Ranking of Scientific Papers for World Universities 2009,"by Higher Education Evaluation & Accreditation Council of Taiwan, Retrieved September 2, 2010 from http://ranking.heeact.edu.tw/en-us/2009/TOP/100; "Academic Ranking of World Universities - 2009," by Shanghai Ranking Consultancy, Retrieved September 4, 2010 from http://www.arwu.org/ARWU2009.jsp; " Top 200 world universities," by Times Higher Education, Retrieved August 26, 2010 from http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=438

However, three systems showed disagreement in the ranking of European universities in the top 20 universities. The THE-QS ranking obviously favored European universities more than the other two systems. Four of the top five universities in the THE-QS ranking were European universities. Further, ranks of some universities were rather large between the three ranking systems. The University of Oxford was ranked fifth in THE-QS, but it was ranked 17th in HEEACT and 10th in ARWU. Similarly, the Imperial College London was listed as the fifth best university in THE-QS, while in the other two rankings it was not even in the top 20 (ARWU: 26rd; HEEACT: 22th).

Disagreement also lies in the rankings of the Asian universities. In the ARWU and HEEACT rankings, Tokyo University was the only Asian university entering the top 20. But it was ranked as the 22nd in THE-QS. The three systems also show rather huge difference in ranking the Australian National University. The THE-QS ranked Australian National University as the world's top 17th university, while ARWU ranked it as the 59th; and HEEACT the 159th. This suggests that biases of peer review had influenced the ranking results in the THE-QS. Quantitative data based rankings obviously varied greatly from the subjective peer review ranking.

4.2. Rankings of the European Universities

Table 5 shows the numbers of European universities in the top 500 by country. Comparisons of the three rankings showed slight differences among the three rankings. Overall, there were 208 European universities in the top 500 countries in ARWU, 215 in HEEACT, and 214 in THE-QS. Germany and United Kingdom had more universities entering the top 500. However, the THE-QS again strongly favored universities of the United Kingdom. Fifty of the 214 European universities were in U.K., while only 40 and 36 U.K. universities entered the top 500 lists of ARWU and HEEACT. Similar bias also favored Irish universities (8 Irish universities in THE-QS; 3 in ARWU and HEEACT). The rankings also disagreed with each other over the Italian universities. Only 13 universities were included in THE-QS, but 21 and 29 Italian universities entered the ARWU and HEEACT lists.

Aside from the THE-QS biases in the aforementioned three countries, the three rankings seemed to show a good level of consensus on the other European countries' universities in terms of which were able to enter the top 500 lists, although each individual university may get different rank in each ranking system.

Rankings of the German universities by the three systems showed greater differences.

Table 5. Numbers of the European Universities in the Top 500

Countries	HEEACT	ARWU	THE-QS
Germany	45	40	41
United Kingdom	36	40	50
Italy	29	21	13
France	20	23	20
Netherlands	12	12	12
Sweden	11	11	9
Spain	10	11	8
Switzerland	8	8	8
Belgium	7	7	7
Finland	6	5	7
Austria	5	7	5
Greece	5	2	4
Denmark	4	4	4
Norway	4	4	4
Ireland	3	3	8
Portugal	3	2	2
Hungary	2	2	1
Poland	2	2	3
Czech	1	1	3
Russia	1	2	4
Slovenia	1	1	1
Total	215	208	214

Note. Data sources are the same with Table 4.

When observing the top 500 lists, the three rankings included similar numbers of German universities. But when observing the top 200 lists, the numbers varied to a greater extent (20 in HEEACT, 15 in ARWU, 10 in THE-QS). Also, ranks given to each German university by the three systems varied greatly both on the national and the global scale. Table 6 shows the rankings by the three systems in top 200. Comparisons showed that some universities were considered the top universities by all the three rankings, i.e., the Technical University

Table 6. Global Ranks and Country Ranks of the German Universities in the Top 200 of the Three Ranking Systems

Germany	HEEACT	ARWU	THE-QS
University of Munich	42(1)	55(1)	98(4)
University of Heidelberg	66(2)	63(3)	57(2)
Technical University of Munich	103(3)	57(2)	55(1)
Humboldt University of Berlin	108(4)	-	146(6)
University of Tübingen	122(5)	135(9)	149(7)
University of Erlangen-Nuremberg	131(6)	206(15)	317(22)
University of Mainz	137(7)	147(11)	-
University of Frankfurt	139(8)	106(7)	-
University of Freiburg	140(9)	102(6)	122(5)
Free University of Berlin	141(10)	-	94(3)
University of Göttingen	150(11)	90(4)	186(10)
University of Hamburg	152(12)	183(13)	-
University of Bonn	153(13)	98(5)	-
University of Würzburg	158(14)	124(8)	-
University of Münster	166(15)	140(10)	-
University of Cologne	171(16)	174(12)	-
University of Düsseldorf	184(17)	-	-
RWTH Aachen University	188(18)	-	182(8)
University of Kiel	-	184(14)	-
Karlsruhe Institute of Technology	-	-	184(9)

Note. Data sources are the same with Table 4.

* The numbers outside the parenthesis were global ranks, those inside were the country ranks.

of Munich, the University of Munich, & the University of Heidelberg. But the ranks of some universities given by each system varied to a greater extent. For example, the Humboldt University of Berlin and Free University of Berlin were ranked well in HEEACT and THE-QS, while ARWU excluded it from the world's top 200. Other examples included the University of Mainz and University of Frankfurt; both were ranked well in HEEACT and ARWU but not even included in THE-QS. Some universities were ranked similarly in the three rankings, e.g.,

Table 7. The Ranks of the UK Universities in the Three Rankings in TOP 200

UK	HEEACT	ARWU	THE-QS
University of Cambridge	15(1)	4(1)	2(1)
University of Oxford	17(2)	10(2)	5(3)
University College London	20(3)	21(3)	4(2)
The Imperial College of Science, Technology and Medicine	22(4)	26(4)	5(3)
The University of Manchester	55(5)	41(5)	26(7)
The University of Edinburgh	61(6)	53(6)	20(5)
King's College London	63(7)	65(8)	23(6)
University of Bristol	92(8)	61(7)	34(8)
University of Birmingham	107(9)	94(11)	66(10)
The University of Glasgow	114(10)	144(15)	79(13)
University of Nottingham	127(11)	83(10)	91(16)
The University of Sheffield	136(12)	81(9)	82(14)
University of Southampton	146(13)	166(17)	95(17)
University of Newcastle upon Tyne	149(14)	-	-
University of Leeds	157(15)	137(13)	99(18)
University of Liverpool	177(16)	110(12)	137(22)
University of Durham	181(17)	187(19)	103(19)
Cardiff University	187(18)	153(16)	135(21)
University of Dundee	192(19)	-	-
University of Sussex	-	140(14)	166(27)
University of Warwick	-	176(18)	58(9)
University of Leicester	-	191(20)	196(29)
Queen Mary, U. of London	-	193(21)	164(26)
University of East Anglia	-	193(21)	-
University of St Andrews	-	199(24)	87(15)
London School of Economics	-	-	67(11)
University of York	-	-	70(12)
University of Aberdeen	-	-	129(20)
University of Bath	-	-	144(23)
Newcastle University	-	-	158(24)
Lancaster University	-	-	162(25)
University of Reading	-	-	191(28)

Note. Data sources are the same with Table 4.

* The numbers in the parenthesis were the country ranks.

University of Hamburg, University of Bonn, University of Würzburg, University of Münster, and University of Cologne, etc.

The UK ranking (see Table7) showed similar discrepancy as the Germany ranking in numbers and ranks. In HEEACT, 19 UK universities entered the top 200, 24 in ARWU, and 29 in THE-QS. University of Cambridge was ranked as the 15th in HEEACT; 4th and 2nd in ARWU and THE-QS respectively. Similar ranking differences occurred with the University of Oxford. Five universities in ARWU and THE-QS top 200 were not included in HEEACT. On the Contrary, University of Newcastle upon Tyne and University of Dundee, ranking 149th and 192nd in HEEACT, were not included in ARWU and THE-QS. Some universities were ranked more differently in THE-QS. For instance, the Imperial College of Science, Technology and Medicine was ranked as 22nd in HEEACT, and 26th in ARWU; but it was ranked as the top 5th in THE-QS. Moreover, there were seven UK universities that entered the top 200 only in the THE-QS.

Table 8 showed the global and national rankings of the Dutch universities. All the three rankings saw 12 Dutch universities in their lists, although the THE-QS ranking once again differed from the others. For example, the Utrecht University was ranked as the 3rd

Table 8. The Dutch Universities Ranked Within Top 500 in the Three Rankings

Netherlands	HEEACT	ARWU	THE-QS
Utrecht University	56(1)	52(1)	70(3)
Leiden University	67(2)	72(2)	60(2)
University of Amsterdam	69(3)	119(4)	49(1)
Erasmus University Rotterdam	81(4)	196(9)	108(5)
University of Groningen	103(5)	112(3)	138(8)
Free University of Amsterdam	106(6)	137(5)	165(10)
Radboud University Nijmegen	128(7)	174(7)	220(12)
Wageningen University	193(8)	150(6)	155(9)
Maastricht University	203(9)	385(10)	116(6)
Delft University of Technology	238(10)	193(8)	83(4)
Eindhoven University of Technology	344(11)	430(12)	120(7)
University of Twente	407(12)	389(11)	200(11)

Note. Data sources are the same with Table 4.
* The numbers in parenthesis means the country ranks.

nationally in THE-QS, but in ARWU and HEEACT it was ranked as the first. Similar ranking differences can be found in the University of Amsterdam, the Delft University of Technology, the University of Twente, and the Eindhoven University of Technology.

Table 9 shows the rankings of Swiss universities. All the three rankings included eight Swiss universities in the world's top 500. ARWU and HEEACT included exactly the same eight institutions, while THE-QS differed in ranking the University of Fribourg and the University of St. Gallen. The three rankings were slightly different at the national scale. The THE-QS unfavorably ranked the University of Zurich as the 92nd. A contrasting example is the rankings of the Swiss Federal Institute of Technology-Lausanne, which was ranked 124th in ARWU, 167th in HEEACT but 42nd in THE-QS.

4.3. Rankings of the Universities in Four Chinese-Speaking Regions

Table 10 shows the rankings of the universities in four Chinese speaking regions – Taiwan, China, Hong Kong, and Singapore. ARWU, HEEACT, and THE-QS all saw five Taiwan's universities in the top 500 universities, but the included universities were ranked quite differently in the three systems. All of the five Taiwan universities were located between the 150th and the 450th in ARWU. In HEEACT, the National Taiwan University was ranked as the 102nd, while the other four were located between the 300th and the 500th. The National Taiwan University was ranked slightly better in

Table 9. The Swiss Universities Ranked Within Top 500 in the Three Rankings

Switzerland	HEEACT	ARWU	THE-QS
Swiss Federal Institute of Technology- Zurich	53(1)	23(1)	20(1)
University of Zurich	65(2)	54(2)	92(4)
University of Geneva	98(3)	129(5)	72(3)
University of Basel	124(4)	85(3)	108(5)
Swiss Federal Institute of Technology- Lausanne	167(5)	124(4)	42(2)
University of Bern	172(6)	169(6)	193(7)
University of Lausanne	175(7)	243(7)	168(6)
University of Fribourg	430(8)	481(8)	-
University of St. Gallen	-	-	337(8)

Note. Data sources are the same with Table 4.
* The numbers in parenthesis show the rank in country.

Table 10. Rankings of the Universities in Four Chinese Speaking Regions

	University Name	HEEACT	ARWU	THE-QS
Taiwan	National Taiwan University	102	150	95
	National Cheng Kung University	307	262	281
	National Tsing Hua University	347	297	223
	National Chiao Tung University	456	327	389
	Chang Gung University	479	408	-
	National Central University	483	441	401-500
	National Yang Ming University	493	449	306
	National Sun Yat-sen University	-	-	401-500
China	Tsinghua University	144	206	49
	Peking University	147	223	52
	Zhejiang University	179	216	247
	Shanghai Jiao Tong University	216	246	153
	University of Science and Technology of China	222	226	154
	Fudan University	250	315	103
	Nanjing University	292	297	168
	Sun Yat-sen University	346	403	-
	Nankai University	376	403	-
	Sichuan University	393	403	-
	Jilin University	416	459	401-500
	Shandong University	432	398	-
	Wuhan University	468	-	-
	Huazhong University of Science and Technology	491	475	-
	Harbin Institute of Technology	496	475	-
	Lanzhou University	-	408	-
	China Agricultural University	-	421	-
	Dalian University of Technology	-	441	-
	Tianjin University	-	491	401-500
	Tongji University	-	-	401-500
	Xi'An Jiaotong University	-	-	401-500
	Southeast University	-	-	-

(Continued)

	University Name	HEEACT	ARWU	THE-QS
Hong Kong	University of Hong Kong	185	212	24
	The Chinese University of Hong Kong	231	235	46
	Hong Kong University of Science & Technology	325	278	35
	City University of Hong Kong	420	385	124
	The Hong Kong Polytechnic University	470	327	195
Singapore	National University of Singapore	93	140	30
	Nangyang Technological University	276	323	73

Note. Data sources are the same with Table 4.

the THE-QS (95th), while the other four were ranked between 200th and 400th.

While the rankings of Taiwan universities seem fairly consistent in the three rankings, it was not the case for the universities of China, Hong Kong, and Singapore. For example, THE-QS ranked Peking University and Tsinghua University as within the top 100 universities, but it was not the case in ARWU and HEEACT. Disagreement widened further in the rankings of Hong Kong's universities. THE-QS ranked the University of Hong Kong and Hong Kong University of Science and Technology as the top 24th and 35th in the world, while they were ranked as the 212th and 278th in ARWU, and 185th and 325th in HEEACT. The Chinese University of Hong Kong was another controversially ranked institution. It was ranked 46th in THE-QS. However, it was located near the 230th in ARWU and HEEACT rankings. Singaporean universities also experienced similar bigger ranking differences. The National University of Singapore and the Nangyang Technological University were ranked as 30th and 73rd by THE-QS, but ARWU ranked them as 140th and 323rd; HEEACT 93rd and 276th. This suggests that the THE-QS ranking had impressionistically favored the universities in these three regions.

A noteworthy finding is that, although China's universities were ranked better in THE-QS, 7 of the 11 Chinese universities within the top 500 universities experienced rank drop from 2008 to 2009. For example, Nanjing University was ranked 143rd in the 2008 THE-QS ranking, but 168th in 2009. In contrast, all the Taiwan universities in the top 500 rose up in ranking from 2008 to 2009, e.g., National Cheng Kung University was the 354th in 2008 and 281st in 2009. The rank rise and drop among the Chinese

and Hong Kong universities was inconsistent between the three rankings.

5. Conclusion

The comparisons in this paper revealed that ranking results can vary, sometimes dramatically, due to methodologies and emphases of various criteria. Peer review can impressionistically favor certain universities and produce results drastically different from quantitative data-based rankings. University College London is a good example. It was ranked 4th by THE-QS, only 20th in HEEACT, and not included in top 20 in ARWU. A leap in ranking also occurred to Imperial College London, which ranked 5th by THE-QS, and is not included both in top 20 by HEEACT and ARWU.

One can argue that THE-QS employs measures more holistic than the ARWU and HEEACT rankings and thus variations are natural. While ARWU and HEEACT both focus only on research performance as shown in bibliometric data, THE-QS has additionally focused on other aspects such as a university's reputation, teaching, and internationalization. However, the major concern here is how ranking can be affected by – and its objectivity suffered from – impressionistic human interference. It is not to say that peer review is an inferior method to bibliometric methods; both methods offer important information for university institutions. Peer review does, however, have its limitations and the results require careful examination. Furthermore, the application of peer review in university ranking/evaluation can be thorny in the relationship-oriented Asian/Chinese cultures, where many academics have seen peer review (PR) and public relation (PR) as twins. In contrast, bibliometric methods are free from human reviewer interference and thus more objective than peer review.

Even when bibliometric data are used as the basis for ranking, the ranking criteria and indicators of a ranking system must be carefully planned in order to generate reasonable and informative assessment. For example, ARWU's indicators overemphasize extreme achievements, and the indicators' applicability and validity is limited in certain traditional subject disciplines. In contrast, the HEEACT indicators seek to reflect a wider range of research performance including both the quantitative and qualitative performances as well as the long-term and short-term research impacts. In contrast to ARWU, which tends to favor a small number of universities already at the top of the world's academic pyramid, and which may fail to represent a wider, diverse scholarly world, the HEEACT system offers a multidimensional assessment of both the top and the "general folks" universities through

carefully weighted use of measures such as highly cited papers, fields of excellence, h-index, and so on. That is, in terms of research performance, the HEEACT ranking may be a fairer and a more informative ranking system for the majority of the world's universities.

The size of a university is closely related to its quantitative performance in all aspects, including research. All the three ranking systems have taken the size factor into consideration to some extent. For example, THE-QS considers the size influences in using the average citations number per faculty member and the ratio of faculty to students. ARWU uses the number of full-time equivalent academic staff to adjust the raw scores of a university in the ranking criteria. However, the designs are not free from problems. For instance, the number of each university's faculty members or the full-time equivalent academic staff may not be accessible to the ranking agency; furthermore, each university may define the academic staff differently, thereby erroneously affect the overall ranking. HEEACT responds to problems by bypassing direct uses of faculty size; instead, it uses the average number of citations, the number of subject fields where a university demonstrates excellent performances, and the *h-index* in the calculation of performance score. The inclusion of these three measures accounts for 40% of the total score and provides a fairer representation of a university's performance regardless of its size.

Both ARWU and THE-QS offer subject discipline based rankings in addition to the overall ranking. ARWU categorizes subject disciplines into five areas: mathematics, physics, chemistry, computer science, and economics/business. THE-QS's categories include arts and humanities, life sciences and biomedicine, natural sciences, social sciences, and technology. Starting from 2008, HEEACT has also implemented subject discipline based rankings and used the six categories supplied by the Current Contents database: agriculture & environment sciences, clinical medicine, engineering, computer& technology, life sciences, natural sciences, and social sciences.

As to the sources of bibliometric data, previously both THE-QS and ARWU used ISI databases such as SCI, SSCI, & ESI. In 2007 THE-QS abandoned ISI databases and used Scopus instead. However, considering the perceived authoritativeness of the bibliometric data sources, HEEACT continues to use ISI databases including SCI, SSCI, JCR, & ESI.

In conclusion, the three ranking systems adopting different criteria and indicators make different ranking results. HEEACT, emphasizing on current research performance, makes fairer ranking than ARWU and THE-QS favoring universities with long histories,

ARWU focuses on university performance in research with outstanding achievement, and THE-QS considering both quality and quantity of universities is the only one with peer review in the three ranking systems. Therefore, readers should recognize different criteria and indicators using in ranking system to interpret the result appropriately.

References

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: A comparative study at a Norwegian university. *Research Evaluation, 13*(1), 33-41.

Aguillo, I., Bar-Ilan, J., Levene, M., & Ortega, J. (2010). Comparing university rankings. *Scientometrics, 85*(1), 243-256.

Bookstein, F., Seidler, H., Fieder, M., & Winckler, G. (2010). Too much noise in the Times Higher Education rankings. *Scientometrics, 85*(1), 295-299.

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71*(3), 349-365.

Campbell, D. F. J. (2002). Conceptual framework for the evaluation of university research in Europe. Retrieved May 9, 2008, from http://www.gwu.edu/~cistp/research/publications/campbell_2002.pdf

Daniel, H. D., & Fisch, R. (1990). Research performance evaluation in the German university sector. *Scientometrics, 19*(5-6), 349-361.

Higher Education Evaluation & Accreditation Council of Taiwan (2008). Performance ranking of scientific papers for world universities 2008. Retrieved September 2, 2010, from http://ranking.heeact.edu.tw/en-us/2008%20by%20field/page/methodology

Higher Education Evaluation & Accreditation Council of Taiwan (2009a). Performance ranking of scientific papers for world universities 2009. Retrieved September 2, 2010, from http://ranking.heeact.edu.tw/zh-tw/2009/Page/Methodology

Higher Education Evaluation & Accreditation Council of Taiwan (2009b). Performance ranking of scientific papers for world universities 2009. Retrieved September 2, 2010, from http://ranking.heeact.edu.tw/en-us/2009/TOP/100

Hong, D. R. (2009). A critical study on the university and academic assessment system in Korea. *Inter-Asia Cultural Studies, 10*(2), 292-302.

Huang, M. H. (2005). Research evaluation of research-oriented universities in Taiwan.

Bulletin of Library and Information Science, 55, 9-23. [Text in Chinese]

Huang, Z. J. (2003). Controversial issues of academic evaluation. *Teacher Welfare 438*. Retrieved May 30, 2008, from http://web.nutn.edu.tw/gac110/presidentsay/20031115.doc. [Text in Chinese]

Leimu, R., & Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology & Evolution, 20*(1), 28-32.

Liu, N. C., Cheng, Y., & Liu, L. (2005). Academic ranking of world universities using scientometrics - A comment to the "Fatal Attraction". *Scientometrics, 64*(1), 101-109.

Liu, Y. (1998). Problem pedigree comparison method of peer review: A new approach of peer review. *Studies in Dialectics of Nature, 14*(10), 32-36. [Text in Chinese]

Kokko, H., & Sutherland, W. J. (1999). What do impact factors tell us? *Trends in Ecology and Evolution, 14*, 382-384.

Kruytbosch, C. E. (1989). The role and effectiveness of peer review. In D. Evered & S. Harnett (Eds.), *The evaluation of scientific research* (pp.69-85). Chichester, N. Y.: J. Wiley.

Makino, J. (1998). Productivity of research groups: Relation between citation analysis and reputation within research communities. *Scientometrics, 43*(1), 87-93.

Meho, L. I., & Sonnenwald, D. H. (2000). Citation ranking versus peer evaluation of senior faculty research performance: A case study of Kurdish scholarship. *Journal of the American Society for Information Science, 51*(2), 123-138.

Norris, M., & Oppenheim, C. (2003). Citation counts and the research assessment exercise V: Archaeology and the 2001 RAE. *Journal of Documentation, 59*(6), 709-730.

Rinia, E. J., van Leeuwen, Th. N., van Vuren, H. G., & van Raan, A. F. J. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria evaluation of condensed matter physics in the Netherlands. *Research Policy, 27*(1), 95-107.

Shanghai Ranking Consultancy (2009a). Ranking methodology. Retrieved September 4, 2010, from http://www.arwu.org/ARWUMethodology2009.jsp

Shanghai Ranking Consultancy (2009b). Ranking methodology. Retrieved September 4, 2010, from http://www.arwu.org/ARWUFieldMethodology2009.jsp

Shanghai Ranking Consultancy (2009c). Ranking methodology. Retrieved

September 4, 2010, from http://www.arwu.org/ARWUSubjectMethodology2009.jsp

Shanghai Ranking Consultancy (2009d). Academic ranking of world universities – 2009. Retrieved September 5, 2010, from http://www.arwu.org/ARWU2009.jsp

So, C. Y. K. (1998). Citation ranking versus expert judgment in evaluating communication scholars: Effects of research specialty size and individual prominence. *Scientometrics, 41*(3), 325-333.

Staropoli, A. (1987). The comite national d' evaluation: Preliminary results of a French experiment. *European Journal of Education, 22*(2), 123-131.

Times Higher Education (2009). *Top 200 world universities*. Retrieved August 26, 2010, from http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=438

Times Higher Education (2010). *Robust, transparent and sophisticated*. Retrieved November 4, 2010, from http://www.timeshighereducation.co.uk/world-university-rankings/2010-2011/analysis-methodology.html

Thomas, P. R., & Watkins, D. S. (1998). Institutional research rankings via bibliometric analysis and direct peer review: A comparative case study with policy implications. *Scientometrics, 41*(3), 335-355.

van Leeuwen, T. N., Moed, H. F., & Reedijk, J. (1999). Critical comments on institute for scientific information impact factors: A sample of inorganic molecular chemistry journals. *Journal of Information Science, 25*(6), 489-498.

van Raan, A. F. J. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. *Scientometrics, 36*(3), 397- 420.

van Raan, A. F. J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics, 62*(1), 117-131.

Wong, B. B. M., & Kokko, H. (2005). Is science as global as we think? *Trends in Ecology & Evolution, 20*(9), 475-476.

(Received: 2011/5/9; Accepted: 2011/5/27)