University rankings: A review of methodological flaws

Muhammad Ashraf Fauzi
Universiti Malaysia Pahang, Malaysia
Christine Nya-Ling Tan
Auckland Institute of Studies, New Zealand
Mahyuddin Daud
International Islamic University Malaysia
Muhammad Mukhtar Noor Awalludin
Universiti Malaysia Pahang, Malaysia

University rankings have gradually become an issue for concern in the academic community worldwide. Several mechanisms with different methodologies have been developed to rank the universities appropriately. However, some ranking tools have notable issues, especially with the indicators adopted. Some are based merely on research performance, whilst others have focused solely on specific fields, such as science and technology – which could have deprived those in the arts and social sciences. This paper uses a narrative review to highlight a number of inconsistencies in the methodologies applied to rank universities. Five main ranking tools commonly applied to the world's universities are reviewed, namely *Quacquarelli Symonds* (QS), *Times Higher Education* (THE), *Academic Ranking of World Universities* (ARWU), *Leiden University* ranking and *Webometrics* ranking. We found that several flaws in the rankings caused inconsistencies in university placings in different rankings. Suggestions for integrating multiple criteria and indicators for better ranking exercises are proposed.

Introduction

Expanding access to education opportunities has raised the demand for better academic quality and attention to reputation issues. This has led to the introduction of university ranking systems (Dill, 2005). Since higher education is a long-term investment that may be costly for some, prospective students demand information from universities or other institutions of higher learning that could help them decide. Since then, an increasing number of universities have competed to become highly placed in university rankings.

Individuals refer to university rankings to decide on future employment opportunities, while investors decide on their contributions (Chan et al., 2015). Stakeholders and policymakers need to evaluate the research performance of universities and make long-term goals and decisions (Huang, 2012). University rankings also improve public awareness of the importance of higher education and help to bring increased transparency into how universities projecting their offers of education mass audiences (Millot, 2015).

Studies have shown that prospective students and their parents rely on university rankings to make choices about higher education (Beine et al., 2014). Governments and non-government agencies prefer to fund universities in the higher ranks with proven records of accomplishment. However, some ranking tools have pressing issues, especially with the indicators adopted. Little work has been carried out to review the university ranking

methodologies. At present, research into university ranking is still in its infancy, and therefore a critical review of the methods used to create rank university rankings is timely.

Research aim and direction

The main objective of this review is to highlight inconsistencies in the methods applied to rank universities. Furthermore, the article raises a voice from the industry and insights from the university perspective. After a definition of ranking, its importance and implications for the universities, the article reviews the five main ranking tools, namely, Quacquarelli Symonds (QS), Times Higher Education (THE), Academic Ranking of World Universities (ARWU), Leiden University ranking and Webometrics ranking.

What is 'university ranking'?

University ranking is an 'organisational report card' that provides unequivocal ranking mechanisms among universities (Aguillo et al., 2010). Rankings impose accountability to members of the public and provide criticism to service providers. Organisational report cards are well-known strategic instruments, in the context of becoming significant in customer decisions, and an increasing impatience of policymakers with instances of low quality and high cost of social service delivery (Gormley, 1999). Therefore, ranking is essential for strategic planning and enhancement of institutional transparency and stimulating the quality culture in education (Berbegal-Mirabent & Ribeiro-Soriano, 2015). From the public policy point of view, ranking systems are inevitable as stakeholders demand to know how funds invested with universities are managed (Marginson & van der Wende, 2007). Leading universities in the United States and the United Kingdom enforce the university ranking system. This allows them to elevate their academic status and strengthen their brands globally. Consequently, these countries could attract young, talented students and generate more income for their universities and the nation as a whole (Marginson, 2014).

Implications of university rankings

University ranking exercises have provided both favourable and unfavourable consequences for universities. Generally, universities have made significant contributions to a country's development towards sustainable economic growth (Jabnoun, 2015). Universities contributed significantly to Taiwan's economy from 1965 to 2000, where a 1% increment of universities' stock resulted in a 0.19% real output increase (Lin, 2004). The study also found that engineering and natural sciences contributed more in terms of nation-building, compared with the humanities fields. Levin (2010) argued that a few countries in East Asia, such as Japan, Singapore, South Korea, Taiwan and Hong Kong, recognised the importance of educated human resources to national economic growth. Therefore, investment in education had reaped the rewards in these countries, particularly with China and India, which are seeking to gain control of the regional economy by 2050.

In the chase to elevate their rankings, universities have been under constant pressure to perform their very best (Mussard & James, 2018). This includes increasing the volume of

publications in reputable academic journals indexed in leading bibliometric data bases, such as *Web of Science* and *Scopus* (Shin et al., 2011, p.10). University managements, as well as governments, aspire to increase the ranking of their universities. This is achieved through increasing human resources, such as the number of researchers and lecturers, and facilities, which are leading indicators in the university ranking exercises (Marconi & Ritzen, 2015). These measures have incurred significant expenditures, which, if improperly managed, could lead to a significant decline in a university's progress. Some reputable universities prefer to invest money in other dimensions that lie in that particular university's vision in terms of principles, socio-economics and culture – which are not considered as essential dimensions or indicators in university ranking systems. In other cases, universities are not able to spend because of inefficiencies in using their available funds. In these circumstances, even though a university spends, it does not improve its ranking.

In terms of impact on student fees, university rankings may come with high costs. A university may tend to charge students more if it becomes higher ranked in the available ranking system. In the UK, for instance, universities are increasing their fees by up to 9000 pounds per year (Broecke, 2015). This is creating an unsustainable environment for future generations seeking a chance to study in prestigious and well-known universities, as higher-ranked universities tend to be costlier. Students are increasingly struggling to cope with the fees. Alongside this, students may become trapped in a high amount of debt even before completing their studies. For some people, it does not matter where they have graduated. What matters is getting the job and having less debt. Sometimes, the lower the rank, the lower the fees, and the lower the minimum requirements. In other words, excellent education for less cost is sometimes more attractive than a high ranking institution and a potentially high amount of debt in the future.

University rankings become a pressure for individual universities because it does affect their activities. Some may respond to the imposed pressure by "overshadowing" their status with historical information with the purpose of "creating narratives that manipulate their rankings to promote their own strengths" (Heffernan & Heffernan, 2018, p.29). The "overshadow psychology" approach may improve public perceptions and help the university compete, but this is a form of "media spin" and ultimately is not a credit to academic society.

University rankings in the private sector may or may not affect graduates' status in finding a job or career advancement in the present times, as many certifications may be regarded as not specifically relevant to the actual skills required for the employment they are seeking (Reddy et al., 2016). However, high expectations from employers towards recruits graduating from high ranking universities are inevitable. When one is out there in the job market, the university's ranking may or may not help. It might help in certain countries, but not all (Pizarro Millian & Rizk, 2018). The labour market outcome for students was among the expected critical implications of university rankings. However, the question of whether a university's high ranking can guarantee the success of its graduates remains unanswered. In some industries, a university's ranking may or may not help a student in

job searching, though for some industries or professions, a university's ranking may play a role, especially in reaching the interview stage of a selection process.

As many graduates may encounter difficulty in finding employment, and may be in debt, some may develop feelings about being trapped in debt as if through an education "scam" in which university ranking was the "bait". In other cases, unless the student gets a full scholarship, getting into a high ranking university could lead to feelings about a waste of money. In the case of Jiangsu in China, for example, Li and Thige (2017, p.183) commented that students who "did not meet their academic obligation within the set timelines were forced to spend more money to finish" (Li & Thige, 2017, p.183).

University ranking may also create a competitive environment among groups of scholars. A study about salary versus ranking has revealed greater inequalities in salary structure among research-orientated institutions, compared to teaching-oriented universities (Roth & McAndrew, 2018). The discrimination appeared to arise from university ranking implications. A university's good ranking may boost the confidence levels of most its academics, but it may incur a risk of academic corruption. A hypothetical example could be a senior lecturer so anxious to attain the title and salary of an associate professor, that he or she is tempted to become named as an author on a student's research paper, though not having made a sufficient contribution warranting authorship. It could be almost impossible to determine whether such an action has occurred, as in many universities postgraduates have to publish as a requirement to graduate. Questioning "How corrupt are universities?", Shore (2018) drew attention to the discourse about corruption in universities.

University rankings may distort public opinion with a misleading goal (Mussard & James, 2018). For example, a student seeks specific criteria to decide on the most suitable university, giving thought to his or her learning style, intellectual interests, and location. Some students may use university ranking as a short cut to finding the perfect fit, or maybe just for the sake of seizing an opportunity for a government scholarship. University ranking is also a marketing strategy to attract international students' (Olcay & Bulu, 2017). University ranking can be misleading; after all, we do live in a society that values brands and labels. University ranking is being used as a tool to reveal information about the university or its students. However, a more useful measurement tool is needed to match global needs, and be a foundation for how, today, we judge the success of a specific university. Exercising caution when making a choice would be the best action since one size does not fit all, and it varies from person to person. It is now unwise to spend a fortune to gain an education, if selection of university was overly dependent on its ranking. A university project its ranking in a way that hides its core limitations. This is because university ranking typically does not have a good capacity to evaluate universities from all angles (Olcay & Bulu, 2017). False judgments about university ranking may occur, due to diversities amongst university ranking methods, so ranking should not be the ultimate guide to decision making about applying to be admitted.

University ranking methodology

A number of organisations provide different methods for measuring university rankings. The Quacquarelli Symonds (QS), Shanghai Jiao Tong University, Taiwan's Higher Education and Accreditation Council and Web visibility by the CSIC Cybermetrics Lab are some of the famous university ranking systems (Aguillo et al., 2010). Although they have received criticisms from universities, they have managed to provide an overall basis for ranking institutional quality, affordability and ability to deliver the best education, research outcomes, and intrinsic values for higher education.

Adaptions of university ranking systems by these independent bodies gives a platform for universities to display their quality and potential, which encourages enthusiasm and commitment towards research productivity (Huang, 2012), though different measurement tools are evident in the various systems. For instance, The Higher Education Evaluation and Accreditation Council of Taiwan (HEEACT) Ranking focuses more on the quality and quantity of institutional research performance. Shanghai Jiao Tung University, which operates the Academic Ranking of universities, emphasises the presence of Nobel Prize winners. The QS World University Ranking has a high weight on peer reviews (Huang, 2011). These inconsistencies have motivated this article's review of flaws and differences in a number of worldwide ranking systems that are listed below.

Quacquarelli Symonds (QS) ranking

Quacquarelli Symonds produces the QS World University ranking. It is known as a top university ranking system worldwide. QS has published its university rankings from 2004 to 2009 in collaboration with Times Higher Education World University ranking (Quacquarelli Symonds, 2016). In the 2016/2017 ranking, QS assessed more than 3,800 universities and ranked more than 900 of these. The universities were ranked in the top 400s with individual rankings, and the next was ranked in groups starting from 401-410 until 701+. The published rankings were categorised in an interactive manner where users can sort according to country/region with the indicator performance of each university listed. QS was the first international ranking system to be audited and approved by the Observatory on Academic Ranking and Excellence (IREG Observatory) in 2013, making it one of the most trustworthy systems. The QS World University Rankings, QS University Rankings: Asia and QS University Rankings: Latin America are "IREG approved."

QS ranks universities based on six leading indicators. Each of them carries different weights. Four indicators rely on hard data, which comes from each participating university, with two indicators being based on a global survey. The global survey was taken from academics and employers, which were the most extensive data of this kind. In the 2016/2017 edition, QS obtained a survey from 74,651 academics worldwide from 140 different countries. These votes included five previous years forwarded if no recent data came from the same respondent. A respondent can nominate up to 30 universities excluding their own. The median nomination was about 20, which resulted in an overall dataset of 500,000 points. QS argued that the survey is reliable, with respondents'

academic experience averaging about 20.4 years, and 81% having more than ten years of experience. As for the employer review, QS sampled 37,781 employers from 130 countries who recruit graduates from around the world.

The QS ranking categorises universities by the number of stars. The "star rating" is different from rankings in that it includes all assessments that are not recognised or overlooked in the ranking, and it suits universities that have distinct strengths which are not evaluated in the rankings. The star rating ranges from one to five stars - with five being the highest and one being the lowest. This star granting method is an alternative for universities to be well recognised as it avoids focusing excessively on reputation and research. A specific and thorough way of assessment based on the multiple indicators has emerged from the star rating, which provides a holistic approach to university excellence.

The surveys employed in the QS constitute 50% of its indicator. The survey serves as a market research tool, which opens a debate on response rate, fairness, and weights, which remain questionable (Marginson, 2014). Also, regarding the information on graduate employers, the survey results have never been revealed (Marginson, 2014). The QS ranking has two major flaws. Firstly, due to the multiple indicators of the QS. universities that are close to one another in terms of scores and indicators are in low correlations. Slight changes in one of the indicators may cause drastic changes in a university's position. Because of the low correlation indicator, random errors might occur. Secondly, with half of the score involving surveys, it is subjected to year by year changes. This will induce the universities to change their strategies, notably marketing of a particular university so that it will be more visible to the potential respondent; thus, it will have more favourable survey returns for itself.

Times Higher Education (THE) ranking

Times Higher Education (THE) ranking was founded in the year 2004. The tools have thirteen indicators for assessing university performance, which, among others, comprise teaching, research, citations, international outlook, and industry income (Times, 2016). The 13 different indicators used in THE show a high level of comprehensiveness in university evaluation. Times ranking is the only tool that incorporates teaching as one of its indicators, weighted at 30%, which is high compared to other tools. Times ranking considers publications from journal articles and conference proceedings in the Elsevier's *Scopus* databases. Times ranking excludes universities not teaching undergraduates, and universities having less than 200 articles per year during the period from 2011-2015 (Times, 2016).

Similar to the QS, THE applies surveys as the main component in assessing teaching and research. Weighted at 15% and 18% respectively, the surveys contribute a total of 33%. This percentage is still considered high, even though it is less than the QS, which is 50%. The survey method is open to debate on the bias and random errors of data collection, which can be easily manipulated. The income component for research in the THE ranking is incomplete and confusing over standardising among different countries (Marginson, 2014).

It is claimed that THE is one of the famous ranking systems because of the reputational survey it uses. However, data in the survey obtained from university staff is a clear threat to the ranking's reliability owing to validity concerns (Waltman et al., 2012). This is addressed in the next part of this paper. Most academics know in detail only their internal networking and only a few reputable universities. They may know the top ten universities in the ranking, but further down the list there may be a prestigious and reputable university, which they do not know because it has concentrated on different fields not relating to their areas of expertise. The great majority of university rankings may be based mainly on universities' images attained through their marketing in a few countries, and also on hearsay evidence that might be misleading.

One of the main problems in THE ranking is normalisation. According to THE, data are normalised to account for the variation in the volume of citations. The variations spread across diversified areas and field subjects. THE claims that universities with high research activities in subjects with areas having a high number of citations, do not gain an advantage over those with less cited subjects, taking, for example, engineering over history. This claim was disputed by Waltman et al. (2012), as the issue of procedure had never been documented precisely. Apart from citations, 52.75% of the overall data were normalised, where this took into consideration the number of PhDs awarded to staff, institutional income, research income, research productivity, and international collaboration (THE, 2016). This is a problem where more than half the percentage of the indicators has been normalised.

Academic Ranking of World Universities (ARWU)

Academic Ranking of World Universities (ARWU), commonly known as the Shanghai Ranking, was the first ranking tool available worldwide. Initially published by Shanghai Jiao Tong University, it is currently published by Shanghai Consultancy. ARWU released its first university rankings in 2003. ARWU focuses on four main criteria, quality of education, quality of faculty, research output, and per capita performance. Each of the four main criteria has indicators with different weights being given. The quality of education utilises the indicator of a university's alumni winning prestigious Nobel Prizes and Fields Medals. The qualities of the faculty or academics have two indicators: staff who had won the Nobel Prizes and Field Medals, and having citations categorised in 21 different subjects. Research output provides two indicators: published papers in *Nature* and *Science* and indexed papers in the *Science Citation Index and Social Sciences Citation Index*.

The ARWU has an indicator of the per capita academic performance of a university, the criteria being the weighted scores of a combination of five indicators relating to the number of the full-time academic staff of a university. It is the only one that is size-adjustable and is considered to be a measurement of efficiency.

The highest score that a university could earn in this ranking is 100. ARWU evaluates more than 1200 universities, but only publishes the top 500 universities on its website. ARWU does not rank arts and humanities fields due to the difficulties in obtaining comparable indicators using a completed, reliable dataset. Several cross-disciplinary fields,

for example psychology, are excluded because of their interdisciplinary nature, which cannot be precisely distinguished (Liu, 2009). The data used in ARWU are all objective, which means data that can be counted and observed. ARWU is not wholly comprehensive, as it excludes essential aspects of universities, namely teaching, social work, and internationalisation (Marginson, 2014).

There have been many criticisms about ARWU. Firstly, the ranking weights are arbitrary, which means the indicator score is fixed. One of the significant shortcomings in the ARWU is that all the data are collected by acquiring them through websites. This includes the official site of Nobel Prizes, the official site of the International Mathematics Union, and various Thomson Scientific sites. The authors of ARWU (2003-2009) did not reveal the specification of the data obtained from the sources, i.e., for the number of the academic staff of each university in the ranking (Billaut et al., 2010).

The majority of universities tend to score less with ARWU, primarily due to the Nobel Prize weight, which contributes 30% of the score. Only a few countries in the past have Nobel award recipients, which included the USA, Europe, and from Asia, only Japan. The Nobel Prize could not be said to be a good reflection of university status and performance, which relatively divides the universities into elite and non-elite, which has nothing to do with rankings. Apart from only using Nobel Prizes and Field Medals, other scientific awards should be included as well (Zitt & Filliatreau, 2007).

ARWU focuses only on the research dimension, without any reference to teaching quality. A reputable university would have given attention to the quality of its students through its development of quality of teaching. ARWU also neglects to consider the employability dimension as an essential indicator, given the fact that the great majority of the students do not become academics after graduation (Saisana et al., 2011).

The basic methodology in ARWU is that it applies a nonlinear transformation of the best performing university and other universities are normalised according to the best performer. It is classified as score-driven tables. Docampo and Cram (2014) have proven that ARWU has poor score-driven rankings by identifying the numerical values of the expected power laws and gains in the method. They recommended that "... utility and usability of the ARWU could be greatly improved by replacing the unwanted dynamical effects of the annual re-scaling based on raw scores of the best performers" (Docampo & Cram, 2014, p.1347).

Leiden Ranking

Leiden University Ranking is the only current ranking published by a university. It is based on the Thomson Reuters' *Web of Science* database, which includes all the sciences, social sciences, and arts and humanities. It excludes all the publications in books, conference proceedings, and journals not indexed in the *Web of Science*. It includes only the core publications from the *Web of Science*, which considers only two document types, articles and reviews. The fundamental methodology of this ranking depends solely on the bibliometric data obtained from the database.

Leiden ranking is a ranking based entirely on research productivity. It is known as a ranking based on citations, strictly limited to excellence in research (Boulton, 2011). It includes the science papers' volume (number), citations of the papers, per paper citations, the number of papers in the top 10% in the area by frequency of citations, and the proportion of a university's publications in that 10% region. Even though Leiden ranking is less known compared with QS and THE, it is easier to read and more transparent (Marginson, 2014). A university's ability to produce more research output would enable them to score higher in the Leiden Ranking, which implies that fundamentally strong research universities can compete in this ranking.

The Leiden rankings are superior to other rankings based on several factors, as pointed out by Docampo and Cram (2014). Firstly, it replaces the highly-cited researchers with the counts of an university's highly cited publications. Secondly, it adopts a method to count, fractionally, for publications that are produced by collaboration. Thirdly, it explores the possibility of including non-English publications. These show that the Leiden ranking is the best method for measuring university publications and citation-based ranking compared with other systems.

According to Waltman et al. (2012), Leiden ranking provides several advantages compared with other ranking systems. Firstly, it refrains from combining multiple dimensions of the university's performance in a single aggregate indicator. Secondly, Leiden ranking does not apply questionable data using surveys and data given by universities. Thirdly, Leiden ranking is presumed to be documented extensively, making it more transparent.

Leiden ranking focuses only on the output of research (i.e., publications and citations). The ranking does not take into account inputs of the research, such as academic staff and research income or grants (Fauzi et al., 2019; Waltman et al., 2012). An ideal research activity performance should take the input and output processes as indicators for a university's excellence. The most probable reason for the Leiden ranking not including the input is due to it being incapable of acquiring accurate data for the university's research input. No international organisation or agency, such as Thomson Reuters and Scopus (which measure the research output), can be accessed by Leiden ranking to obtain the data, which makes the assessment challenging.

Webometrics

The *Ranking Web* or *Webometrics* is one of the most extensive academic rankings available, with nearly 12,000 universities included in the list - more than any other ranking tool available. This ranking is provided by Cybermetrics Lab, an agency under the Spanish National Research Council (CSIC), which started its ranking process in 2004. Universities are included based on an institution's web presence and its impact on education (Webometrics, 2016). According to Aguillo et al. (2010), there are significant "reasonable similarities" between Webometrics rankings and other ranking tools in listing the university's ranking, where there is not much difference in the institution's positioning, despite using a completely different set of methods. This is even more prevalent when rankings are compared using European universities, suggesting that the Webometrics

method, based solely on the university's web access, presence, and visibility, is a reliable tool.

A unique indicator in the Webometrics ranking is that it uses data from pages retrieved via significant search engines (Yahoo, Google, and MSN) together with the unique sites in total that are externally linked to a university (Lukman et al., 2010). This has created a significant shortcoming in Webometrics, namely that many universities do not have policies for uniformly naming their affiliated institutions, thus resulting in impaired identification due to different host domains (Aguillo et al., 2010). This is an issue in Webometrics because different universities apply different approaches in naming websites and sometimes are customised according to the socioeconomics, cultures, and nation-building approaches of a country.

Using only website visibility as the ranking criteria has resulted in a major flaw in Webometrics. A university could be ranked high in Webometrics when it is marketing vigorously through its website, compared to other universities. A good university may not have good marketing through its website, thus degrading its ranking in Webometrics.

Summary of university rankings

Table 1: Summary of university rankings

Indicators	QS 2018	TIMES 2018	ARWU 2018	Leiden	Webometrics
No. unis ranked	1,000	1,400	1,000	963	11,997
Method-ology	6 metrics	13 performance indicators	6 objective indicators	Bibliometric sci-entific impact indicators	Web presence and impact
Owner	Quacquarelli Symonds	Times Higher Education	Shanghai Jiao Tong Uni.	Leiden University	Cybermetrics Lab
Advan-tage	Included other rankings such as by subject and region	Includes teach-ing in ranking. Comprehensive indicators	Strength in research aspects	Includes only the core public-ations from *Web of Science*	Institution web presence and its impact on education
Disadvan-tage	40% comes from academic reputation	Normalisation problem in subject differences	Limited unis worldwide having Nobel winners	Limitation of diversity in ranking university	Universities low in marketability will be ranked lower

Discussions and findings

For an empirical comparison, Table 2 lists the top 10 universities ranked according to QS ranking 2019, together with rankings for each according to the four other tools. While THE and QS show relatively small differences, Leiden and Webometrics show large differences from QS. The most substantial difference was for California Institute of Technology, where the Leiden ranking positioned it in 193rd place, very different from QS and THE, fourth and third respectively. Leiden ranking included only publications and

citations counts, in contrast with QS and THE, which included feedback from academic staff and a global survey (Waltman et al., 2012).

Table 2: Other rankings for top 10 universities in the world according to QS ranking

University	QS 2018	TIMES 2018	ARWU 2018	Leiden	Webometrics
MIT	1	5	4	54	3
Stanford	2	=3	2	10	2
Harvard	3	6	1	1	1
Caltech	4	=3	9	193	34
Cambridge	5	2	3	21	10
Oxford	6	1	7	13	7
UC London	7	16	17	15	18
Imperial	8	8	24	40	59
Chicago	9	9	10	98	26
SFITZ	10	=10	19	71	27
MIT - Massachusetts Institute of Technology; Caltech - California Institute of Technology; UC London - University College London; Imperial - Imperial College London; SFITZ - Swiss Federal Institute of Technology Zurich					

For the top 10 universities, large differences are shown in the positioning. Increasing contrasts and discrepancies seems to occur when proceeding down the rankings. It was further reported that universities from developing countries in the Asian and African regions find it challenging to compete, which creates the notion of "unfairness" in rankings. It is not practical to say that ranking systems should be abolished, but rather ask how can ranking systems be further improved with minimisation of its adverse effects, to provide the best possible information for comparing universities.

Ranking systems have drawn many criticisms and comments due to their statistical flaws. Inaccuracies in methodology and measures for assessing the quality of academic and scholastic outputs, which have negative consequences for the general achievement of universities, have been debated extensively (Fauzi et al., 2018; Bowden, 2000). Technically, the bases of methodology in all of the significant ranking systems are similar, as they focus on research activities of IHL, which tend to be accorded the most weight in ranking processes.

Some have suggested that the best action for improving university rankings is to incorporate measures known to be associated with student learning outcomes esteemed and valued by society (Dill & Soo, 2005). Student outcomes that give positive influences to society are the best contributions to human capital (Becker, 1964). The development of human capital among students would enable their maximum contributions to society.

Criticisms of university ranking

It was observed that ARWU, QS, and THE appeared to provide rankings "... in favour of greater system stratification and the concentration of elite researchers" (Marginson & van der Wende, 2007, p.306). These rankings were fabricated to suit institutions that are

science-based and significant English language universities. ARWU, in particular, only includes publications in *Nature* and *Science*, suggesting that this ranking does not take into perspective other fields such as humanities and arts. Though ARWU has stated that they have difficulties in obtaining such data, it does not imply the ranking has done enough.

Bibliometric data that consist of numbers of publications and citations of faculty members tends to dominate the weightings (van Raan, 2005). This has created the issue of English major universities dominating the rankings because all take their bibliometric data from major English publishers (Amsler & Bolsmann, 2012). Certain countries have much of their academic publishing in the native language, which could have a considerable impact on research rankings. This includes countries publishing in French (France and Canada) and Japanese, for example, and are in countries very advanced in technology and research activities.

It is therefore submitted that rankings contain under-representations and cannot reflect accurately the capabilities of universities and countries. Take Germany, for instance, accepted by the world in the fields of engineering and technology, but German universities do not enter into the world's top rankings. The ranking criteria are not measuring Germany's adequately (Hazelkorn, 2009).

Daraio et al. (2015) suggested a model that could address four criticisms of university rankings, including (1) mono-dimensional nature; (2) lacking statistical robustness (3) overly dependent on a university's subject mix and size; and (4) lacking full consideration of input and output structure.

One of the reasons that different university rankings give different results is differences in their methods for values of ranks being normalised (Aguillo et al., 2010). Peer review and bibliometrics are the two most influential methods for academic evaluation and assessment, with peer review as qualitative and bibliometric as quantitative (Geuna & Martin 2003; Huang 2012).

With the rapid development of ranking systems worldwide in recent years, there has been debate and intense, cross-border criticisms between different regions of the world. Countries are different in their history, traditions, and culture and different perspectives upon university establishment. This prompts arguments from many authors of comparisons of rankings from different countries, which are problematic because of diversity in sociocultural and politico-economic influences (Berbegal-Mirabent & Ribeiro-Soriano, 2015). Some universities have other important foci besides research-based education. For example, universities may have an extensive focus on teaching and learning which can be an excellent contribution to nation-building, though it may lack proper weighting in ranking criteria.

Research activities

As pointed by Vidal and Filliatreau (2014), it may be considered that a "good" university is based on research activities with indicators on research that can be obtained correctly on the international level (Vidal and Filliatreau 2014). The databases that have been used for

rankings, such as Thomson Reuters and *Scopus*, are based on university research activities, but databases are not available for other activities, such as teaching and community service.

Several rankings use Elsevier's *Scopus* as the bibliometric data source, with over 51 million citations with 11.3 million journal articles published over five years. Data were extracted from more than 23,000 journals indexed within *Scopus*. THE citation includes journals, conference proceedings, and review articles, which is different from Leiden ranking, which excludes conference proceedings. Taking only journals and omitting conference proceedings can give a considerable difference in scoring for a university.

Nobel Prize and Fields Medal

The most controversial of the criteria is Nobel Prize winners (Marginson & van de Wende, 2007). It is debatable due to the claim by some that decisions are determined by a "politicking factor" and merit of the scientific contribution is not the key factor in determining the winner. The prize is also based on submission, but sometimes a university may have had a notable candidate who was suitable for the prize, but was not submitted for reasons relating to the university's philosophical and cultural values.

The Nobel Prize is awarded to a person long after research has been conducted. This is problematic when the Nobel Prize is included in a university ranking. For instance, Albert Einstein's Nobel Prizes were awarded when he was an academic at the University of Berlin, though his research work on photoelectric effects was conducted when he was working in Zurich at the Swiss Patent Office (Billaut et al. 2010).

ARWU is the only tool currently that awards points based on faculty members and alumni being awarded a Nobel Prize. This is the major flaw in ARWU where Nobel Prizes and Fields Medals are science-based and mathematics-based respectively. Even though the ARWU made several changes since the first inception, it is still considered biased towards the science stream when Nobel Prizes and Field Medals are included in the ranking system. If these are to stay, some other awards in all fields should be considered even though not as prestigious as the Nobel Prize.

Survey method

The survey conducted by THE as well as QS is based on a small sample and under-represented numbers of countries, which means the results are biased based towards a limited number of countries (USA, UK, or Australian universities especially). (Aguillo et al., 2010). THE and QS are based in the UK and US, respectively, which means that their sampling of the respondent could be covering only certain countries, owing to accessibility for date collection. It is unfair to other countries, especially the developing countries in Asian and Africa, which methodologically should be evenly represented. Rankings which use surveys as the main tool could be subject to psychological, sociological, and anthropological limitations. Rankings that use composite indicators, such as ARWU, QS and THE (Vidal and Fillatreau, 2014), may contribute to problems with the reputation of surveys. A more valid and reliable system could be use holistic statistical

tools that integrate with the survey method. Soh (2016) stressed that ranking must be based on statistical grounds; without it, the discussion would be based on abstract and verbal only.

Peer review

Peer review is where experts in particular fields of knowledge and area of competence evaluate and appraise the output of professional and scientific work by others in the domain of performance, creativity, and quality (Lee et al. 2013; Dobrota et al. 2016). Several biases have been identified in the peer-review process, including affiliation, nationality, language, and gender bias (Lee et al., 2013). With the majority of journals being based in Western countries, biases in the evaluation of articles submitted by authors from Asia and Africa may occur, if prioritisation is accorded to articles from Western region authors. For instance, the USA has the most universities ranked in the top 100 of many ranking tools, perhaps especially those where peer review weighting is high.

Service to community

Universities build knowledge-bases for their communities, which will benefit from the knowledge being developed, distributed and applied by their university. This aspect is not included in any of the ranking systems available, although academics are experts in their fields, which in many ways can be utilised by their local and regional communities (Fauzi, 2019). It is considered a blessing when an IHL is brought up to a specified location.

Conclusion

This paper concludes that the ranking exercises applied by the above tools can be prone to preferences and favouritisms in terms of cultural, social and historical values (Berbegal-Mirabent & Ribeiro-Soriano, 2015; Liu (2009). Universities that are prestigious and have previous historical value tend to have higher ratings, in part due to global survey problems. English and Western universities tend to have higher rankings than their Asian and African counterparts, reflecting the possibility that survey respondents tend to think that English is always better than the rest. Differences in each university's key objectives in research, teaching and community service, and the great diversity of programs and courses offered in universities worldwide, have contributed to inconsistencies in ranking systems. A university which is poorly ranked may be excellent in teaching or in other qualities that contribute to nation-building compared to universities with have higher rankings. Not all tools have comprehensive metrics and weights that represent all the indicators on the respective target concepts.

Therefore, no ranking tool should be considered as perfect, and continuous improvement should be called for. Relevant experts must re-examine the methodologies adopted by the major ranking systems, because decisions by prospective future students, research funding bodies, and others, should be made with informed reference to all indicators, avoiding over-reliance on university rankings. This study hopes to encourage researchers to develop more comprehensive and holistic ranking tools, aiming especially to better integrate all the

factors of cultural and cross-border heritage values that reflect a university's real mission, vision, and objectives.

References

Aguillo, I. F., Bar-Ilan, J., Levene, M. & Ortega, J. L. (2010). Comparing university rankings. *Scientometrics*, 85(1), 243-256. https://doi.org/10.1007/s11192-010-0190-z

Amsler, S. S. & Bolsmann, C. (2012). University ranking as social exclusion. *British Journal of Sociology of Education*, 33(2), 283-301. https://doi.org/10.1080/01425692.2011.649835

Berbegal-Mirabent, J. & Ribeiro-Soriano, D. E. (2015). Behind league tables and ranking systems: A critical perspective of how university quality is measured. *Journal of Service Theory and Practice*, 25(3), 242-266. https://www.emerald.com/insight/content/doi/10.1108/JSTP-04-2013-0059/full/html

MCDM view. *Scientometrics*, 84(1), 237-263. https://doi.org/10.1007/s11192-009-0115-x

Becker, G. (1964). *Human capital*. New York: Columbia University Press.

Beine, M., Noël, R. & Ragot, L. (2014). Determinants of the international mobility of students. *Economics of Education Review*, 41, 40-54. https://doi.org/10.1016/j.econedurev.2014.03.003

Billaut, J. C., Bouyssou, D. & Vincke, P. (2009). Should you believe in the Shanghai ranking? An Boulton, G. (2011). University rankings: Diversity, excellence and the European initiative. *Procedia - Social and Behavioral Sciences*, 13, 74-82. https://doi.org/10.1016/j.sbspro.2011.03.006

Bowden, R. (2000). Fantasy higher education: University and college league tables. *Quality in Higher Education*, 6(1), 41-60. https://doi.org/10.1080/13538320050001063

Broecke, S. (2015). University rankings: Do they matter in the UK? *Education Economics*, 23(2), 137-161. https://doi.org/10.1080/09645292.2012.729328

Centre for Science and Technology Studies (2011). *Leiden Ranking 2016*. [viewed 7 September 2016]. http://www.leidenranking.com/ranking/2016/list

Chan, K. C., Fung, A., Fung, H .G. & Yau., J. (2016). Ranking of institutions and academic journals: A selective review and a conceptual framework. *Managerial Finance,* 42(4), 292-302. https://doi.org/10.1108/MF-09-2015-0247

Daraio, C., Bonaccorsi, A. & Simar, L. (2015). Rankings and university performance: A conditional multidimensional approach. *European Journal of Operational Research,* 244(3), 918-930. https://doi.org/10.1016/j.ejor.2015.02.005

Daraio, C. & Bonaccorsi, A. (2017). Beyond university rankings? Generating new indicators on universities by linking data in open platforms. *Journal of the Association for Information Science and Technology*, 68(2), 508-529. https://doi.org/10.1002/asi.23679

Dill, D. D. & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education*, 49(4), 495-533. https://doi.org/10.1007/s10734-004-1746-8

Dobrota, M., Bulajic, M., Bornmann, L. & Jeremic, V. (2016). A new approach to the QS university ranking using the composite I-distance indicator: Uncertainty and sensitivity analyses. *Journal of the Association for Information Science and Technology*, 67(1), 200-211. https://doi.org/10.1002/asi.23355

Docampo, D. & Cram, L. (2014). On the internal dynamics of the Shanghai Ranking. *Scientometrics*, 98, 1347-1366. https://doi.org/10.1007/s11192-013-1143-0

Fauzi, M. A., Tan, C. N. L. & Ramayah, T. (2018). Knowledge sharing intention at Malaysian higher learning institutions: The academics' viewpoint. *Knowledge Management & E-Learning: An International Journal*, 10(2), 163-176. https://doi.org/10.34105/j.kmel.2018.10.011

Fauzi, M. A., Tan, C. N. L., Thurasamy, R. & Ojo, A. O. (2019). Evaluating academics' knowledge sharing intentions in Malaysian public universities. *Malaysian Journal of Library & Information Science*, 24(1), 123-143. https://doi.org/10.22452/mjlis.vol24no1.7

Fauzi, M. A. (2019). Knowledge sharing in Asia Pacific via virtual community platform: A systematic review. *International Journal of Web Based Communities*, 15(4), 368-394. https://doi.org/10.1504/IJWBC.2019.103186

Geuna, A. & Martin, B. R. (2003). University research evaluation and funding: An international comparison. *Minerva*, 41(4), 277-304. https://doi.org/10.1023/B:MINE.0000005155.70870.bd

Gormley, W. T. & Weimer, D. L. (1999). *Organizational report cards*. Cambridge, Massachusetts: Harvard University Press.

Jabnoun, N. (2015). The influence of wealth, transparency, and democracy on the number of top ranked universities. *Quality Assurance in Education*, 23(2), 108-122. https://doi.org/10.1108/QAE-07-2013-0033

Hallinger, P. (2014). Riding the tiger of world university rankings in East Asia: where are we heading? *International Journal of Educational Management*, 28(2), 230-245. https://doi.org/10.1108/IJEM-11-2012-0126

Hazelkorn, E. (2009). Rankings and the battle for world-class excellence: Institutional strategies and policy choices. *Higher Education Management and Policy*, 21(1), 1-22. https://doi.org/10.21427/D76P7M

Heffernan, T. A. & Heffernan, A. (2018). Language games: University responses to ranking metrics. *Higher Education Quarterly*, 72(1), 29-39. https://doi.org/10.1111/hequ.12139

Huang, M. H. (2012). Opening the black box of QS World University Rankings. *Research Evaluation*, 21(1), 71-78. https://doi.org/10.1093/reseval/rvr003

Huang, M. H. (2011). A comparison of three major academic rankings for world universities: From a research evaluation perspective. *Journal of Library and Information Studies*, 9(1), 1-25. https://core.ac.uk/download/pdf/26749616.pdf

Lee, C. J., Sugimoto, C. R., Zhang, G. & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, 64(1), 2-17. https://doi.org/10.1002/asi.22784

Levin, R. C. (2010). Top of the class: The rise of Asia's universities. *Foreign Affairs*, 89(3), 63-75. https://www.jstor.org/stable/25680916

Li, X., & Thige, J. M. (2017). Benchmarking Jiangsu University to improve its academic ranking. *Journal of Education and Practice*, 8(29), 178-187. https://files.eric.ed.gov/fulltext/ED579140.pdf

Liu, N. C. (2009). The story of academic ranking of world universities. *International Higher Education*, 54, 2-3. https://doi.org/10.6017/ihe.2009.54.8409

Lukman, R., Krajnc, D. & Glavic, P. (2010). University ranking using research, educational and environmental indicators. *Journal of Cleaner Production*, 18(7), 619-628.

Marconi, G. & Ritzen, J. (2015). Determinants of international university rankings scores. *Applied Economics*, 47(57), 6211-6227. https://doi.org/10.1080/00036846.2015.1068921

Marginson, S. (2014). University rankings and social science. *European Journal of Education*, 49(1), 45-59. https://doi.org/10.1111/ejed.12061

Marginson, S. & van der Wende, M. (2007). To rank or to be ranked: The impact of global rankings in higher education. *Journal of Studies in International Education*, 11(3-4), 306-329. https://doi.org/10.1177/1028315307303544

Millot, B. (2015). International rankings: Universities vs. higher education systems. *International Journal of Educational Development*, 40, 156-165. https://doi.org/10.1016/j.ijedudev.2014.10.004

Mussard M. & James, A. P. (2018). Engineering the global university rankings: Gold standards, limitations and implications. *IEEE Access*, 6, 6765-6776. https://doi.org/10.1109/ACCESS.2017.2789326

Olcay, G. A. & Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting and Social Change*, 123, 153-160. https://doi.org/10.1016/j.techfore.2016.03.029

Pizarro Milian, R. & Rizk, J. (2018). Do university rankings matter? A qualitative exploration of institutional selection at three southern Ontario universities. *Journal of Further and Higher Education*, 42(8), 1143-1155. https://doi.org/10.1080/0309877X.2017.1349889

Quacquarelli Symonds (2016). *QS World University Ranking 2016-17*. [viewed 9 September 2016]. http://www.topuniversities.com/university-rankings/world-university-rankings/2016

Reddy, K. S., Xie, E. & Tang, Q. (2016). Higher education, high-impact research, and world university rankings: A case of India and comparison with China. *Pacific Science Review B: Humanities and Social Sciences*, 2(1), 1-21. https://doi.org/10.1016/j.psrb.2016.09.004

Roth, M. G. & McAndrew, W. P. (2018). To each according to their ability? Academic ranking and salary inequality across public colleges and universities. *Applied Economics Letters*, 25(1), 34-37. https://doi.org/10.1080/13504851.2017.1290783

Saisana, M., d'Hombres, B., Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research Policy*, 40(1), 165-177. https://doi.org/10.1016/j.respol.2010.09.003

Shore, C. (2018). How corrupt are universities? Audit culture, fraud prevention, and the Big Four accountancy firms. *Current Anthropology*, 59(S18), S92-S104. https://doi.org/10.1086/695833

Soh, K. (2016). Nearing world-class: Singapore's two universities in QSWUR 2015/16. *Journal of Higher Education Policy and Management*, 38(1), 78-87. https://doi.org/10.1080/1360080X.2015.1126897

Times Higher Education (THE) (undated). *World university rankings*. https://www.timeshighereducation.com/world-university-rankings

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. https://doi.org/10.1007/s11192-005-0008-6

Van Vught, F. & Ziegele, F. (2013). U-Multirank: A user-driven and multi-dimensional ranking tool in global higher education and research. In P. T. M. Marope, P. J. Wells & E. Hazelkorn (Eds), *Rankings and accountability in higher education: Uses and misuses*, pp.257-280. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000220789

Vidal, P. & Filliatreau, G. (2014). Graphical comparison of world university rankings. *Higher Education Evaluation and Development*, 8(1), 1-14. https://pdfs.semanticscholar.org/67a3/90548a4eca7c05f6ae305a51692d54f4ee03.pdf

Waltman, L., Calero-Medina, C., Kosten, J., Noyons, E. C. M., Tijssen, R. J. W., van Eck, N. J., van Leeuwen, T. N., van Raan, A .F. J., Visser, M. S. & Wouters, P. (2012). The Leiden Ranking 2011/2012: Data collection, indicators, and interpretation. *Journal of the American Society for Information Science and Technology*, 63(12), 2419-2432. https://doi.org/10.1002/asi.22708

Webometrics (2016). Ranking web of universities: Methodology. [viewed 7 September 2016] http://www.webometrics.info/en/Methodology

Wu, Y. C. J., Shen, J. P. & Kuo, T. (2015). An overview of management education for sustainability in Asia. *International Journal of Sustainability in Higher Education*, 16(3), 341-353. https://doi.org/10.1108/IJSHE-10-2013-0136

Zitt, M. & Filliatreau, G. (2007). Big is (made) beautiful: Some comments about the Shanghai-ranking of world-class universities. In J. Sadlak & N. C. Liu (Eds), *The world-class university and ranking: Aiming beyond status*, 141-160. https://unesdoc.unesco.org/ark:/48223/pf0000151009

Muhammad Ashraf Fauzi is a senior lecturer at Faculty of Industrial Management, Universiti Malaysia Pahang. He completed his PhD in the Faculty of Management, Multimedia University. His research interests include knowledge management, organisational behaviour, marketing, leadership, entrepreneurship, quality management, and project management.
Email: ashrafauzi@ump.edu.my

Christine Tan Nya-Ling is a senior lecturer of the Information Technology Programmes at the Auckland Institute of Studies (AIS), Auckland, New Zealand. Her current research interests focus on knowledge management, knowledge sharing, service quality, electronic commerce and research methodology.
Email: christinet@ais.ac.nz

Mahyuddin Daud is an Assistant Professor at Ahmad Ibrahim Kulliyyah of Law, International Islamic University Malaysia. He obtained his PhD in Cyber Law (Internet Content Regulation) from International Islamic University Malaysia (IIUM).
Email: mahyuddin@iium.edu.my

Muhammad Mukhtar Noor Awalludin is currently a supply chain manager in Al-Ekhwan Marine Engineering Sdn Bhd. Prior to that, Mukhtar was a full-time PhD researcher at the Faculty of Mechanical Engineering, Universiti Malaysia Pahang in collaboration with Universiti Teknologi Petronas.
Email: mmukhtar1981@gmail.com

Please cite as: Fauzi, M. A., Tan, C. N. L., Daud, M. & Awalludin, M. M. N. (2020). University rankings: A review of methodological flaws. *Issues in Educational Research,* 30(1), 79-96. http://www.iier.org.au/iier30/fauzi.pdf