

What the *Overall* doesn't tell about world university rankings: examples from ARWU, QSWUR, and THEWUR in 2013

Kaycheng Soh*

Independent Consultant, Singapore

In the various world university ranking schemes, the *Overall* is a sum of the weighted indicator scores. As the indicators are of a different nature from each other, *Overall* conceals important differences. Factor analysis of the data from three prominent ranking schemes reveals that there are two factors in each of the ranking systems. For Academic Rankings of World University, the two factors are both academic in nature with a time element and are *negatively* correlated. For the QS World University Ranking (QSWUR) and Times Higher Education World University Rankings (THEWUR), there are two factors each, one for academic excellence and one for managerial excellence. The two factors in QSWUR have a low negative correlation, but the two factors in THEWUR are not correlated. These findings do not support the assumption that the indicators are mutually supporting and additive as conceptualised by the ranking providers. To avoid misinterpretation of the *Overall*, it is suggested that world university ranking needs be re-conceptualised to have a higher-order *institutional* excellence, comprising two lower-order measures based on *academic* excellence and *managerial excellence*. Possible future development is discussed.

Keywords: academic excellence; ARWU; institutional excellence; managerial excellence; QSWUR; THEWUR; world university rankings

World university ranking is an annual affair which is eagerly awaited by a range of consumers, especially vice-chancellors, presidents, rectors, administrators, and perhaps academics of the participating universities the world over. Usually, the consumers of university rankings care almost exclusively about their own institutions' respective rankings in comparison with others' and the changes in rankings from one year to the next. This is evident in the many commentaries posted on the Internet. These consumers care about the ranking but not the *Overall* based on which the universities are ranked. The tacit assumption in so doing is that the ranking tells all about the academic excellence of the university. This simplistic approach to university ranking is understandable since ranking is based on *Overall* which in turn is based on indicator scores supposedly quantifying the various aspects of the university's academic and non-academic or managerial qualities. So it appears that ranking summarises what has taken place in the university and is therefore a trustworthy indication.

In 2003, the Academic Rankings of World University (ARWU) (Shanghai Ranking Consultancy [SRC], 2013) was released for the first time and that set in motion the cumulative frenzy of university ranking. Conducted at the Shanghai Jiaotong University, the original purpose was to find out what makes great universities, but this seems to have since been forgotten. The QS World University Ranking (QSWUR) (Quacquarelli Symonds Ltd. [QSL], 1994–2014a, 1994–2014b) appeared in 2005, and

*Email: sohkaycheng@hotmail.com

it was said to have the purpose of bridging between the academic and industry/business communities in the United Kingdom. In the beginning, QS Limited and the *Times Higher Education Supplement* collaborated and jointly published the ranking results. A few years later, they announced their parting; QSWUR continued and the new (modified) Times Higher Education World University Rankings (THEWUR) (*Times Higher Education*, 2013) appeared. There are others in the ranking arena which propose to provide alternatives to these three systems with new perspectives or contexts. In fact, Rauhvargers (2011) included no less than 11 systems in his comprehensive review commissioned by the European University Association. A more recent system is that of the Centre for World University Ranking (Centre for World University Ranking [CWUR], 2013) and the latest is the U-MultiRank (Osborn, 2013), which is supported by the European Commission with an expectation of 500 participating universities and with a methodology based on a wider range of performance indicators. While the impacts of the new systems are being observed, to date, ARWU, QSWUR, and THEWUR remain the three most popular world university ranking systems, partly because they were in the game earlier.

Three world university ranking systems

Table 1 shows the indicators used in ARWU, QSWUR, and THEWUR classified to indicate those of an academic nature and those of a non-academic or managerial nature. ARWU uses five indicators and one which is derived from the five. All five indicators are academic in nature, pertaining to a university's past and recent academic achievements in terms of alumni and staff winning prestigious prizes and medals, publication in highly

Table 1. Indicators and assigned weights.

Indicator	Academic	Managerial
<i>ARWU</i>		
Alumni winning Nobel Prizes and Fields medals	10	—
Staff winning Nobel Prizes and Fields medals	20	—
Highly cited researchers	20	—
Articles published in <i>Nature</i> and <i>Science</i>	20	—
Science Citation Index and Social Sciences Citation Index	20	—
Per capita academic performance on the above	10	—
Total	100	—
<i>QSWUR</i>		
Academic peer review	40	—
Employer/recruiter review	—	10
Citation per academic staff member	20	—
Academic staff–student ratio	—	20
International students	—	5
International academic staff	—	5
Total	60	40
<i>THEWUR</i>		
Teaching	30	—
Research	30	—
Citations	30	—
International mix	—	7.5
Industry income	—	2.5
Total	90	10

specialised journals, and inclusion in premium citation indices. Thus, the 100 per cent weights are devoted to academic indicators.

For QSWUR, two of its six indicators are academic (i.e., academic and citation) and are given a total weight of 60 per cent. While academic is based on surveys, citation is based on objective counts. The remaining 40 per cent is distributed among the other four indicators (citation, academic staff–student ratio, international academic staff, and international students) which are essentially non-academic, although they may be indirectly related to academic matters. It is obvious that QSWUR has a broader scope by including non-academic indicators to reflect its conception of the modern-day university where relationships with the business/industry community are synergic and internationalism is valued.

As for THEWUR, three of its five indicators (teaching, research, and citations) are academic in nature and are assigned a total weight of 90 per cent. In fact, Teaching is a composite made up of five sub-indicators (i.e., academic reputation survey 15 per cent; student–staff ratio 4.5 per cent; doctorate to bachelor degrees awarded 2.25 per cent; number of doctorates awarded 6 per cent; and institutional income per academic 2.25 per cent). Likewise, research is made up of three sub-indicators (i.e., academic reputation survey 18 per cent; institutional income/total number of academic 6 per cent; and number of papers per academic 6 per cent). THEWUR also includes indicators which reflect interest in internationalism and community relations. Therefore, international mix and industry income are given a total weight of 10 per cent. When compared, THEWUR obviously places more importance on academic indicators than QSWUR does.

Weight-and-sum

These three systems use the weight-and-sum methodology. In fact most, if not all, systems use this method for calculating *Overall*. The exception to this is U-MultiRank which does not provide just one total score but is multi-dimensional as the name implies. In a sense, this is replacing single-score ranking with profiling (Soh, 2012). Weight-and-sum is a simple methodology: a set of indicators selected to fit the conceptualisation of a system are chosen as an operationalisation of academic excellence and data are gathered for these indicators. The indicator scores are then weighted, summed, and scaled. This yields the *Overall* which is then used to compile the league table in which the participating universities are arranged in descending order.

However, a system may group universities beyond the top 100 group in, say, lots of 50 and later 100 yet another system may group those beyond 200 according to some other methodology. The reason for doing this is not given by the systems' designers, presumably because the differences in *Overall* among these low-ranking universities are too small (i.e., in the first decimal) for ranking to be meaningful. Even if two universities differ in the first whole number of the *Overall*, and are thereby ranked one higher than the other, this does not take into account measurement error which is an inevitable component of social measurement. This is another issue of spurious precision, dealt with in an earlier article (Soh, 2012). Anyway, the weight-and-sum approach gives the impression that university ranking is objective and scientific; and this enhances the ranking results with unconditional trust, if not awe, by the consumers of ranking.

Underlying the weight-and-sum methodology is the first assumption that all the indicators are mutually supporting and that they all contribute, though not necessarily in equal proportion, to the measurement of academic excellence (hence, the need for differential weightings). In other words, the relationships between the indicators are

assumed to be additive. Related to this is the second assumption that the indicators compensate one another such that a weakness in one indicator is made good by a strength in another, for instance, a poor showing for citation (an academic indicator) can be compensated by having more international students (a non-academic indicator). Moreover, it assumes that the actual weightings of the indicator scores in the *Overall* are exactly those assigned by the ranking provider. If this is not actualised, the meaning of *Overall* is different from the original conceptualisation and is misinforming and misleading.

As the last-mentioned issue has been recently dealt with (Soh, 2013a), the present study will focus on the first two assumptions by using data from the 2013 rankings from the ARWU, QSWUR, and THEWUR schemes. For this purpose, exploratory factor analysis is a suitable way to uncover the underlying organisation or inter-relatedness (i.e., factor structure) of a set of indicators.

Factor-analytical studies

Factor analysis is a family of correlation-based statistical procedures of which principal components analysis is commonly used. Operationally, principal components analysis combines variables which are more highly correlated and creates factors (or dimensions, components) which are broader concepts at a higher level above the individual variables. This re-organisation reveals the underlying relationships among the original variables (in the context here, the indicators). This simplifies the complex of many variables into a smaller set of factors, each comprising a subset of the original variables. Thus, revealing the hidden structure makes it easier to understand the inter-relations among the variable for clearer interpretation (Tabachnick & Fidell, 2001).

Based on the correlations, variables are organised into a smaller number of factors with each factor denoting the distinct construct defined by the high-loading variables. When two or more such factors are found, the question of their relationship arises. There are two commonly used approaches to this issue. First is the Varimax rotation which results in orthogonal (uncorrelated) factors that are independent of one another. This is achieved by organising the variables such that the contribution of a variable is maximal on one factor and, at the same time, minimal on the remaining factors in the same analysis; hence, the term Varimax for *variance maximisation*. The independence among the factors makes interpretation easier as each factor can be interpreted with no reference to other factors. This factorial independence may be a true reflection of physical objects (e.g., the three dimensions of height, length, and depth of a cube) but may not truthfully reflect the many dimensions of human minds and feelings which tend to be mutually influencing. In contrast, the Promax rotation allows the resultant factors to correlate. This may be more sensitive to human minds and feelings, but it poses a problem for interpretation in that when interpreting a factor it is necessary to bear in mind the other factors which correlate with it. Of course, when the oblique (correlated) factors show low correlations, the results may be just as good as those of a Varimax rotation. For this reason, Promax rotation was used in the present study.

Before studying the factor structures of the indicators used in ARWU, QSWUR, and THEWUR, it is useful to see what has been reported in the pertinent literature on factor analytical studies on university ranking. However, the author of this paper was able to locate only two such studies investigating the underlying factor structure of ranking, by Steiner (2007) and Saisana and d'Hombres (2008), respectively.

Steiner's (2007) study analyses the data for 178 universities common to ARWU and THEWUR in the 2005 rankings. In addition to the original 12 indicators of the two systems, the Webometrics ranking is included; this is a measure of the extent to which the university uses information resources, structures, and technologies on the World Wide Web. Steiner obtains a first factor of academic performance made up of publications, citations, awards, and reviews. The second factor is formed by international academic staff and international students and is thereby indicative of internationalism. A third factor is associated with academic staff–student ratios. The results further show that academic performance is *not* correlated with either internationalism or the staff–student ratio. The author does not indicate whether Varimax or Promax was used to rotate the obtained factors. As gathered from the way the analysis was reported, it is safe to assume that Varimax was employed.

Saisana and d'Hombres's (2008) study analysed the 2007 data of the top 200 universities in the ARWU and the top 200 universities in the THEWUR. However, the data pertain only to the 88 universities common to the two ranking systems. The authors found four factors which together explained 78 per cent of the total variance. The first factor, explaining 15 per cent of total variance, is contributed by THEWUR's international students and staff, indicating internationalism. The second factor explains 34 per cent of total variance and is contributed by ARWU's prize-winning publications and citations of *both* alumni and academic staff and it is clearly an academic factor having to do with *publish-or-perish*. The third factor, explaining 18 per cent of total variance, is contributed by two THEWUR indicators combined with one ARWU indicator, reflecting reputations assessed by academics and employers. This is the only factor for which the two systems came together. Finally, the fourth factor is contributed by two THEWUR indicators of academic staff–student ratio and citation and it explains 11 per cent of total variance. This factor presumably reflects the quality of instruction and quality of academic staff.

By way of summary, the findings of the two independent studies converge. Their results show that the ARWU and THEWUR have their own emphases and their indicators form separate factors that do not correlate.

Present study

The data used in this analysis were downloaded from the websites of the three ranking systems for 2013. They present the ultimate rankings of the universities based on the total score or *Overall*. For this, the indicator scores are also available on the websites.

The present study looks into the factor structures of ARWU, QSWUR, and THEWUR indicators and explores their implications for the interpretation of *Overall* and ranking. Factor analysis, as explained earlier, takes the correlations among the variables (indicators) as a starting point and organises them into a smaller set of factors (or components) to facilitate understanding of the underlying structure among the variables. However, the researchers have a choice of whether or not to allow the obtained factors to correlate, by using Varimax rotation or Promax rotation, respectively.

In this analysis, Promax rotation was chosen, that is, resultant factors are allowed to correlate as they may have mutual influences to some extent. Of course, had the factors been orthogonal (i.e., uncorrelated), the result is the same as for Varimax rotation. For the first run, factor scores were generated and interpreted. Then, the factor scores were submitted for a second run to obtain second-order factors and factor scores; these

represent a factor at a higher level. This approach was applied to all three systems separately.

ARWU

The analysis returned two factors (Table 2). The first factor (explaining 48 per cent of total variance) is heavily contributed to by alumni and awards and the second (explaining 36 per cent total variance) by highly cited researchers, articles published in *Nature* and *Science*, and citation. Although both factors are academic in nature, there is a time gap, with the first factor pertaining more to achievements in the past and the second more to recent achievements. Note that there is a sizable *negative* correlation between the two factors, indicating that universities which were high on one factor tended to be low on the other. This suggests that universities with past glory may be neglecting their recent development or, alternatively, universities unknown in the past are catching up in terms of reputation.

The second-order factor (explaining 78 per cent of total variance) is contributed equally by both factors of achievement but in the opposite direction, with past achievement having a negative loading and recent achievement a positive loading. As this factor was formed by both the first-order factors, it is interpreted as an institutional factor covering both the past and present achievements. The negative loadings suggest that the university's *Overall* (and hence ranking) for 2013 depends on its recent academic achievement and that past glory has a detracting effect. This suggests that a university with high scores for alumni and awards may be neglecting its recent academic development – an effect of resting on the laurels?

As all ARWU indicators are academic in nature, it is surprising that the two factors (one for the past and the other recent achievements) show a negative correlation. Methodologically, it is of note that indicators which are all supposed to be measures of academic achievement and therefore additive can work against one another because of time gap between them. This also suggests that the inclusion of alumni and awards not only did not contribute to *Overall* in the manner expected but, on the contrary, subtracts from indicators which represent recent academic achievement.

Table 2. Factor structures of ARWU.

	Past achievement	Recent achievement
Alumni winning Nobel Prizes and Fields medals (alumni)	.820	.314
Staff winning Nobel Prizes and Fields medals (awards)	.861	.297
Highly cited researchers	.522	.761
Articles published in <i>Nature</i> and <i>Science</i>	.614	.695
Science Citation Index and Social Sciences Citation Index (citation)	.042	.944
Per capita academic performance on the above (PCP)	.905	.075
Total variance explained	48.07%	35.76%
Inter-factor correlation	–.562	
<i>Second-order factor</i>		
Past achievement	–.884	
Recent achievement	.884	
Total variance explained	78.11	

Note: Factor loadings in bold are for the defining indicators.

Table 3. Factor structures of QSWUR.

	Managerial factor	Academic factor
Academic peer review (academic)	.195	.810
Employer/recruiter review (employer)	.600	.521
Citations per academic staff member (citation)	-.079	.601
Academic staff–student ratio	.858	-.095
International students	.838	-.059
International academics	-.231	.361
Total variance explained	31.64%	24.48%
Inter-factor correlation	-.209	
<i>Second-order factor</i>		
Managerial factor	.778	
Academic factor	-.778	
Total variance explained	60.47%	

Note: Factor loadings in bold are for the defining indicators.

QSWUR

The six indicators used in QSWUR were organised into two factors (Table 3). The first factor is contributed to by three non-academic managerial indicators (i.e., employer, academic staff–student ratio, and international students) and explains 32 per cent of total variance. This is clearly a managerial factor. The second is contributed heavily by two academic indicators (i.e., academic and citation) with a minor loading by international academics and explains 25 per cent of total variance. This is obviously an academic factor. The two factors together explain only 57 per cent of total variance, indicating that the set of indicators are more diverse in nature. Note that there is a low negative correlation between the two factors, indicating that universities with higher scores for one factor tend to have somewhat lower score for the other. Thus, there is an antagonistic relation between emphasising academic excellence and emphasising managerial efficiency.

The second-order factor (explaining 61 per cent total variance), as an institutional factor, is formed by the two sets of first-order factor scores but in opposite direction. This, of course, suggests that one aspect of institutional excellence works against the other to some extent.

THEWUR

The five indicators of THEWUR formed two factors together explaining 68 per cent total variance (Table 4). The first factor is obviously an academic factor loaded heavily by teaching and research, with a lesser loading by citations. The second is clearly a managerial factor contributed by international mix and industry income, two non-academic indicators. The inter-factor correlation is a negligible one, indicating that universities high on one factor may and may not be high on the other; that is, the two factors are independent.

For the second-order institutional factor, both first-order factors loaded onto it equally and in the same direction. This indicates that the two factors can be additive and may be mutually supporting.

With the underlying factors revealed above, a question can be asked is, *How do the two factors predict the original Overall in the three systems?* This question can be answered by running a multiple regression of *Overall* on the factors scores. In so doing, the *Overall* was used as the criterion variable and the factor scores as the

Table 4. Factor structures of THEWUR.

	Academic factor	Managerial factor
Teaching	.958	-.012
Research	.943	.157
Citations	.423	-.694
International mix	.075	.500
Industry income	.199	.783
Total variance explained	40.62%	27.38%
Inter-factor correlation	.036	
<i>Second-order factor</i>		
Academic factor	.720	
Managerial factor	.720	
Total variance explained	51.80%	

Note: Factor loadings in bold are for the defining indicators.

Table 5. Regression of *Overall* on factor scores.

	b-weight	SE	Beta-weight	<i>t</i>	<i>p</i>
ARWU ($R = .9735$; adjusted $R^2 = .9729$)					
Past achievement	14.806	.269	1.099 (50.9%)	54.948	.001
Recent achievement	13.508	.269	1.002 (46.4%)	50.134	.001
QSWUR ($R = .7171$; adjusted $R^2 = .7113$)					
Managerial excellence	3.429	.514	.367 (21.6%)	6.672	.001
Academic excellence	7.893	.514	.844 (49.6%)	15.358	.001
THEWUR ($R = .9825$; adjusted $R^2 = .9821$)					
Managerial excellence	-12.848	.231	-1.072 (-42.2%)	-55.679	.001
Academic excellence	17.085	.231	1.426 (56.1%)	74.044	.001

Note: Percentages in parentheses are variance components.

predicting variables. How well the obtained model fits the data is indicated by the proportion of the criterion variable explained (predicted) by the predicting variables, that is, the coefficients of multiple determination adjusted for the number of variables (namely, the adjusted R^2).

As the results in Table 5 show that for ARWU both past achievement and recent achievement factors (Table 2) predict the *Overall* with a very high regression coefficient predicting 97 per cent of the variance in *Overall*. It is of note that the past achievement factor predicts more than the recent achievement factor does, 51 per cent versus 46 per cent.

For QSWUR, 71 per cent of *Overall* variance is predicted by the managerial and academic factors, with the latter predicting more: 50 per cent versus 22 per cent. And, for THEWUR, 98 per cent of variance in *Overall* variance is predicted by the two factors, but note that while academic factors predict 56 per cent, but managerial factor predicts in the negative direction (-42 per cent), indicating that higher managerial excellence not only *does not* enhance but in fact detracts from the prediction. This is a finding not to be dismissed lightly.

Exercise caution for the present

For the time being, with the currently available rankings, the consumers of world university ranking would be well advised to be cautious in their interpretation and use

of the original *Overall* and the rankings based on it, by taking into due consideration the underlying factor structures of the three systems. The following points based on the findings above are worthy of note:

- For ARWU, there are factors relating to past glory and current academic achievement. These two factors have a sizable negative correlation. The past glory factor plays a more prominent role in the factor structure. This implies that past academic achievement has greater influence on the *Overall* (and ranking) and, because of the negative correlation, the institutional excellence might have been underestimated for the ‘younger’ universities. By the way, there have been efforts to offset the advantage of history by excluding long-established universities from ranking in the *QS Top 50 Under 50* (QS TOPUNIVERSITIES, n.d.) and *THE 100 Under 50* (Times Higher Education, 2014).
- For QSWUR, there are two factors, with the managerial factor playing a more prominent role than the academic factor, and there is a low negative correlation between the two. Thus, universities paying more attention to managerial excellence might have unintentionally neglected their academic development to some degree, and vice versa. A balance between the two aspects of institutional excellence is obviously something to work for.
- For THEWUR, the two factors of academic excellence and managerial excellence are quite independent of each other and they both contribute to institutional excellence, separately. Thus, universities high on one factor may not be so on the other factor. Therefore, there is therefore a need to find a way that the two can come together to contribute jointly to institutional excellence, so that the two subsets of indicators are made mutually supporting and additive.

Re-conceptualising for the future

The findings of the present study corroborate in a general sense the findings of the two factor-analytical studies which used 2005 and 2007 data (Saisana & d’Hombres, 2008; Steiner, 2007), although the findings are not exactly replicated. By way of summary, in the present study, the QSWUR and THEWUR 2013 data show that an academic factor and a non-academic managerial factor underlie the conglomeration of indicators. Moreover, the two factors may not correlate and might even be antagonistic. Even for ARWU which has only academic indicators, there are two factors which are differentiated with a time element. In short, the findings do not support the assumption that the indicators within each system are mutually supporting and additive as conceptualised by the ranking providers. These call for a re-conceptualisation of university ranking.

As the history unfolded, university ranking began with ARWU focusing only on academic excellence operationalised by indicators of an academic nature and assigned 100 per cent weight to them. Then came QSWUR with a broader scope to include non-academic managerial indicators which took up as much as 40 per cent of the weights, leaving 60 per cent to academic indicators. Later, THEWUR came to the ranking arena as a modification of QSWUR and included both academic and non-academic excellence but with 90 per cent weight for academic indicators, supposedly returning to the focus on academic excellence but allowing for an element of managerial excellence.

These changes and development in the ranking of universities are not chance occurrences but have historical precedence. As gathered from the response to the annual ranking results, it seems that universities the world over have now accepted that university ranking is not only a matter of academic excellence but also is about other conditions which are managerial in nature and may contribute to academic excellence indirectly. Therefore, it is herewith suggested that instead of continuing to see the ranking of the world's universities as an evaluation of academic excellence (which was the original purpose), the concept needs be revised and expanded as an evaluation of *institutional excellence* which is made up of two components of *academic excellence* and *managerial excellence* both of which can be evaluated independently. This is a simple formulation represented thus:

$$\text{Institutional excellence} = \text{Academic excellence} + \text{Managerial excellence}$$

With this re-conceptualisation, ARWU as it is covers only academic excellence while QSWUR and THEWUR include both academic excellence and managerial excellence and therefore are an evaluation of institutional excellence. Therefore, the ARWU will need to decide whether it will expand the coverage of its survey to include managerial excellence and thereby evaluate institutional excellence or continue to focus specifically on only academic excellence. As for QSWUR and THEWUR, there is a need to decide on the relative significance between academic excellence and managerial excellence so as to continue evaluating institutional excellence (which they have been doing without invoking the concept and its namesake) and to look into how the two first-order sources of excellence can be balanced. To do so would ensure the avoidance of have a confusing situation where non-equivalents are equated.

In this conceptualisation, institutional excellence is represented by a composite score made up of scores for academic excellence and managerial excellence. The tacit assumption is that the two lower-level forms of excellence are mutually supporting and therefore additive. However, as the analysis has shown, this may or may not be the case and care needs be exercised when interpreting the score for the second-order factor (i.e., institutional excellence). At the moment, the QSWUR data show the two first-order factors to be mildly antagonist, whereas the THEWUR data show them to be uncorrelated. Thus, there is a need for the systems to find ways and means to ensure a moderate degree of positive correlation between academic excellence and managerial excellence so that these can be combined as a measure of institutional excellence and yet yield information pertaining to the two aspects with sufficient independence for making decisions on academic or managerial matters.

Moving ahead

Ranking consumers the world over seem always to accept the ranking results of the three currently most popular systems as if the ranking results are perfectly trustworthy without any reservation. However, world university ranking has been criticised on its conceptualisation and usefulness so much so that Lincoln (2012) is of the view that world university ranking's heyday has passed.

Since 2000, criticism has been directed at world university ranking systems on several grounds. For instance, Bowden (2000), after reviewing a number of university league tables appearing in the United Kingdom in 1998, concludes that the league tables did not provide the types of useful information potential students need to make

informed decisions about choosing where to go for study. More technically, Usher and Savino (2006) consider it arbitrary to select and weight indicators for the calculation of a total score used for ranking universities against one another. They concluded that there may be other legitimate indicators or combinations of indicators that are neglected. Stella and Woodhouse (2006, p. 23) are more concerned with the problem of interpretation of university ranking results and conclude that ‘...one main problem with rankings is that they are measuring one thing, but most of their reader think they are measuring something else’. Criticism of university ranking continues (Hazelkorn, 2007, 2009; Rauhvargers, 2011; Staff Writers, 2012), and Harvey (2008) summarily dismissed the trustworthiness and usefulness of ranking on methodological ground, thus:

The construction of indices by which institutions or departments are ranked is *arbitrary, inconsistent and based on convenience measures*. The operationalisation of the concept of quality is cursory at best. (p. 189; emphasis added)

Most of the criticisms cited above are general discussion focusing on the conceptualisation and utility of ranking results. Specific problems of a measurement and statistical nature have only been uncovered recently (Soh, 2013b). Such problems pertain to spurious precision (i.e., assigning relative ranks to universities with only small differences in *Overall* different rankings), the assumption of mutual compensation of indicators (e.g., high international student proportion ‘makes good’ a low citation rate), discrepancy between assigned and obtained indicator weights (i.e., an indicator assigned a weight which turns out to be much lower/higher), and so on. A more recent criticism concerns the multicollinearity or overlapping of indicators which render some indicators redundant (Soh, 2014).

World university ranking has been around for more than a decade and so have criticisms of it. However, there is no evidence that ranking providers have heeded the criticisms or taken steps to improve their systems. A comparison of the systems as they were when first conceived and as they now operate would show that they are practically the same. In short, ranking has been static and self-perpetuating while the world moved ahead. The rank providers are either oblivious to the methodological/statistical issues raised or, for reasons better known to themselves, prefer to hide their heads in the sand. Until such time as the flaws are rectified, world university ranking results will continue to misinform.

Ranking providers and consumers have been preoccupied with rankings one year to the next with the competitive drama played out in the same manner with the same excitement. Having several systems around and with more coming, they have neither the time nor the desire to pay sufficient attention to the conceptual and technical problems underlying the ranking of the world’s universities, or the need to improve and refine them for the benefit of all interested parties. This is rather like the Caucus Race in the *Adventures of Alice in Wonderland* (Carroll, 1865) which keeps everybody busy running in circles leading to nowhere.

Even if the type of information provided by the various ranking systems is needed by universities, there is always the need for caution and discernment so as to avoid misinterpretation when evaluating their institutions’ strengths and weaknesses so that efforts and resources are devoted to the correctly identified aspects of a university’s endeavours. At the same time, rank providers need to take heed of the various criticisms and thereafter render a better service to their clientele by regularly reviewing

and modifying their operations to improve and refine their products – the rankings of the world's universities.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Bowden, R. (2000). Fantasy higher education: University and college league tables. *Quality in Higher Education*, 6(1), 41–60. doi:10.1080/13538320050001063
- Carroll, L. (1865). *Adventures of Alice in Wonderland*. London: Macmillan.
- Centre for World University Ranking. (2013). *CWUR: World University Ranking 2013*. Retrieved from <http://cwur.org/2013/>
- Harvey, L. (2008). Editorial: Rankings of higher education institutions: A critical review. *Quality in Higher Education*, 14(3), 187–207. doi:10.1080/13538320802507711
- Hazelkorn, E. (2007, November 11). OECD: Consumer concept becomes a policy instrument. *University World News*, (5). Retrieved from <http://www.universityworldnews.com/article.php?story=20071108144018483&query=university+ranking>
- Hazelkorn, E. (2009). *The impact of global rankings on higher education research and the production of knowledge*. Dublin: Centre for Social and Educational Research, Dublin Institute of Technology.
- Lincoln, D. (2012). Rankings: An idea whose time has come, and gone. *Inside Higher Ed*. Retrieved from <https://www.insidehighered.com/blogs/world-view/rankings-idea-whose-time-has-come-and-gone>
- Osborn, A. (2013, February 1). U-Multirank is launched, 500 universities expected to sign up. *World University News*, (257). Retrieved from <http://www.universityworldnews.com/article.php?story=20130201144343699>
- Quacquarelli Symonds Ltd. (1994–2014a). *QS Top 50 Under 50*. Retrieved from [http://www.topuniversities.com/top-50-under-50/2013#sorting=rank+region="+country="+faculty="+stars=false+search=](http://www.topuniversities.com/top-50-under-50/2013#sorting=rank+region=)
- Quacquarelli Symonds Ltd. (1994–2014b). *QS World University Rankings 2013*. Retrieved from [http://www.topuniversities.com/university-rankings/world-university-rankings/2013#sorting=rank+region="+country="+faculty="+stars=false+search=](http://www.topuniversities.com/university-rankings/world-university-rankings/2013#sorting=rank+region=)
- QS TOPUNIVERSITIES. (n.d.). *QS University Rankings: Top 50 under 50 2013*. Retrieved from [http://www.topuniversities.com/top-50-under-50/2013#sorting=rank+region="+country="+faculty="+stars=false+search=](http://www.topuniversities.com/top-50-under-50/2013#sorting=rank+region=)
- Rauhvargers, A. (2011). *Global university ranking and their impact. EUA report on rankings 2011*. Brussels: European University Association.
- Saisana, M., & d'Hombres, B. (2008). *Higher education rankings: Robustness issues and critical assessment: How much confidence can we have in higher education rankings?*. Luxembourg: European Commission Joint Research Centre, Institute for the Protection and Security of the Citizen, and Centre for Research on Lifelong Learning.
- Shanghai Ranking Consultancy. (2013). *Academic ranking of world universities 2013*. Retrieved from <http://www.shanghairanking.com/ARWU2013.html>
- Soh, K. C. (2012). Profiling universities, not only ranking them: maximising the information of predictors. *Higher Education Review*, 44, 27–42.
- Soh, K. (2013a). Rectifying an honest error in world university rankings: A solution to the problem of indicator weight discrepancies. *Journal of Higher Education Policy and Management*, 35, 574–585. doi:10.1080/1360080X.2013.844670
- Soh, K. (2013b). *Social and educational rankings: Problems and prospect*. New York, NY: Untested Ideas Research Center.
- Soh, K. (2014). Multicollinearity and indicator redundancy problems in world university rankings: An example using THEWUR 2013–2014 data. *Higher Education Quarterly*, 69, 158–174.
- Staff Writers. (2012, June 5). Twelve questionable factors behind traditional university ranking. *Online Colleges*. Retrieved from <http://www.onlinecolleges.net/2012/06/05/12-questionable-factors-behind-traditional-university-rankings/>

- Steiner, J. E. (2007). *World university rankings: A principal component analysis*. Sao Paulo: Instituto de Estudos Avancados da Universidade de Sao Paulo. Retrieved from <http://arxiv.org/ftp/physics/papers/0605/0605252.pdf>
- Stella, A., & Woodhouse, D. (2006). *Ranking of higher education institutions* (Occasional Publications Series No.6). Melbourne: AUQA. Retrieved from http://www.auqa.edu.au/files/publications/ranking_of_higher_education_institutions_final
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.
- Times Higher Education. (2013). The world university rankings 2013–2014. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/world-university-rankings/2013-14/world-ranking>
- TSL Education Ltd. (2014). The 100 under 50 university rankings: Results. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/419908.article>
- Usher, A., & Savino, M. (2006). *A world of difference: A global survey of university league tables* (Canadian Education Report Series). Toronto: Educational Policy Institute. Retrieved from <http://www.educationalpolicy.org/pdf/World-of-Difference-200602162.pdf>

Copyright of Journal of Higher Education Policy & Management is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.