High Educ (2010) 60:269–286
DOI 10.1007/s10734-009-9299-5

Obtaining a picture of undergraduate education quality: a voice from inside the university

Chia-Wei Tang · Cheng-Ta Wu

Published online: 27 December 2009

Abstract This study aims to construct ranking indicators from the perspective inside of the university and shift the ranking target from overall university quality to undergraduate education quality. In dealing with the complexity of the concept of undergraduate education quality, two-stage questionnaire survey was conducted to gain comprehensive opinions from 20 higher education evaluation experts. The Fuzzy Delphi Method then was introduced to perform data analysis and help final indicator selection as well as the distribution of weights. The results compared with ranking systems conducted by US NEWS and Guardian showed weighting differences and greater comprehensiveness in terms of the types of measures, sources of data and different perspectives from the students, employers and the academics. Most important of all, this study provided a more transparent ranking system and detailed ranking methodology that are crucial for users' understanding and use of the ranking system.

Keywords University ranking · Quality of undergraduate education · Fuzzy Delphi Method (FDM)

Introduction

While higher education institutions cannot reach consensus on whether rankings are desirable or not, there is increasing acknowledgement that the practice is here to stay (Merisotis 2002; Merisotis and Sadlak 2005; Usher and Savino 2006). Rankings, in one way or another, satisfy a public demand for comparable information about higher education while institutions and governments pay less attention to it (Federkeil 2002). To this end, rankings hold universities accountable (Hazelkorn 2006; van Dyke 2005), especially

C.-W. Tang (✉)
Graduate School of Education, University of Bristol, 35 Berkeley Square, Clifton, Bristol BS8 1JA, UK
e-mail: edcwt@bristol.ac.uk

C.-T. Wu
College of Education, National Chengchi University, Taipei, Taiwan

to their clients such as students and parents. However, scholars are frequently critical of the methodology of current national ranking systems in the UK and the United States. Dill and Soo (2004) and Van Dyke (2005) claim that the choice of indicators and weightings are subjective and biased, with little or no empirical basis. Additionally, these rankings do not measure what the publishers think they are measuring (Stella and Woodhouse 2006a). Instead, they reflect only the views of different publishers about what is good university (Marginson 2007).

This study aims to respond actively to discourse surrounding university ranking systems. Its purpose is to explore the issues regarding the measurement of undergraduate education quality from the perspective of higher education evaluation experts and presidents of top universities in Taiwan. As a preliminary study, this article will focus on indicator construction and weightings but not the other aspects of ranking systems. A two-phase questionnaire survey and the Fuzzy Delphi Method (FDM) were employed to gain comprehensive opinion from higher education experts in order to construct a ranking system for Taiwan. Its implication for programs as well as institutional ranking in other countries is also discussed.

Literature review

University rankings (also called league tables, and report cards) are weighted combinations of performance indicator scores to rank institutions (Bowden 2000; Morrison et al. 1995). According to publishers, the commercial rankings measure "university quality" which is a multi-dimensional concept that embraces key functions and activities, such as teaching, research, staffing, students, facilities, and services to the community (Harvey 2002; Salih 2003; Stephenson 2004; UNESCO 2006). The concept of university quality is so broad that ranking systems differ extensively in the type of indicators selected and the particular definition as well as measures of university quality adopted by the publishers vary (Buela-Casal et al. 2007). To date there is limited discussion or explanation of what commercial ranking systems measure provided by the ranking publishers. Except for the *Washington Monthly* ranking which has a clear focus regarding how universities contribute to society, most publishers claim that their rankings are "good university guides" that help students identify the best university. Nevertheless the concept of "the best university" is questionable and not well justified by the publishers.

Empirical investigations of the introduction and comparison of current ranking systems are global in scope (Buela-Casal et al. 2007; Clarke 2005; Dill and Soo 2004, 2005; Usher and Savino 2006; Van Dyke 2005). Many studies examined the effectiveness of rankings (Altbach 2006; Billing 2003; Bowden 2000; Brooks 2005; Butler-Adam 2007; CHE Report 2000; Clarke 2002, 2004; Dill and Soo 2004; Eccles 2002; Federkeil 2002; Morrison et al. 1995; Stella and Woodhouse 2006a). A common criticism of scholars is that most of the current rankings lack transparency and a complete rationale to help readers understand why particular measures are included and why they are important as well as necessary to understand university quality. For example, according to Dichev (2001), approximately 70–80% of the variance in the US NEWs' rankings from 1 year to the next is due to changes in the ranking formula and weight assigned to categories. With little or no theoretical or empirical basis, assigning weights to different indicators is regarded as being biased and invalid (Bowden 2000; Carrico et al. 1997; Clarke 2002; Dill and Soo 2005; Eccles 2002).

An alternative ranking target: quality of undergraduate education

Rankings and league tables play a useful role in many ways, but the challenge is to ensure that they provide accurate and relevant assessments that measure the right elements and proper weighting (Altbach 2006). The concept of "university quality" might be too broad to have a clear focus if rankings are designed to help students' choose a university. University quality is a too broad concept to evaluate through use of only a few measures, since a university's mission is no longer limited to teaching. To use this term for ranking could be misleading for readers and even the ranking publishers themselves. According to Dill and Soo (2005), rankings provide comparative information about "college education" which is an increasingly important and expensive decision for students and their families. This definition can be further explained by Altbach's (2006) argument that rankings ignore key academic roles such as teaching and do not consider how students are affected by their academic experiences. In other words, instead of employing the concept of "university quality", "the quality of undergraduate education" becomes a more concise and appropriate target in the rankings complied for students.

It is important to note, however, that an actual consensus regarding what education quality means remains questionable in the higher education domain (Mullin and Wilson 2000). In his book, 'The Uses of the University', former University of California President, Clark Kerr, laid out his views on the concept of "multiversity" and its application to the practices of higher education during the mid-twentieth century. He argued that the university is no longer an ivory tower, but a service station responsive to multiple social forces and the varied needs outside the university (Miyoshi 2000). The definition of undergraduate education quality here varies due to the perspectives of different stakeholders thus making it difficult or even impossible to propose and apply a single universal definition (Harvey 2002; Harvey and Green 1993; Lagrosen et al. 2004; Srikanthan and Dalrymple 2003).

Mapping undergraduate education quality: from whose perspective?

Instead of continuing to pursue a single definition of quality, a more adaptable and broadly adopted approach in the quality assurance literature is to identify basic elements of education quality. Green (1994) argues that, given the difficulties in defining quality in higher education it is necessary to consider views from relevant stakeholders. That is to say, a ranking system should reflect different angles in order to offer a comprehensive view of the complexity of modern higher education. Lewis' (2006) reflection on Harvard University's shift to an overemphasis on the customer and the market's needs is an important reminder of the dilemmas associated with universities' roles as knowledge and skill factories in the modern age:

> Harvard strives to be the best at many things and it often succeeds. But Harvard has protected its reputation for excellence at the expense of its sense of a larger purpose. Harvard's leaders have allowed the university's mission to drift from education to customer satisfaction. For them, Harvard is no longer a city upon a hill but merely a brand name (Lewis 2006, p. 301).

Lewis (2006) and Altbach (2006) argue that since university rankings provide publisher viewpoint regarding quality and do no way take into account the tradition values of education, universities that are trying to rise in the rankings have an incentive to pursue dollars, test scores and faculty research credentials, rather than other areas which

actually matter more to undergraduate education, for example the quality of course and teaching. According to Stella and Woodhouse (2006b), universities do tend to focus on the indicators and weights used in rankings and this might lead to institutions attending to certain aspects of institutional quality over other vital aspects, most notably the ones that would ultimately bring them a higher ranking. The discussion above suggests that university ranking is far more than providing comparative information for students and parents; it is an issue of great complexity and has tremendous influence on university policy and development. Due to the lack of professional knowledge and thorough understanding of higher education, current ranking publishers piece together a picture of quality based on their opinions or on available data. This distorts what is education quality as well as rankings and some input measures and reputation can play dominant roles while the core issues of education quality receives little attention (Altbach 2006). Accordingly, this study argues that since the rankings are here to stay, academics should respond to these rankings in a more active way in order to improve the tilted emphasis of current rankings. Students' needs can no longer being ignored, but profession judgment as well as broader societal concerns should also play an important role to inform a balanced way of thinking about the nature and idea of today's undergraduate education quality. Moreover, this study assumed that instead of university quality, using undergraduate education quality as ranking target would make the selecting and weighting of indicators more focused and thus more accurate.

Research design

Research setting

The current higher education system in Taiwan includes 2-year junior colleges, technical colleges, liberal arts colleges, 4-year universities and graduate schools. The ranking system explored in this study mainly focused on that of the public 4-year universities. Since the 1990s, higher education in Taiwan has experienced tremendous expansion which has led to fierce competition for resources among institutions. The influence of marketization has never been greater. Moreover, with Taiwan's entry into the WTO, the increasing pressure to compete internationally constantly challenges higher education institutions. Recently this challenge has been fueled by the 'Development of Outstanding Universities and Research Centres' project launched in 2005 by the Ministry of Education. Similar to many other countries, Taiwan's higher education faces an unparalleled challenge to redefine its role among tensions of massifacation, marketization, and globalization.

Methodology

The aim of this study was shaped by an insider's perspective of higher education. This focus shifted the current discourse on rankings from the "overall university quality" to the "undergraduate education quality". In other words, the potential to provide good quality education specifically at the undergraduate level was of concern.

Rankings involve subjective value judgment and great debate; therefore, this study employed the FDM. The FDM is an effective tool to gather data generated from opinion that usually involves imprecision and uncertainty. It is a methodology by which subjective data can be transformed into quasi-objective quantitative data and to facilitate decision-making

of controversial issues. The FDM is a combination of the traditional Delphi method and the fuzzy set theory. The fuzzy set theory was introduced by Lotfi Zadeh in 1965 and served as an extension of the classical set theory in which the membership of elements in a set is assessed in binary terms (Yes or No). The fuzzy set theory permits the gradual assessment of the membership of elements in a set; this is described with the aid of a membership function valued in the unit interval [0, 1] (Ragin 2007). According to Chang et al. (2000), the FDM provides the following merits: (a) It processes the fuzziness in relation to the forecast item and the information contents of respondents and (b) Individual attributes of participants may be elucidated because the fuzzy forecasts are utilized and thus preserved. In brief, this method tries to reach a quantitative consensus in a way taking into account of the variance of respondents'. It is more than just a mean generated mechanically from given single values but an *agreeable* score based on the "acceptable value ranges" (as Table 2 shows) given by respondents.

Procedures

This study employed two rounds of questionnaire survey to gain opinion from higher education experts. The participants were presidents from the top 12 universities in the 'Development of Outstanding Universities and Research Centers' project and 10 higher education evaluation experts in Taiwan who specialize in the field of higher education evaluation. The rationale for selecting presidents and evaluation experts are based on the assumption that they are experts who understand the quality of undergraduate education well and familiar with how it can be better evaluated. President candidates in Taiwan are required to (1) have been (were) tenure professors in the university more than 4 years; (2) have administration experiences (being superintendent at departmental and institutional level) in universities more than 4 years. In other words, presidents in Taiwan are not only familiar with administration in the university but also teaching, curriculum and how these are or should be evaluated. The 10 higher education evaluation experts are selected due to their great involvement (as panel members) in the national university evaluation practices initiated by the Taiwan MoE.

First round questionnaire: the construction of ranking determinants and measures

Present research has compiled 16 university ranking systems worldwide (Table 1) in order to develop a preliminary ranking model. The number of individual measures used in these ranking systems runs well into the hundreds. In Fig. 1, they have been categorized as 11 larger quality determinants according to the aspect of university quality being measured.

The first questionnaire was developed based on the preliminary model. Annexes stating how those determinants and indexes might contribute to the understanding of education quality were also attached to the questionnaire. The purpose of Annexes was to help respondent understand the determinants and indexes better in order to assist them in making more relevant and concise comment on the preliminary ranking model. The main purpose of this questionnaire was to gather higher education experts' opinion on whether the main determinants and measures in question were appropriate for the evaluation of "undergraduate education quality" in the Taiwanese context and moreover, whether or not the classification of measures into the given categories was a suitable one. More importantly, the questionnaire respondents were requested to provide suggestions for additional measures or determinants that could improve the validity of the ranking system and to provide reasons to justify their ideas.

Table 1 Schematic representation of existing ranking systems classified by purpose

	Evaluation of research quality	Undergraduate choice	Evaluation of the contribution to the society	Evaluation of international university quality
National universities/ programs	The Center RAE New Zealand PBRF	US News Maclean's Times Guardian GUG(UK) GUG(AU) ISAU CHE-Der Stern Education18	Washington Monthly	
International universities/ programs				ARWU The Times HES Asiaweek

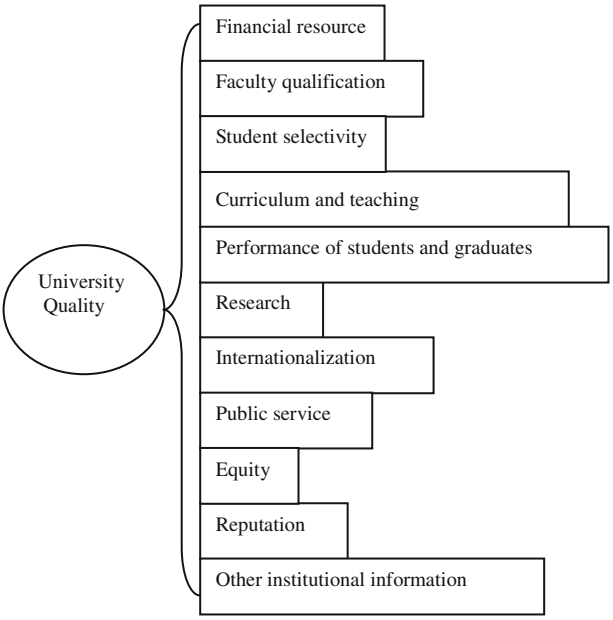
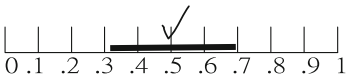


Fig. 1 Composite view of university quality determinants of current ranking systems

Second round questionnaire: Fuzzy Delphi Method questionnaire design

The Fuzzy Delphi questionnaire was designed to gather participant opinion on the second ranking model and was based on the first questionnaire. Respondents were required to grade the relative importance of main categories and measures to determine indicator selection and weight distribution. In the Fuzzy Delphi questionnaire, experts were asked to give a three-point estimate with the left-hand value representing the most pessimistic value, the right-hand value representing the most optimistic value and a mark representing the most suitable value in the interval between the left and right value (Table 2).

Table 2 A sample of FDM questionnaire

Determinants/measures	Compared with other measures, to what extent does this measure help understand of the quality of the determinant
Determinant 1. Resources	
Measure 1.1 revenue per student	

Source: The author

Data analysis

With the reply from 10 of the 12 presidents and 10 of 13 evaluation experts, the response rate for both questionnaires was 80%. The FDM consisted of two main steps. First, respondents' opinions were transformed to triangular fuzzy numbers to represent their views on the importance of each factor and measure. After that, a process of defuzzification was employed to synthesize respondents' opinion into a single value representative of the consensus of all respondents.

Step 1:

Based on participant responses, the triangular fuzzy number for each measure was identified as: $A = (L, M, U)_{L-R}$, where $L \leq M \leq U$, L and U stand for the lower and upper value and M for the modal value. The membership function of the triangular fuzzy number was defined as follows:

$$\mu_{\underset{\sim}{A}}(X) = \begin{cases} (x-l)/(m-l), & l \leq x \leq m \\ (x-u)/(m-u), & m \leq x \leq u \\ 0 \quad \text{otherwise} \end{cases}.$$

Step 2: Defuzzification

This study adopted Chen and Huang's (1992) method to process defuzzification and transformed the fuzzy number into a definite value to represent the respondent opinions. The computational procedure was summarized as follows:

1. Define fuzzy max and fuzzy min as:

$$\mu_{\max}(X) = \begin{cases} x, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\mu_{\min}(X) = \begin{cases} 1-x, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

2. Evaluate the left utility score $\mu_L(A)$:

$$\mu_L(A) = \sup_x [\mu_A(X) \wedge \mu_{\text{MIN}}(X)]$$

3. Evaluate the right utility score $\mu_R(A)$.

$$\mu_R(A) = \sup_x [\mu_A(X) \wedge \mu_{\mathrm{MAX}}(X)]$$

4. Given the left and right scores, the total score of a fuzzy number is defined as:

$$\mu_T(A) = [\mu_R(A) + 1 - \mu_L(A)]/2$$

5. Higher total scores means indicate a higher level of factor importance. This study choose α-cuts $= 0.5$ as the threshold to select the ranking system measures. The value of 0.5 was the middle point (median) of interval [0, 1]. It reflected the logical reasoning that only those elements from the support of a fuzz set with "sufficiently large" membership grades in a fuzz set were included (Bodjanova 2006).
6. To determine the weights of each factor and measure, total scores of each main factors and measures were normalized. Weights were distributed in a total amount of 100.

Results

The main purpose of the first questionnaire was to gather higher education experts' comments on the preliminary model and identify other potential determinants or measures that help predict the quality of undergraduate education in the Taiwan context but not included in the preliminary model. The determinants identified from the literature, such as "public service" and "other institutional information" were generally regarded invalid or irrelevant for ranking in the present study. Figure 2 is the conceptual framework of the undergraduate education ranking system based on respondents' opinions.

Among measures included in the preliminary model, 36 measures were considered invalid for the higher education context in Taiwan or for the understanding in their corresponding quality determinants. These measures were either disregarded or modified according to respondents' suggestions. Moreover, respondents recommended six additional measures; finally, 35 measures in total were selected to develop the Fuzzy Delphi questionnaire.

Tables 3 and 4 summarize the Fuzzy Delphi questionnaire results. In Table 3, the total score represented respondents' opinion regarding the relative importance of each factor for "understanding the quality of undergraduate education". In Table 4, the total score

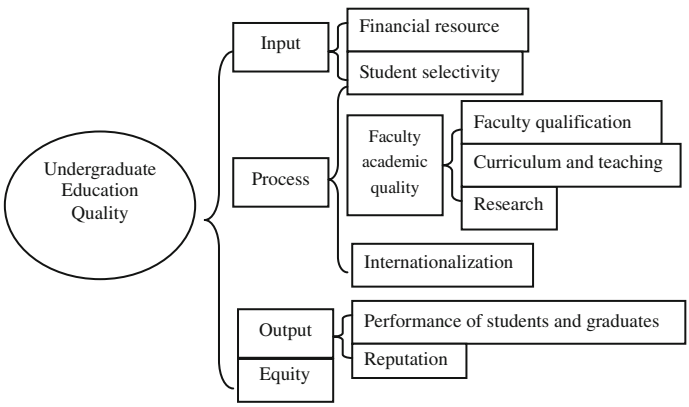


Fig. 2 Higher education experts' view of undergraduate education quality determinants

Table 3 Results of Fuzzy Delphi questionnaire (determinants)

Determinants	Triangular fuzzy number	Left utility score	Right utility score	Total score
1. Financial resource	(0.417 0.730 0.955)	0.444	0.780	0.668
2. Student selectivity	(0.682 0.796 0.917)	0.285	0.818	0.766
3. Faculty qualification	(0.750 0.806 0.955)	0.237	0.831	0.797
4. Curriculum and teaching	(0.830 0.869 0.955)	0.164	0.879	0.858
5. Research	(0.682 0.813 0.955)	0.276	0.851	0.787
6. Internationalization	(0.500 0.664 0.917)	0.430	0.732	0.651
7. Performance of students and graduates	(0.667 0.796 0.955)	0.295	0.824	0.765
8. Reputation	(0.682 0.806 0.955)	0.283	0.831	0.774
9. Equity	(0.409 0.574 0.917)	0.507	0.683	0.588

indicated the relative importance of each measure for understanding the performance of associated factors. For instance, the total score of "curriculum and teaching" was the highest; respondents regarded it as the most important factor to predict the quality of undergraduate education.

To delete measures which were less relevant/important, this study choose α-cuts $= 0.5$ as the threshold to select the measures for the ranking system. Measures including "graduation rates", "drop-out rate", "computer and internet spending" and "scholarship spending" scored 0.487, 0.494, 0.476 and 0.497 were crossed out. To determine the weights of each determinant and measure, total scores were normalized and weights were distributed by a total percentage of 100. The final ranking system is presented in Table 5.

Discussion

It is argued that due to the lack of professional knowledge and thorough understanding of higher education displayed by ranking publishers, the measure selection and weight distribution in commercial rankings fail to have their focus on what actually matters for student's education and learning (Bowden 2000; Carrico et al. 1997; Clarke 2002; Dill and Soo 2005; Eccles 2002). In response, this study narrowed down the evaluation target on the quality of undergraduate education in order to gain a clearer focus. Figure 2 is the conceptual framework of the undergraduate education ranking system based on respondents' opinions. The determinants identified in the literature and applied to the preliminary model, such as "public service" and "other institutional information" were generally regarded irrelevant for ranking in the present study. It was noted that other institutional information, such as location and costs of study, could be informative for students making their university choices but they did not directly contribute to undergraduate education quality. Some respondents also commented that instead of putting this information into the ranking system in a rather brief manner, universities were responsible for providing this information on web pages in a detailed, user-friendly format.

One fundamental assumption of this study is that the inclusion of professionals inside universities and the use of quality of undergraduate education as ranking target would provide another landscape concerning the indicator selection and weighting for university ranking. In the light of this assumption, Table 6 presents a comparison/overview of the

Table 4 Results of Fuzzy Delphi questionnaire (measures)

Measure	Triangular fuzzy number	Left utility score	Right utility score	Total score
1.1 Revenue per student	(0.519 0.689 0.917)	0.411	0.747	0.668
1.2 Computer and internet spending	(0.283 0.500 0.591)	0.589	0.542	0.476
1.3 Library spent per student	(0.591 0.765 0.917)	0.348	0.796	0.724
1.4 Scholarship spending	(0.365 0.417 0.864)	0.604	0.597	0.497
2.1 Acceptance rate	(0.283 0.671 0.917)	0.517	0.736	0.610
2.2 Entry score	(0.750 0.830 0.955)	0.231	0.849	0.809
3.1 Percent of full-time faculty with top terminal degree	(0.682 0.825 0.955)	0.278	0.845	0.783
3.2 Percent of full-time faculty as professor	(0.667 0.760 0.917)	0.305	0.793	0.744
3.3 Percent of full-time faculty	(0.583 0.784 0.917)	0.347	0.809	0.731
4.1 Staff-student ratio	(0.394 0.597 0.830)	0.504	0.673	0.585
4.2 Senior student assessment score	(0.394 0.774 0.864)	0.439	0.793	0.677
4.3 National teaching assessment score (by academic peers)	(0.691 0.805 0.955)	0.277	0.830	0.777
4.4 Graduation rates	(0.383 0.436 0.720)	0.586	0.561	0.487
4.5 Drop-out rate	(0.294 0.507 0.654)	0.582	0.570	0.494
5.1 Research grants per academic staff member	(0.591 0.788 0.864)	0.342	0.803	0.731
5.2 Percent of academic staff member with National Faculty Awards	(0.500 0.768 0.917)	0.394	0.798	0.702
5.3 Percent of academic staff member with Academy membership	(0.500 0.764 0.917)	0.396	0.795	0.700
5.4 Publications on Nature, Science, SCI, SSCI, TSSCI[a], EI and A&HCI per academic staff member	(0.417 0.738 0.917)	0.441	0.778	0.668
5.5 Citations per article on Nature, Science, SCI, SSCI, TSSCI, EI and A&HCI	(0.500 0.788 0.917)	0.388	0.812	0.712
5.6 Articles in peer-reviewed journals per academic staff member	(0.500 0.764 0.864)	0.396	0.785	0.695
5.7 Articles in international conferences per academic staff member	(0.591 0.774 0.864)	0.346	0.793	0.723
5.8 Publications of book per academic staff member	(0.500 0.761 0.917)	0.397	0.793	0.698
6.1 Percent of international students (who major in Chinese)	(0.394 0.612 0.773)	0.498	0.666	0.584
6.2 Percent of international students (excludes those who major in Chinese)	(0.417 0.721 0.917)	0.447	0.767	0.660
6.3 Percent of international academic staff member	(0.345 0.681 0.917)	0.490	0.742	0.626
6.4 International cooperation projects	(0.583 0.749 0.955)	0.358	0.792	0.717
7.1 The success of the student body at winning national academic awards within 5 years	(0.500 0.702 0.917)	0.416	0.755	0.669
7.2 Graduate employment	(0.667 0.769 0.864)	0.302	0.789	0.743
7.3 Correspondent	(0.417 0.764 0.917)	0.433	0.795	0.681
8.1 Peer subjective assessment of undergraduate education quality	(0.591 0.798 0.955)	0.339	0.825	0.743
8.2 Employer subjective assessment of undergraduate education quality	(0.500 0.777 0.917)	0.392	0.804	0.706

Table 4 continued

Measure	Triangular fuzzy number	Left utility score	Right utility score	Total score
8.3 Graduate subjective assessment of undergraduate education quality	(0.591 0.778 0.917)	0.345	0.805	0.730
9.1 Inclusiveness	(0.417 0.725 0.955)	0.446	0.776	0.665
9.2 Total budget on student stipends/total student population	(0.500 0.651 0.864)	0.434	0.712	0.639
9.3 Expense as subvene for the poor students	(0.750 0.758 0.955)	0.248	0.798	0.775

[a] Taiwan Social Science Citation Index (TSSCI)

differences in relation to the commercial rankings published by the US NEWS and Guardian. It is interesting to note that some factors were not included in US NEWS and Guardian's rankings: research, internationalization and equity. Table 7 summarizes agreed reasons by respondents on how selected determinants contribute to the understanding of undergraduate education quality.

In this study, some *interesting categorizations of these determinants* (as Fig. 2 shows) were different from current commercial rankings. Firstly, respondents saw student selectivity as a determinant of dual characteristics: it is not only an input but also a process indicator of undergraduate education quality and receives a moderately high weighting. This finding pertains to the idea that the influence of outstanding peers is even greater than that of good professors. Chinese people have a saying that is similar in essence to this idea that good friends as well as good teachers are mentors. Outstanding peers on one hand suggest a university is of good enough quality to attract a high standard of student while on the other hand, they could be deemed as informal mentor that have the potential to facilitate a student's learning on and off the campus. Another interesting categorization is that of internationalization as a process indicator. Through means of international exchange of knowledge, academic staffs and students, it also enriches student learning experience in a way that is different from that universities themselves can provide alone. Finally, research has also been included as a process determinant. Although it is suggested in the teaching-research nexus literature that whether or not an academic's ability to conduct high quality research is essential for good teaching (Jenkins et al. 2003), the respondents considered research is an essential factor in determining academic quality in their teachers. In other words, from a respondent's perspective, the linkage between research and teaching should be emphasized and is considered beneficial for the improvement of overall education quality.

According to the results (see Table 5), the weight of determinants indicates that the respondents generally believed the factors relating to the education process (faculty qualification, curriculum and teaching, research, internationalization) were more important than both the output (reputation, performance of students and graduates) and input factors (financial resource, and student selectivity). Overall, the input factors received the least respect. This result supports Dill and Soo's (2005) claim that valid university rankings focus on process measures that link student learning and student output, while input measures are given minimal weight. Input measure should be considered but it should not play the predominant role in university ranking. However, why did the respondents consider the process factors as more important than the output ones (or value-added measures)? There were three possible reasons: first, reputation (peer, employer and graduate subjective assessment) was generated from survey data that could have been manipulated by

Table 5 The final ranking system

Determinant		Measure	Weight (%)		
Input	Financial resource	6.1 Revenue per student	21.55	10.04	4.82
		6.2 Library spent per student			5.22
	Student selectivity	1.1 Acceptance rate		11.51	4.95
		1.2 Entry score			6.56
Process	Faculty qualification	2.1 Percent of full-time faculty with top terminal degree	46.48	11.98	4.15
					3.95
		2.2 Percent of full-time faculty as professor			3.88
		2.3 Percent of full-time faculty			
	Curriculum and teaching	3.1 Staff-student ratio		12.89	3.70
		3.2 Senior student assessment score			4.28
		3.3 National teaching assessment score (by academic peers)			4.91
	Research	5.1 Research grants per academic staff member		11.83	1.54
		5.2 Percent of academic staff member with National Faculty Awards			1.48
					1.47
		5.3 Percent of academic staff member with Academy membership			1.40
					1.50
		5.4 Publications on Nature, Science, SCI, SSCI, TSSCI, EI and A&HCI per academic staff member			1.46
					1.52
		5.5 Citations per article on Nature, Science, SCI, SSCI, TSSCI, EI and A&HCI			1.46
		5.6 Articles in peer-reviewed journals per academic staff member			
		5.7 Articles in international conferences per academic staff member			
		5.8 Publications of book per academic staff member			
	Internationalization	7.1 Percent of international students (who major in Chinese)		9.78	2.21
					2.50
		7.2 Percent of international students (excludes those who major in Chinese)			2.37
					2.70
		7.3 Percent of international academic staff member			
		7.4 International cooperation projects			
Output	Performance of students and graduates	4.1 The success of the student body at winning national academic awards within 5 years	23.13	11.50	3.68
					4.08
		4.2 Graduate employment			3.74
		4.3 Correspondent			
	Reputation	8.1 Peer subjective assessment of undergraduate education quality		11.63	3.97
					3.77
		8.2 Employer subjective assessment of undergraduate education quality			3.89
		8.3 Graduate subjective assessment of undergraduate education quality			
Equity		9.1 Inclusiveness	8.84	8.84	2.83
		9.2 Total budget on student stipends/total student population			2.72
					3.29
		9.3 Expense as subvene for the poor students			

institutions in order to inflate their rankings (Ehrenberg 2002). Second, output measures might reflect student selectivity of one university as opposed to the quality of education (Montgomery and Canaan 2004). Finally, it is not easy to define what is the output of higher education and to find precise way to measure (or evaluate) it. These would make the attempt to understand the output or value-added effects resulted from education more difficult.

Table 6 A comparison of this study's ranking system with those of the US NEWS and Guardian

	US NEWS	Guardian	This study
Financial resource	Financial resources (10 percent); faculty salary (7%)	Spending per student (17%)	Revenue per student (4.82%); library spent per student (5.22%)
Student selectivity	Student selectivity (15%)	Entry score (17%)	Acceptance Rate (4.95%); entry score (6.56%)
Faculty qualification	The proportion of professors with the highest degree in their fields (3%); the proportion of faculty who are full time (1%)	None	Percent of full-time faculty with top terminal degree (4.15%); percent of full-time faculty as professor (3.95%); percent of full-time faculty (3.88%)
Curriculum and teaching	Class size (8%); the student-faculty ratio (1%)	Staff/student ratio (17%)	Staff-student ratio (3.70%)
	Alumni giving rate (5%); retention (20%)	Teaching quality—as rated by final year students on the course (10%); feedback (assessment)—as rated by final year students on the course (5%)	Senior student assessment score (4.28%)
	None	None	National teaching assessment score (by academic peers) (4.91%)
Research	None	None	11.83%
Internationalization	None	None	9.78%
Performance of students and graduates	Graduation rate performance (5%)	Value added—comparing students' degree results with their entry qualifications (17%)	The success of the student body at winning national academic awards within 5 years (3.68%)
	None	Job prospects (17%)	Graduate employment (4.08%)
	None	None	Correspondent (3.74%)
Reputation	Peer assessment (reputation) (25%)	None	Peer subjective assessment of undergraduate education quality (3.97%)
	None	None	Employer (3.89%) and graduate (3.77%) subjective assessment of undergraduate education quality
Equity	None	None	8.84%

In contrast, the distribution of the weights into the three categories in the rankings by US NEWS and Guardian is fairly even (see Table 8). The process score in the US NEWS ranking is also the highest. However, the index used for this category (especially the retention rate which receive the highest weighting of 20%) were considered by respondents in this study as invalid for they do not directly explain the quality of teaching and in fact a variety other reasons could lead to student dropout rates, such as financial problems.

Table 7 A summary of how the determinants contribute to the understanding of undergraduate education quality

	Determinant	Contribution to the understanding of education quality
Input	Financial resource	The more resources one university possesses, the more likely it provides a high quality learning environment, teaching staff and learning support
	Student selectivity	Whether a university can attract good students depends on its structure and quality (at least from students and parents' perspective). Moreover, students are enriched by the input of their peers. The influence of outstanding peers is far greater than first imagined and is even more important than good teachers
Process	Faculty qualification Curriculum and teaching Research	Teaching and the curriculum are the core of education quality and should be the leading concern in terms of rankings for school choice. Ideally, teaching and research should have a positive correlation in order to increase the education quality
	Internationalization	Quality of higher education should also be characterized by its international dimension: exchange of knowledge, interactive networking, mobility of teachers and students. Moreover, an international network, exchange of academics and facilitated student interaction between themselves and the world will enrich their learning experience at school
Output	Performance of students and graduates	Since US News was criticized for overemphasizing on input performance in university rankings, output indicators could make the ranking league table more persuasive
	Reputation	Different stakeholder perspectives should be elicited within the discourse of ranking to provide more comprehensive and up-to-date information and increase the validity of ranking. It is also recognized that a university with a good reputation will have a branding effect on their students thus promoting the value of their degree
Equity		Excellence, appropriateness and equity should be considered in all educational policy in a balanced way. As the access to higher education grows, more attention should be paid to issues of equity in higher education.

Table 8 The weight distribution of three categories given by the US NEWS, Guardian and this study

	US NEWS	Guardian	This study (%)
Input	32%	34%	21.55
Process	38%	32%	46.48
Output	30%	34%	23.13
Equity	0	0	8.84

It is argued by Green (1994) that important perspectives should be included in the ranking system to ensure a thorough and balanced understanding of the complex concept of undergraduate education quality today. This was firmly shown through the inclusion of differing perspectives from academics, clients (students, employers), different dimensions of quality in terms of equity as well as excellence and viewpoints from international students/institutions. Moreover, in response to the emerging role of market forces, it was also suggested that the inclusion of measures of graduate assessment (curriculum and teaching), graduate employment, employer and graduate subjective assessment of undergraduate education quality should be considered. However, if we take a look at the relative weights among different measures under each determinants, interesting findings can be

found. For example, under the same determinants of reputation, the employer and graduate subjective assessment of undergraduate education quality receive weightings of 3.77 and 3.89%, respectively while peer subjective assessment of undergraduate education quality gains the highest weighting of 3.97%. Similar findings were presented in the determinant of curriculum and teaching, where the graduate assessment weighting scored 4.28%, and national teaching assessment 4.91%. These results suggest that education today has to take into account the needs of clients but importantly also rely more on expert knowledge when defining the quality, meaning, and content of education.

As Table 6 shows, the measures in the other two commercial ranking systems, especially in the category of "curriculum and teaching" presented a customer oriented way of thinking about teaching quality: students were the main/only judge in determining a department/university's teaching/curriculum score. While this article argues that education is not a commercial product or merely a process of job preparation, there is also intellectual and moral aspects of education which are usually ignored under the pressure of marketization but matter more in the development of a whole person and citizen (Lewis 2006; Mok 2000). Students are undoubtedly in the best position to know what they want from school and teachers but not necessarily what they need. It would therefore be not only a threat for universities but also a danger for students themselves if too much weight was given to their opinions. On the other hand, the result suggest that employers' opinion might also be beneficial for understanding the deficiencies in today's undergraduate education. This would be correlated by the fact that nowadays, many students than ever before have undergraduate education most of who are going to be employed after graduation (in Taiwan, this issue is even more obvious by the fact that according to 2009 data, more than 90% of high school graduates entered higher education). Undergraduate education has to be reasonably responsive to the changes and societal needs and this cannot be done if universities continue to think of themselves as Ivory Towers.

Although this study's ranking system can by no means transplanted to any other country, based on the literature and findings of this study, implication for the ranking practice in other countries can be summarized as follows. First, higher education academics and researchers should continue to play active roles in the discourse and practice of university rankings. Rankings are here to stay, and will continue to have tremendous impact on readers and the development of higher education. Academics could respond in a passive manner and try to forget and dismiss rankings, but this will not make them disappear.

As for the methodological issues of rankings, the findings suggest that in a time of change, a better way to understand the quality of undergraduate education is the continual effort to incorporate important factors and perspectives from the discourse. In this way, ranking achieves a comprehensive picture of quality instead of a partial image in favor of particular dimensions. The quality of education is an abstract concept involving more philosophical thinking and debate. Unlike the height of one person, education quality of a university is not directly measurable; therefore, there must be statistical error in all rankings. In response, a ranking system with a focus on measuring the right things and providing relevant perspectives is suggested. Moreover, in considering the domain differences (Altbach 2006) and the irreplaceable value of comparative data at the departmental level of undergraduate choice, it would be better that all measures are compared at the departmental level first and then transferred to institutional ranks (if ranking on the level of university is to be made). In general, data was collected at both departmental and institutional levels. Departments were categorized according to subjects, and each measure score was derived from a comparative performance of departments within the same group. For example, a particular education department was ranked along side education departments from other

universities. Finally, in order to generate an institutional ranking, the department scores within each institution were averaged to determine an institutional score.

Ranking or evaluation concerns activity making value judgment and there must be trade-off among the long list of possible determinants and measures. The FDM is a technique that gathers subjective opinion from experts and converts them into quasi-objective data. This can be considered better than simply generating a mean of expert response in the form of a five-point scale or given percentages and it could be a useful tool in enhancing ranking methodology. Moreover, the criticism that current rankings are usually based solely on what data is currently available and not what is relevant and important also gives rise to the need for the institutions, departments and the governments to build a longitudinal higher education database which could serve as a tool and reference for self-study and accountability.

Given the exploratory nature of this study, several challenges have arisen in terms of future research. First, each determinant within the study's ranking system is in itself a huge concept. Even though the higher education experts were asked to make suggestions on what measures could be included, there is still room for further improvement concerning the reliability and validity of the measures under each category. Moreover, some of the measures rely on survey data, for example, the measure "employer's subjective assessment of undergraduate education quality". Further study is required on how to conduct these surveys in terms of sampling and method of data collection which are crucial for validity and reliability.

Conclusion

Given the distinct nature of education from commercial products, this study has aimed to construct an undergraduate education ranking system from the insiders' perspectives. In response to the complexity of education quality in higher education during a time of transformation, the results suggest a comprehensiveness of the ranking system in terms of measure types, data sources, and different stakeholder perspectives. This study addresses the idea of undergraduate education quality in a manner different from commercial rankings that usually reflect the thinking of market and consumer needs. However, it cannot be denied that quality is in the eye of the beholder (Marginson 2006). As such, a final recommendation is to suggest an overall ranking system based on expert opinion with special attention to the transparency of the ranking system in terms of the rationale, methodology and detailed scores of each determinants and measures.

Acknowledgments The authors would like to thank Dr. Amy Roberts from University of Wyoming for her full support and assistance during the preparation of this article.

References

Altbach, P. G. (2006). The dilemmas of ranking. *Higher Education Digest, 55*, 9–10.

Billing, D. (2003). Evaluation of a trans-national University Quality Assessment Project in Bulgaria. *Perspectives, 7*(1), 19–24.

Bodjanova, S. (2006). Median alpha-levels of a fuzzy number. *Fuzzy Sets and Systems, 157*(7), 879–891.

Bowden, R. (2000). Fantasy higher education: University and college league tables. *Quality in Higher Education, 6*(1), 41–60.

Brooks, R. L. (2005). Measuring university quality. *The Review of Higher Education, 29*(1), 1–21.

Buela-Casal, G., Gutiérrez-Martínez, O., Bermúdez-Sánchez, M. P., & Vadillo-Muñoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71*(3), 349–365.
Butler-Adam, J. (2007). Response to questionnaire on developing a quality rating mechanism for African universities. In Neil Butcher and Associates (Eds.). Johannesburg.
Carrico, C. S., Hogan, S. M., Dyson, R. G., & Athanassopoulos, A. D. (1997). Data envelopment analysis and university selection. *The Journal of the Operations Research Society, 48*(12), 1163–1177.
Chang, P. T., Huang, L. C., & Lin, H. J. (2000). The fuzzy Delphi method via fuzzy statistics and membership function fitting and an application to the human resources. *Fuzzy Sets and Systems, 112*(3), 511–520.
CHE Report. (2000). Towards a new higher education landscape: Meeting the equality, quality and social development imperatives of South Africa in the 21st century. Retrieved 1, March, 2008, from http://www.polity.org.za/govdocs/reports/education/chereport.html.
Chen, S. J., & Huang, C. L. (1992). *Fuzzy multiple attribute decision making methods and applications*. New York: Springer.
Clarke, M. (2002). Some guidelines for academic quality rankings. *Higher Education in Europe, 27*(4), 443–459.
Clarke, M. (2004). *International issues in quality assessment: Limitations of the weight-and-sum approach to ranking the academic quality of institutions of higher education*. Paper presented at the Proceedings of the Australian Universities Quality Forum 2004.
Clarke, M. (2005). Quality assessment lessons from Australia and New Zealand. *Higher Education in Europe, 30*(2), 183–197.
Dichev, I. (2001). News or noise? Estimating the noise in the US News university rankings. *Research in Higher Education, 42*(3), 237–266.
Dill, D. D., & Soo, M. (2004). *Is there a global definition of academic quality?: A cross-national analysis of university ranking systems.* Paper presented at the Public Policy for Academic Quality Research Program.
Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education, 49*, 495–533.
Eccles, C. (2002). The use of university rankings in the United Kingdom. *Higher Education in Europe, 27*(4), 423–432.
Ehrenberg, R. G. (2002). Reaching for the brass ring: The US News & World Report rankings and competition. *The Review of Higher Education, 26*(2), 145–162.
Federkeil, G. (2002). Some aspects of ranking methodology—The CHE-ranking of German universities. *Higher Education in Europe, 27*(4), 389–397.
Green, D. (1994). What is quality in higher education? Concepts, policy and practice. In D. Green (Ed.), *What is quality in higher education?* (pp. 3–20). Buckingham: Open University press and Society for Research into Higher Education.
Harvey, L. (2002). The end of quality? *Quality in Higher Education, 8*(1), 5–22.
Harvey, L., & Green, D. (1993). Defining quality. *Assessment and Evaluation in Higher Education, 18*(1), 9–34.
Hazelkorn, E. (2006). *Impact and influence of league tables and ranking systems on institutional decision-making*. Ireland: Dublin Institute of Technology.
Jenkins, A., Breen, R., Lindsay, R., & Brew, A. (2003). *Re-shaping teaching in higher education: Linking teaching and research*. London: Routledge.
Lagrosen, S., Seyyed-Hashemi, R., & Leitner, M. (2004). Examination of the dimensions of quality in higher education. *Quality Assurance in Education, 12*(2), 61–69.
Lewis, H. (2006). *Excellence without a soul: How a great university forgot education*: PublicAffairs.
Marginson, S. (2006). *Global university rankings at the end of 2006. Is this the hierarchy we have to have?* Paper presented at the OCEDE.
Marginson, S. (2007). *Global university comparisons: The second stage*. Paper presented at the International Trends in University Rankings and Classifications—Griffith University/IRU Symposium.
Merisotis, J. (2002). Summary report of the invitational roundtable on statistical indicators for the quality assessment of higher/tertiary education institutions: Ranking and league table methodologies. *Higher Education in Europe, 27*(4), 475–480.
Merisotis, J., & Sadlak, J. (2005). Higher education rankings: Evolution, acceptance, and dialogue. *Higher Education in Europe, 30*(2), 97–101.
Miyoshi, M. (2000). Ivory tower in Escrow. *Boundary 2, 27*(1), 7–50.
Mok, K. H. (2000). Impact of globalization: A study of quality assurance systems of higher education in Hong Kong and Singapore. *Comparative Education Review, 44*(2), 148–174.

Montgomery, L. M., & Canaan, J. E. (2004). Conceptualizing higher education students as social actors in a globalizing world: A special issue. *International Journal of Qualitative Studies in Education, 17*(6), 739–749.

Morrison, H. G., Magennis, S. P., & Carey, L. J. (1995). Performance indicators and league tables: A call for standards. *Higher Education Quarterly, 49*(2), 128–145.

Mullin, R., & Wilson, G. (2000). Quality of undergraduate education: Examining the current paradigm and system. *Journal of Quality Management, 5*, 225–246.

Ragin, C. C. (2007). Qualitative comparative analysis using fuzzy sets(fsQCA). In *Configurational comparative analysis*. London: Sage Publications.

Salih, H. (2003). *Discussions.* Paper presented at the Commission II—Quality of higher education.

Srikanthan, G., & Dalrymple, J. (2003). Developing alternative perspectives for quality in higher education. *The International Journal of Educational Management, 17*(3), 126–136.

Stella, A., & Woodhouse, D. (2006a). Ranking of higher education institutions [Electronic Version]. Retrieved February 3, 2008, from http://goliath.ecnext.com/coms2/gi_0199-5479051/The-mission-of-the-university.html.

Stella, A., & Woodhouse, D. (2006b). Ranking of higher education institutions [Electronic Version]. *AUQA Occasional Publications Number 6*.

Stephenson, S. L. (2004). Saving quality from quality assurance. *Perspectives, 8*(3), 62–67.

UNESCO. (2006, 6–8 February 2008). *International conference on accreditation, quality assurance and recognition of qualifications in higher education in Africa.* Paper presented at the International Conference on Accreditation, Quality Assurance and Recognition of Qualifications in Higher Education in Africa, Nairobi, Kenya.

Usher, A., & Savino, M. (2006). *A world of difference: A global survey of university league tables*. Toronto: Educational Policy Institute.

Van Dyke, N. (2005). Twenty years of university report cards. *Higher Education in Europe, 30*(2), 103–125.