Worldwide university ranking and its underlying basis: a perspective of university orientation towards excellence

Wenjie Xu

MA thesis
University of Iceland
School of Education

HÁSKÓLI ÍSLANDS

Worldwide university ranking and its underlying basis: a perspective of university orientation towards excellence

Wenjie Xu

60 ECTS thesis submitted in partial fulfillment of a Master degree in International Studies in Education

Supervisor: Prof. Allyson Macdonald
Co-supervisor: Dr. Amalía Björnsdóttir

Faculty of Education Studies, School of Education
University of Iceland
Reykjavik, November 2012

Worldwide university ranking and its underlying
basis: a perspective of university orientation
towards excellence

This thesis is a 60 credit final project towards the
MA degree in International Studies in Education,
School of Education, University of Iceland.

Preface

I would like to express my sincere gratitude to all these who support me to complete this master project.

First and foremost, I would like to thank my supervisors, Prof. Allyson Macdonald and Dr. Amalía Björnsdóttir, for their excellent guidance, constructive suggestions and warmhearted assistance throughout this study. I also want to thank Prof. Börkur Hansen, for his valuable comments and recommendations.

I would like to extend my grateful thanks to Prof. Pall Skúlason, professor of philosophy and former rector of the University of Iceland, for giving me valuable insights and suggestions.

My special thanks to Prof. Ólafur Páll Jónsson and Anna María Hauksdóttir, for all their heartfelt assistance during this study.

In addition, I want to thank my husband, for his valuable advices, steadfast endorsement and encouragement. And, thanks to my beloved parents and all friends, for their moral support during my study aboard.

Last but not the least, I would like to thank all tutors from Faculty of Education Studies, for their devoted teaching and enlightening lectures helping me to develop my background in education studies.

Abstract

Worldwide university ranking (WUR) has aroused heated debate and a great concern in the field of Higher Education Institutions. Both the meaning and accuracy of WURs give rise to arguments because of the high diversity in universities across nations.

This study reviews the social and historical background of WURs by examining their birth, present situation, future direction and the soil for university ranking development. To understand WUR in depth, the conception of ranking is analyzed and compared with "evaluation" and "clarification". Different types of WURs are featured and summarized. The methodologies of WUR involving related indicators and weightings are deciphered as well as key steps in producing rankings. The impacts of university rankings are examined and both advantages and disadvantages of WURs are investigated in many aspects.

In this multiple-case study, representative WURs, i.e., *Times Higher Education (THE), Academic Ranking of World Universities (ARWU), QS together with US News & World Report (QS-USNWR)* and *U-Multiriank* project, are selected based on their characteristics and influence. Both results and methodologies of rankings are well documented. The top 10 as well as the top 50 universities ranked by *THE, ARWU* and *QS-USNWR* are extracted and agreement among rankings of top-tier universities is evaluated by both qualitative and quantitative comparisons. It is found that different WURs lead to similar consequences. The WUR system is utilized as one perspective and toolkit to study university excellence. Top-tier universities in WURs are categorized and the results show that English-speaking and developed countries have an obvious advantage in WUR systems.

All in all, WUR offers us a new angle to access to world-class universities and would serve as an effective tool to visualize university excellence. However, WUR should not be abused since not everything can be measured in regard to university and higher education.

Ágrip

Röðun háskóla á lista yfir bestu háskóla heims hefur mikið verið til umræðu og valdið áhyggjur ýmissa innan háskólasamfélagins. Deilt er um hversu merkingarbærir og nákvæmir þessir listar eru m.a. vegna þess að háskólar eru mjög fjölbreyttar stofnanir. Í þessari rannsókn er farið yfir sögu þessara röðunarlista, hvernig þeir urðu til og hvaða hlutverki þeir genga í dag. Í framhaldinu er velt upp hver framtíð þeirra gæti orðið og í hvað átt þeir gætu þróast.

Til að öðlast betri skilning á röðunarlistunum er röðunin (ranking) borin saman við mat (evaluation) og skýringar (clarification). Gerð er grein fyrir ólíkum röðunarlistum, hvað þeir eiga sameiginlegt og hvað skilur þá að. Aðferðafræðin að baki röðunarlistum er skoðuð t.d. hvaða árangursvísar eru notaðir og hvaða vægi þeir fá í röðun háskóla á lista. Velt er upp áhrifum röðunar á starf háskóla og hvaða kostir og gallar fylgja henni.

Þeir röðunarlistar sem skoðaðir eru sérstaklega í rannsókninni eru: Times Higher Education (THE), Academic Ranking of World Universities (ARWU), QS ásamt US News & World Report (QS-USNWR) og U-Multirank. Þessir röðunarlistar voru teknir til skoðunar vegna ákveðinna eiginleika þeirra og áhrifa sem þeir hafa. Gerð er grein fyrir þeirri aðferðafræði sem þeir byggja og farið yfir niðurstöður röðunarinnar. Skoðað er sérstaklega samræmi í röðun efstu 10 og efstu 50 skólanna á THE, ARWU og QS-USNWR. Þessi samburður er gerður með fjölbreyttum aðferðum. Niðurstaðan er sú að mikið samræmi er í röðunarlistunum þ.e. ólíkir aðferðir gefa svipaða niðurstöðu. Þeir skólar sem raðast ofarlega eru í flestum tilfellum í löndum þar sem enska er móðurmál og í iðnvæddum löndum.

Röðunarlistar veita okkur nýtt sjónarhorn til að meta háskóla. En röðunarlista segja ekki alla söguna og það þarf að varast að misnota þá. Ekki er hægt að mæla allt það sem skiptir máli í starfi háskóla og því gefa röðunarlista aldrei fullkomna mynd af gæðum þeirra.

Contents

List of figures

List of tables

1 Introduction

Recent years have witnessed a substantial increase in the number of university ranking systems. Ranking arouses great concern in the field of Higher Education Institutions (HEIs), and also has a large impact on diverse stakeholders. Every year, various kinds of university rankings are published and updated by governments, academics, eminent magazines and newspapers. University ranking has become a heated issue all over the world. Worldwide university ranking (hereafter called WUR) has generated fierce debates and plenty of discussions especially in relation to ranking accuracy and usefulness. Nevertheless, statistical and multifarious rating criteria have provided us a new perspective on research into university excellence, and have become a toolkit to examine university information from around the world.

The aim of this study is to explore university orientation towards excellence by understanding and analyzing four representative worldwide university rankings published for 2011-2012,

➢ • Academic Ranking of World Universities (ARWU) or Shanghai Jiao Tong University Rankings (SJTU),

➢ •European Multidimensional University Ranking System – EU funded project (U-Multirank),

➢ • Times Higher Education World University Ranking (THE),

➢ • World's Best Universities Ranking – US News and World Report in cooperation with Quacquarelli Symonds (QS-USNWR).

This study reviews historical and social context to help us to comprehend the status of university ranking and its development. Selected WURs are analyzed to determine their common ground and relationships between them. Then, the results are used to discuss and clarify the characteristics of top universities, and the challenges and directions for university development in the future.

1.1 Theoretical context

In order to contextualize the study before discussing university excellence in our time, it is necessary to examine some key issues, such as the university and globalization, university ranking, and the world-class university. The globalization context is the background of the study while university ranking is the carrier for researching the world-class university.

1.1.1 Globalization and university

Educationists have noted that "'internationalization' and 'globalization' have become buzz words in higher education and practice" (Guri-Rosenblit, Sebkova, & Ulrich, 2007, p. 8). Through globalization and informationization, the world is developing into a knowledge-based society. Hazelkorn (2009) has suggested that "the evolution from agricultural to industrial to knowledge production has transformed every aspect of society" and higher education is the top priority of a policy agenda since "knowledge has become the foundation of economic, social and political power" (p. 3). In general, globalization and an evolutionary knowledge-based society have led to similar and significant changes of higher education with regard to functions and characteristics all over the world although the localization is still important (Deem, Mok, & Lucas, 2008, p. 83).

"University" is a term that refers to "an institution of higher learning, conducting teaching and research at the undergraduate and postgraduate level" (as cited in Taylor & Braddock, 2007, p. 246). Three activities belong to the core work of the university: teaching, research and service provided to society derived from teaching and research directly or indirectly (Skulason, 2009).

Globalization has a salient influence on university change and development in modern society. Specifically, universities are becoming both collaborative and competitive partners in international schemes.

Comprehensive collaboration between universities or academic institutions across nations has been more and more popular. Guri-Rosenblit et al. (2007) have pointed out that "students, academic staff and curricula are transferred and exchanged between institutions; accreditation agencies ensure promptness in

accrediting previous experiential learning and previous academic studies; governments append their signatures to cooperative projects in higher education" (p. 8–9). Meanwhile, diverse cultures permeate gradually through the contemporary campus since the proportion of international students has increased in universities.

The global evolution of university has intensified competition among universities. For example, university administrators develop policies to attract excellent students and scholars all over the world. University globalization forces a rearrangement of educational and academic resources and simultaneously leads to constant pressure on universities due to fierce competition. The conduct of such cooperation and competition by universities has in turn promoted step by step the development of globalization. Therefore, the university is not only a recipient or beneficiary of globalization, but also an important driver in the process of globalization.

1.1.2 World-class university

In the past decade, the term "world-class university" has been used to describe research universities on the pinnacle of the higher education hierarchy (Salmi, 2011, p. 323). In fact, pursuing excellence by universities has great impacts on higher education. Williams and van Dyke have written (as cited in Salmi, 2009):

> In the past decade, the term "world-class university" has become a catch phrase, not simply for improving the quality of learning and research in higher education but also, more important, for developing the capacity to compete in the global tertiary education marketplace through the acquisition, adaptation, and creation of advanced knowledge. (p. 3–4)

Williams and van Dyke (as cited in Salmi, 2009) have further commented that global competition partly results from resource distribution in that students as well as their parents want to optimize their choice for university life on a global scale, while government always wants to obtain maximum profit of investment so that a sign of global standing is a top issue for university development.

Indeed, world-class university is highlighted by stakeholders all over the world (Huisman, as cited in Yang & Welch, 2012). However, a definition of 'world-class university' is missing although the target is clear (Mohrman, as cited in Salmi, 2009): "Everyone wants one, no one knows what it is, and no one knows how to get one" (Altbach, as cited in Salmi, 2009, p. 4).

It is argued that the essence of university ranking is to classify and ascertain university excellence by measuring criteria and then determining a list of world-class universities. The proliferation of WUR in the past few years has created systematic ways to indentify and classify world-class universities (Salmi, 2011, p. 324). In this study, WUR will be used as a tool to generalize world-class universities, as well as to discuss key features of top universities from a statistical angle.

University ranking is currently one of the most popular means to determine university excellence. It is worth considering the term "world-class university" in an evolutional and dialectical manner, which could help to recognize the pros and cons of university ranking.

On the one side, the level of evolution of tertiary education makes it necessary to reassess "academic excellence" and "quality differentiation" at diverse levels, e.g. local, national and international levels. Further, the world-class university can be regarded as a distinct feature of HEIs in the globalization age and also over time. Therefore, "the meaning of the concept is then somewhat fluid, dependent on the context and also, if as is often the case, related to league tables, the indicators used" (Deem et al., 2008, p. 85).

On the flip side, it should be pointed out that an evaluation of a modern university is highly complicated and dynamic and any simple definition of world-class university is inaccurate and not suitable. Sadlak and Liu (2009) have suggested that "any categorization is quite arbitrary and carries with it only a limited, brittle notion of 'superior performance' which does not fit well in particular regard to a highly complex organization as the modern university" (p. 13).

Thereby, world-class university is a term of great complexity and it puts forward the notion of "academic excellence" or "appraising university quality". It prompts a debate on the ranking

systems and related results. In addition, "the very term 'excellence' reflects and is an issue of scarcity, that is, not everyone can reach it but everyone can aspire to it" (Sadlak & Liu, 2009, p. 13).

1.1.3 University ranking

Ongoing globalization and associated challenges make progressive and renowned universities interested in proving their performance "through global university league tables or ranking exercises" (Deem, et al. 2008, p. 84). However, it is hard to define university ranking clearly and most people or organizations just explain it as a way of depicting facts according to existing ranking results.

For instance, Usher (as cited in Swedish National Agency for Higher Education, [SNAFHE], 2009) has indicated that "rankings are simply collections of indicators" (p. 13). Lukman, Krajnc, and Glavic (2010) have also claimed that "university ranking belongs to the field of social assessment, which aims to evaluate and rank university quality by combinations of varied impact factors" (p. 619).

Nevertheless, some basic components and crucial principles of university ranking can be summed up. SNAFHE (2009) has summarized six traits as follows:

1. Ranking is based on a series of indicators

2. It is assumed that indicators can assess the quality of higher education

3. Indicators encompass a great deal of information, compiled from very different sorts of data ranging from published statistics to subjective experience

4. The indicators focus on a specific unit that ranges from education programs, disciplines, and departments to whole HEIs

5. ndicators sum up cohesive, aggregated results

6. Ranking yields results which are listed "in order of precedence of the units covered" and makes "a comparison of the results achieved". (p. 14)

The Swedish summary reveals that university rankings use overall scores and these scores are usually formed through a combination of indicators of quality assigned different and predetermined weights by collecting university data that are regarded as indicators of quality.

Besides the various criteria and elaborate results, the goals of university ranking are closely linked to the globalization of higher education since the birth of WUR. According to the conclusions summed up by Lukman et al. (2010), university ranking mainly devotes itself to achieving the following aims:

1. Directing an entrant to higher educational programmes

2. Evaluating the phenomena of the international higher education market

3. Introducing market directions for universities at national levels

4. Enhancing sound and positive competition for students, professors, and the funders of universities. (p. 619)

In addition, university ranking can offer information about the quality and characteristics of HEIs for dozens of stakeholders, e.g. students and their parents, faculty and education policy makers. It should also be noted that, even with objective and reliable indicators, the concept of university ranking is too narrow to judge comprehensive achievements of colleges because of inevitable structure defects. In other words, it is unreasonable to evaluate universities with a uniform standard.

University ranking is regarded as an international transparency tool to probe the principles and new features of preeminent universities and university development in the present era.

1.2   An overview of worldwide university rankings

Numerous academics, policymakers, university administrators and students are eager to update university ranking year by year. Rauhvargers (2011) recognized that "as long as only national university rankings existed, they were popular and important in some countries, while other university systems did not pay much

attention to them" (p. 19). However, the WUR attracts more and more interest from the public.

Kazimi-erz Bilanow, managing director of the International Ranking Expert Group (IREG) Observatory on Academic Rankings and Excellence that was created in 2009 as a Warsaw-based ranking quality-assurance body, has pointed out that "rankings have outgrown the expectations of those who started them" (Butler, 2010, p. 16). The extension or outreach of university ranking has offered sweeping powers to the university research field and even raised "movement in the ranks" (Movement in the ranks, 2011, p. 435).

The first national university ranking was published by the *US News and World Report* as an annual American best colleges' review in 1983 (Dill, 2006). It should be pointed out *US News and World Report* ranking was the first time that "assessment information became easily accessible to prospective undergraduates and their parents, thus creating a dramatic shift in the consumers of quality assessments" (Webster, as cited in Brooks, 2005, p. 6). Dichev, Longden, Marginson and Pike (as cited in Shin & Harman, 2009) believe that "university rankings have increased to a marked degree in importance since the *US News and World Report* published its first rankings" (p. 10).

There was no truly international university ranking until *Shanghai Ranking Consultancy* (initially *Shanghai Jiao Tong University*) published *Academic Ranking of World Universities* (*ARWU*) in 2003. Raughvargers (2011) indicated that the results of ARWU shocked the world and especially in Europe, since US and UK universities overwhelmingly dominated the Top 20 and Top 100 lists of ARWU (p. 19). "And now it is the rankings' turn to be assessed" (Butler, 2010, p. 17). The European University Association has published a report on rankings for evaluating global university rankings and their impacts. The report was named as "EUA report on ranking 2011, global university rankings and their impact" (hereafter called EUA report on ranking 2011).

Although the ranking movement is thriving, the voice of objection or criticism exists especially concerning usefulness and accuracy of rankings.Charon and Wauters (2008) believe that university rankings are debatable and have questioned some of the indicators used, such as the particular weight given to articles

published in *Nature* and *Science* or the number of Nobel Prizes and Fields Medals winners educated at a given institution (p. 62). Marginson believes that "ranking tables conceal a whole array of methodological problems and anomalies. It is often unclear why a particular indicator was chosen, by whom it was decided, and how open and reflective the decision process was" (as cited in Lukman, et al. 2010, p. 619).

There is no doubt that WUR has many flaws and biases and for some it seems nonsensical to take ranking positions too seriously since rankings and university quality together are similar to "the six blind men and an elephant" as shown by the cover story of EUA report on ranking 2011 (Rauhvargers, 2011, p. 66 & p. 5).

There are however two aspects worth pondering in the WUR system. First, world wide assessment systems for higher education underscore a cosmopolitan perspective according to mainstream internationalization, and this sort of quantitative research method provides a viable goal for the improvement of higher education quality despite intense opposition. Furthermore, the diversity and unity reflected in the changing data is evident in relation to the stratified category. These facts encouraged me to dig into the relationship between WUR and modern HEIs.

In the literature to date, there is little in-depth analysis of rating systems and with relation to HEI development. This study introduces a way of thinking about the existing WUR system. In particular, the WUR system is utilized as a perspective to examine the world-class university in the globalization era. In order to develop this research, the "underlying basis" of WUR will be clarified.

1.3   Research questions and methodology

The annually updated ranking results have helped to turn the spotlight to discussions on university performance again and again, with regard to either the league tables or the indicators and their weightings. For one thing, methodologies have became more sophisticated in response to criticism; for another, ranking results are published every year, and influence diverse stakeholders more and more.

Are there any relationships or similarities among different WURs in terms of their results and methodologies? If so, what

relationships or similarities are? Further, with these clues and analysis what will WUR bring us with regard to university excellence as HEIs grapple with great internationalization and highly knowledge-based society in our times?

The aim of this study is to determine the relationships and similarities among selected representative international rankings by comparative and quantitative analysis, and then to discuss the university rankings' effectiveness or discourse concerning university excellence, as well as explore the traits and underlying principles of the concept of world-class university.

In particular, the similar indicators of selected WURs will be examined and analyzed in order to further understand these WURs. And the correlation between selected WURs' results will be calculated and discussed to visualize the relationship between these WURs.

1.4   Structure of the study

This study first reviews key issues in university ranking field in literature. WUR is examined through its birth, present situation and future direction. The conception of ranking is clarified by comparisons with "evaluation" and "classification". The methodologies of WUR as well as key steps in producing rankings are deciphered. Further, the impacts of university rankings are evaluated and both advantages and disadvantages of WUR are summarized.

Then research methodology is provided and followed by sections describing the findings, summary and discussions. Research methodology is introduced briefly. As a multiple-case study, the framework of research design is proposed and specific methods and procedures are determined. Findings are generated by means of understanding and analyzing associated criteria and quantitative data of WURs. After a short summary of the findings, discussions are provided based on the findings as well as literature review. Finally, traits of the top-tier university are demonstrated.

2 Literature Review

University performance in our time has become an issue of growing concern. The masses have been obsessed with "university excellence", despite the fact that the term is hard to define. University ranking can be indentified as a succinct way to recognize world-class universities all over the world. To get a full appreciation of WUR, this section starts with a brief review of the birth and background of WUR and then looks into related issues, such as the concept and impacts.

2.1 The rise of WUR

2.1.1 WUR — the past

University ranking has been developing for a long time. According to Lukman et al. (2010), university ranking is an international phenomenon with more than 25 years of history, ever since the *US News and World Report* started to publish America's annual best colleges' review in 1983 (p. 619). In history, however, university ranking can date back to the 19th century. Salmi and Saroyan (as cited in Rauhvargers, 2011) noted that *The Commission of US Bureau of Education* engaged in classifying and compiling university rankings since 1870 (p. 19). Various university ranking activities were performed sporadically throughout the 20th century (SNAFHE, 2009, p. 11).

In particular, *Chesley Manly of The Chicago Tribune* published six different rankings in 1957, including ten best universities, co-educational colleges, men's colleges, women's colleges, law schools and engineering schools. This was really the first media ranking of universities and higher education institutions (Rauhvargers, 2011; SNAFHE, 2009). *The Fiske Guide to Colleges* in 1982 and the *US News and World Report's rankings* in 1983 started to extend to undergraduate education and began to generate more extensive ranking activities in the higher education sector (SNAFHE, 2009, p. 12). Notably, *US News and World Report* mainly fulfilled the information needs of prospective students, which is usually considered as the starting point of university ranking.

2.1.2 WUR — the present

Globalization affects university development in modern times, and also influences current ranking activities. A notable feature of university rankings at present is that they are in favor of collecting data and appraising university performance on a global basis. Buela-Casal, Gutierrez-Martinez, Bermudez-Sanchez, and Vadillo-Munoz (2007) have pointed out that the increasing global mobility of students due to technological expansion and economical development has changed academic ranking systems from nation-specific approach to providing international characterization although the initial rankings and university analysis were only national, for example, rankings of USA universities, Chinese universities, German universities, and Japanese universities. They (2007) have remarked that the Institute of Higher Education of Shanghai Jiao Tong University was the "precursor of an academic ranking of universities worldwide. After this initiative, the purpose of doing global rankings of universities based on international comparable academic data has been followed by other entities (p. 351).

The *Academic Ranking of World Universities (ARWU)* published by Shanghai Ranking Consultancy (initially Shanghai Jiao Tong University) is the earliest attempt to evaluate and rank university performance on a global scale and mainly concentrates on academic aspects, which was designed in order to "ascertain the relative position of Chinese universities internationally", attracting much interest all over the world (Dehon, McCathie, & Verardi, 2010, p. 516).

Later, *The Times Higher Education Supplement* published *Times Higher Education World University Ranking (THE)* in 2004. In one sense, this ranking was a European response to ARWU, and the number of rankings has grown (Rauhvargers, 2011) since then, for instance, *The Ranking Web (or Webometrics Ranking)* published by *The Spanish Research Council*, *The Performance Ranking of Scientific Papers for World Universities* published by the *Higher Education Evaluation and Accreditation Council of Taiwan*, *Leiden Ranking* published by *Leiden University*, Netherlands, and *CHE University Ranking* published by *Centre for Higher Education Development*, Germany. In fact, the number of international ranking systems is still increasing despite criticism. User and

Medow (as cited in Shine & Toutkoushian, 2011) have pointed out that at the time of their ranking study, there were at least 26 ranking systems all over the world, and new rankings are emerging in many countries.

2.1.3 WUR — the future

It is not surprising that there are intensive concerns regarding the quality of university ranking because of its expanding influence over the field of HEIs (Cheng & Liu, 2008, p. 201).

The International Ranking Expert Group (IREG) was established in 2004 by the *UNESCO European Centre for Higher Education (UNESCO-CEPES)* in Bucharest, Romania and the *Institute for Higher Education Policy (IHEP)* was founded in Washington, DC at the same time (IREG Observatory Academic Ranking and Excellence, 2006). In May 2006, IREG's second meeting was held and a guideline document was created on how to produce university rankings, i.e. the Berlin Principles on Ranking of Higher Education Institutions(here after called the Berlin Principles), providing a set of principles for good practice and quality in HEI rankings (IREG Observatory Academic Ranking and Excellence, 2006).

The Berlin Principles (see http://www.ireg-observatory.org/) have been regarded as groundbreaking principles and also had a profound influence on university ranking development. Sixteen principles are listed, covering four aspects: the target and purpose of ranking; the formation of indicators; the collection and analysis of data; and the presentation of results (Cheng & Liu, 2008, p. 201). To some extent, "the Berlin Principles provide implications for the future of rankings and how they may be improved" (Shin & Toutkoushian, 2011, p. 12).

Based on the Berlin Principles, the IREG Ranking Audit was proposed by the IREG Observatory on Academic Ranking and Excellence (IREG Observatory) on 15 December 2011. According to IREG Observatory (2011), the Ranking Audit is a voluntary audit procedure for various rankings and each ranking that passes robust assessment will be entitled to use the quality label "IREG approved". Ultimately, the purpose of the IREG Ranking Audit is to "enhance the transparency about rankings, give users of rankings a tool to identify trustworthy rankings; and improve the overall quality of rankings" by assessing the ranking criteria

(IREG Observatory, 2011, p. 5). Specifically, the IREG Ranking Audit involves five dimensions to evaluate ranking quality:

1. Purpose, target groups, basic approach

2. Methodology

3. Publication and presentation of results

4. Transparency, responsiveness

5. Quality assurance. (p. 6–8)

According to these five dimensions, twenty criteria have been developed including ten core criteria with double weight and ten standard criteria, and each criterion is assessed on a point scale (see IREG Ranking Audit Manual on http://www.ireg-observatory.org/). It is clear from the procedure that the IREG Observatory has put some efforts into creating a set of principles for more objective and reliable university rankings although there is still a long way to go.

In addition, there are many new attempts to develop ratings, rankings and classifications for all HEIs and their various missions, such as U-Multirank (Rauhvargers, 2011, p. 68). It demonstrates thatthe methodologies of university rankings are changing from a single dimensional to a multi-dimensional orientation.

To sum up, IREG Observatory contributes to auditing the processes of producing university ranking and university rankers are making improvements with some systemic changes in the methodologies of university rankings. Looking ahead, "there are no indicators that ranking will fade" (Shin & Toutkoushian, 2011, p. 67).

2.1.4 Social background — the soil for development

With the advent of university ranking, various rating systems and league tables have became popular in measuring university performance. Their existence and development are driven by complex factors including the external stimulation of social change and the internal desire of higher education. The mass media stirs up the fire of ranking. Rauhvargers (2011) has suggested that:

It would be naive to imagine that the media will ever give up a tool such as the global university rankings, which attract thousands of readers when the new results are published and which allows suspense to be maintained over an entire year, by publishing tiny snippets of information about minimal changes in ranking methodologies. The general public as well as politicians will always like easily readable tables, which clearly state which universities are the very best in the world. (p. 20)

In analyzing the development of ranking systems, "massification, marketization, and globalization of higher education" are three important elements that cause increasing interest in ranking systems (Shin, Harman, & Dill, as cited in Shin & Toukoushian, 2011, p. 3).

Indeed, rapid developments in the higher education market in recent years have given renewed impetus to the display of university performance. Shin and Toutkoushian (2011) have observed the expansion of the higher education market raising the issue of HEIs quality in the 1980s, and suggested it could be an external boost to university evaluation, similar as ranking systems. They have commented as following:

With the rapid growth of higher education markets, policymakers and employers began to raise the issue of quality in the 1980s. Elite universities soon began to compete with each other to attract better qualified students and attract financial resources from donors. The general public also began to be interested in the activities and accomplishments of universities and how they compared to each other. This societal interest led to the emergence of ranking systems for higher education. (p. 3)

It is important to keep in mind that the internationalization of HEIs promotes university rankings, extending from domestic level to global level, as well as making ranking activities public topics. Buela-Casal et al. (2007) have noted that it is not enough any more for universities to only know their rankings compared to peers at

national scale since higher education has developed so internationally and "as universities increasingly compete in a global environment, they tend to compare themselves with world universities. In fact, the expression 'World Class' has been created and many universities expect being considered as 'World-Class Universities'" (p. 350).

Additionally, Shin and Toukoushian (2011) have pointed out that the increasing interest in egalitarianism in higher education could be another reason for the rapid growth of university rankings. They have explained that:

> The concept of egalitarianism in higher education competes with the elitism ideal and argues that higher education should focus on providing services to the general population, as well as the elites. This ideal emerged in the late nineteenth century in the United States with the rapid massification of higher education. Since then, different types of higher education institutions such as community colleges in the United States, polytechs in Europe, and other types of two-year institutions have emerged in many countries as a means of increasing egalitarianism. (p. 2–3)

On the other side, there is no doubt that higher education exports a large quantity of advanced technology and ideology and fosters a great variety of talents every year, which is a strong driving-force for social development. So it is necessary to establish a set of reasonable supervision and restriction to ensure efficient and normal operation despite the university being an open and highly independent higher institution with freedom. As noted by Brown (as cited in Federkeil, 2008, p. 219), "evaluation, assessment and assurance of academic quality is intrinsic to higher education". Moreover, Federkeil has pointed out that the European higher education systems enjoy a high degree of self-governance and freedom, which makes it necessary to involve in an audit and accountability system concerning administration and management practice. A similar point was raised by Shin and Toutkoushian (2011), arguing that policymakers started establishing quality assurance schemes and tried to solve the quality issues of higher education. Shin and Toutkoushian (2011)

have noted that another effort contributing to accountability "came from public sector management with the liberal governments in the United States and the UK in the 1980s developing public accountability systems. This required higher education institutions to report their performance to policymakers according to predetermined performance indicators" (p. 3).

In support of this viewpoint, it could be found that there are a minimum of three different mechanisms co-existing in the HEIs currently, namely "rankings developed mainly by the media, quality assurance measures created by quality assurance agencies, and accountability measures imposed by governments" (Shin & Toutkoushian, 2011, p. 3).

WURs are interesting because higher education is an important vehicle for the rearrangement of social goods and various stake-holders are eager to get an instrument to meet their requirement of updated information in higher education.

High school students and their parents are concerned about university performance and try to get a reliable assessment of national and/or international universities and this is one kind of "public demand for transparency and information that institutions and government have not been able to meet on their own" (Usher & Savino, 2006, p. 38). Elite universities compete with each other to attract more talented students and obtain more financial resources from donors (Shin & Toukoushian, 2011, p. 3). Further, such ranking activities have made university resources and achievements public information. The masses can approach higher education and exploit rich resources in ways that have never happened before.

Last but not least, at the technical level, well-developed information processing techniques have created a possibility for ranking universities in practice. As seen by Boulton (2011), the emerging international rankings of universities were probably triggered by the new communications technologies driving liberalization of international markets and the development of information and knowledge based global economy during the 1990s. With hindsight, university ranking can be regarded as one sort of product of higher education and knowledge-based social development. Sadlak and Liu (2007) have noted as well that

"rankings can be conceived as an imperative of the knowledge society" (as cited in Federkeil, 2008, p. 223).

2.2 The understanding of university ranking

Conceptions of university ranking are a theoretical basis for analyzing and exploit ranking system. So it is necessary to discuss the concepts of university ranking prior to expounding rankings' functions and impacts in this study. The concept will be discussed with the comparative point of view.

2.2.1 Conceptual issues

A. Comparing ranking with evaluation

In considering the comparison of ranking and evaluation, Huang (2011) has found that these two notions are different but associated with their purpose and outcome and stated:

1. Evaluation is not equal to ranking.

2. Evaluation sets a benchmark against which a university performance in certain aspects can be assessed. The goal is to determine if a university passes the assessment, meaning it has achieved at or surpassed a basic level of requirements. Evaluation results do not have to be quantitative. Descriptive evaluation contents in some evaluation contexts and some evaluation results indicate simply final decisions such as pass or fail to pass.

3. Ranking, on the other hand, sorts a group of universities by numerical indicators. Ranking shows a university's relative strength and weakness as compared to its peer institutions in the areas represented by the indicators.

4. Ranking is an efficient, convenient, and easily under-standable evaluation method, even though some have argued about the fairness of quantitative comparisons of universities where each university is unique and differs to the others in some aspects. (p. 5)

It is clear that both evaluation and ranking are for appraising university performance, however, they have different performance measures and targets. The former focuses on basic level of quality requirements for universities, while the latter one focuses on appraising university excellence. Taylor and Braddock (2007) also believe that "a genuine ranking system is one that sets out specifically to measure the excellence of universities" (p. 246).

In addition, rankings entirely neglect reasons for bad performance of HEIs, while examining the reasons for bad performance is exactly the mission of evaluation (Federkeil, 2008). To clarify these two notions, a diagrammatic comparison is shown in **Fig. 1**.
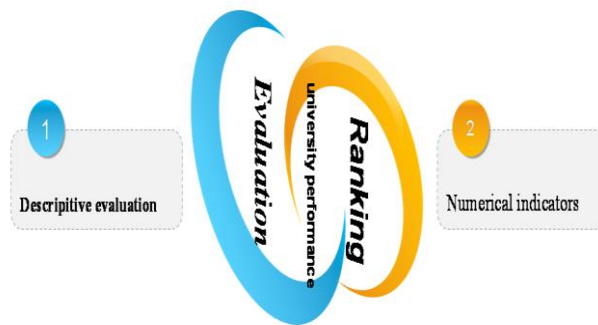
Figure 1 Schematic comparison between evaluation and ranking

Based on these perceptions and the fact that one of the key issues in this study is trying to identify the principles of an outstanding university, the concepts of ranking and evaluation need to be further developed in the context of university appraising mechanisms. Precisely, evaluation primarily focuses on a basic level of quality requirements for the universities with qualitative methods in broad sense, while the notion of ranking primarily focuses on appraising excellence within elite universities by quantitative methods although descriptive or qualitative methods have been utilized in the process of making ranking as well.

B. Comparing ranking with classification

To differentiate ranking from classification, van Vught and Westerheijden (2010) have asserted that classifications allocating objects to groups on the basis of their characteristics only show

horizontal diversity without implying ordinal scales of 'more', 'bigger' or 'better' while rankings display vertical diversity in terms of performance by using various criteria.

Clearly, classification and ranking play different roles in the domain of HEIs at present. The former one emphasizes on grouping and summarizing, while the latter one leads to competition. Further, it could be a new direction and development for university ranking that the ranking makers, e.g. U-multrirank ranking project bring the view of classification into ranking systems.

C. Quality assurance, accountability and ranking

As we quoted earlier, Shin and Toutkoushian (2011) have concluded that "quality assurance", "accountability" and "ranking" are three relative and well-known mechanisms for HEIs' quality measurement currently (p. 3). They have explained the concept of quality assurance and accountability as below:

> Quality assurance refers to national and institutional systems designed to assess and improve the quality of teaching and research, and provide relevant information to key stakeholders on academic standards and employment of graduates, while accountability refers to 'rendering an account' about what an institution is doing in relation to goals that have been set, or legitimate expectations that others may have of one's services or processes, in terms that can be understood by those who have a need or right to understand 'the account'. (p. 36–37)

In respect of ranking systems, Shin and Toutkoushian (2011) suggested that ranking concentrates on accountability and ranking order in order to use public taxes legitimately, while quality assurance aims to enhance institutional quality as stated by institutional mission (p. 25–26).

It is significant that these three mechanisms are different in contrast to their targets, policy links, publishing of results and method of evaluation, although they have much in common since they "provide information to the public" and "enhance institutional quality" (Shin & Toutkoushian 2011, p. 25). As seen in

Table 1, Shin and Toutkoushian gave a summary of comparisons between quality assurance, ranking and accountability (p. 26).

A significant point to draw from the above brief summary or comparisons between quality assurance, ranking and accountability is that ranking provides information to stakeholders, in relation to the quality of HEIs.

Table 1 Comparisons between quality assurance, ranking, and accountability by Shin and Toutkoushian

Characteristics	Quality assurance	Ranking	Accountability
Goals	Enhancing quality	Information providing	Financial accountability
Stakeholder	Government/HEIs/agency	Media/researchinstitute	Government/fundingagency
Actions	Accreditation Quality assessment Quality audit Program review Licensure	Ranking by institution Ranking by region or disciplines	Performance reporting Performance-funding/budg eting
Indicators	Teaching/research/service	Research/teaching/repu tation/internationalizati on	Teaching/research/ service
Data sources	Nationwide data Peer review/survey	Nationwide data Peer review/survey	Nationwide data
Linking with government policy	Institution's legal status Financial aids Research funding Operational funding	Not linking Some developing countries link with policy	Linking or not linking with funding
Customers	HEIs, government	Parents, students, HEIs, enterprise, government	Government

2.2.2 University ranking construction in theory

This part will focus on the issues of indicators and the type of measurement, i.e. methodologies used by ranking systems, with an emphasis on the noteworthy traits of the four selected rankings in the study. It's important to locate suitable indicators or parameter setting before data collection. It is commonly accepted that ranking project design should be considered as the first step in the procedures of ranking systems.

A. Methodologies issues – indicators and weightings

Huang (2012) has suggested that the indicators and measurement play an essential role in international university rankings (p. 2).

Indeed, identifying performance indicators and assigning corresponding weightings are important parameters. Indicators and weightings for ranking criteria are subject to the goals and the types of specific rankings. Or rather, the two segments, i.e. indicators and weightings are associated with other components closely so that examining all key contents can help us comprehend the ranking systems wholly.

The concept of performance indicators is characterized by scholars as "performance indicators designed to provide quantifiable measurements which, having been agreed in advance, reflect the relative success of an organization" (Longden, 2008, p. 79). Validity, reliability and attributes of performance indicators have been discussed in the university ranking research, as they to some extent work upon the attainment of original objectives, as well as raising the controversy over the value of ranking as discussed in general previously.

There are general guidelines behind the Berlin Principle from Williams (2008), in which compilers are supposed to "use outputs rather than inputs, be transparent, use verifiable data and recognize diversity of missions" (p. 52). As evidenced by this statement, Williams (2008) has summarized four attributes which should be stressed in ranking university performance:

a) Research output and its influence

b) The quality of teaching and research training

c) The contribution to the formulation and implementation of national policy

d) Ratings should be undertaken separately for the different attributes before they are combined into a single measure. (p. 52)

Based on above statements, it could be argued that:

1. Choosing indicators should be objective and manifold

2. The main dimensions of performance indicators are teaching and research.

Here, one should be somewhat concerned about the multidimensional feature of performance indicators in U-Multirank project, which are not directly tied to education and research performance - knowledge transfer, regional engagement and international orientation are combined into indicators and proxies to evaluate university performance. Additionally, stakeholder consultation and expert advice are involved in the process of selecting proxies and indicators.

To clarify the measurement spectrum, Williams (2008) put forward a general framework in regard to the performance indicators currently used in ranking systems. He has commented:

> The methods used to measure research performance in universities form a spectrum: from a survey of peers at one end to the use of quantitative measures of output only, such as publications and citations, at the other end. In the middle of the spectrum lies evaluation obtained by providing peers with representative publications and detailed quantitative information. In evaluating the quality of teaching the methodology spectrum ranges from surveys of students and employers to quantitative measures such as progression rates, job placements and starting salaries of graduates. (p. 52)

Williams (2008) has figured out two major measurement means for current methodologies in ranking systems, in which he suggests using publications and related citations in measuring research performance, and peer review in measuring the quality of teaching and learning (p. 52). Aguillo, Bar-Ilan, Levene and Ortega (2010) has also commented on the spectrum of performance indicators in the actual setting and stated that "the different rankings consider different parameters including publication and citation counts, student/faculty ratio, percentage of international students, Nobel and other prizes, number of highly cited researchers and papers, articles published in *Science* and *Nature*, the h-index and web visibility" (p. 244).

For the measurement spectrum discussed above, publication count and peer review are two most common measuring means used to evaluate university performance. One should be reminded

that the new ranking project U-Multirank has tried to rank university quality with a multiple perspective, and therefore attempted to build up a set of more comprehensive indicators and proxies to meet the need of diverse development in HEIs.

Moreover, parameters in different ranking systems are all in relation to some specific components. It has been pointed out by the Expert Group on Assessment of University-Based Research that "indicators must be fit for purpose and verifiable" (European Commission, 2010, p. 12).

The strengths of the indicators used in assessment exercises have been discussed by many researchers. European Commission (2010) report has provided a comprehensive overview of indicators which are most commonly used "relating each indicator to the measurement of a specific aspect or dimension of research, pointing out its strengths and weaknesses, and indicating some further development which should be undertaken to make the indicator in question more robust" (p. 13).

These indicators make important contributions to measuring university excellence. However, some questions were brought by the use of indicators within assessment exercises. Three serious weaknesses were assembled by European Commission (2010):

1. There is no single set of indicators capable of capturing the complexity of research and research assessment.

2. There is no such thing as a perfect indicator; all indicators have their own specific strengths and weaknesses, and assessment exercises have to take this into consideration from the outset.

3. There is no such thing as an objective indicator: Indicators are rarely a direct measurement; more often than not, they are proxies. (p. 12)

It is worth thinking whether or not such indicators of ranking systems could measure HEIs' quality accurately and comprehensively. Charon and Wauters (2008) have elaborated that "even with the best possible indicators, the quality of an institution may not only depend on academic and research performance, but that the quality of education, library,

administration, peculiar regional or national missions, campus culture and quality of life should also, whenever possible, be considered" (p. 64). It could be argued that only limited indicators or proxies cannot display university performance objectively and comprehensively, which is the root cause for the structural defect of university ranking.

To some extent, performance indicators and the related weightings are combined, and utilized in ranking systems. European Commission (2010) has reported that "the choice, interpretation and weighting of indicators are of utmost importance in any assessment exercise or system" (p. 13). In relation to the translation of indicator values into scores or visual results, it can be argued that the weighting of indicators needs to be calculated effectively at a technical level, and to be able to evaluate university performance with a statistical position, attempting to expose the mystification of university performance. However, none of rankers explains what their weights represent, despite use of weights in all published ranking lists (Tofallis, 2012, p. 5).

In conclusion, proxies and indicators in current rankings are not directly correlated with HEIs' achievements but could be used as a new angle to recognize university performance. It's necessary to stress this viewpoint since this study aims to investigate university excellence in the context of internationalization by means of abundant information from university rankings. Indicators and overall scores of different rankings need to be compared and analyzed carefully.

B. The key steps of producing university ranking

Deconstruction of the process of producing university ranking apparently is an important issue in the field of university ranking research. This is a necessary and pragmatic way to help us to realize what data are valuable.

Longden (2011) has claimed that the possibility exists that there are certain key processes adopted by all compilers when we rethink the stages necessary in creating ranking. Longden has argued that it is possible to find some points where potential difficulties happen if we use critical steps of producing ranking as a guide to break down ranking systems. These vital steps could be identified and analyzed in detail in such a way that it could

facilitate an overhaul of ranking systems from a systematic and critical standpoint.

According to Longden (2011), the process of producing university rankings involves:

1. Clarifying reason for ranking

2. Selecting suitable metrics - performance indicators (PI)

3. Collecting data - metrics

4. Adaptation of PIs into a scale

5. Standardizing measures prior to aggregating

6. Weighting PIs prior to aggregating

7. Creating a single index reflecting a university. (p. 78–100)

Merisotis and Sadlak (as cited in Huang, 2012) have figured out that ranking processes are roughly outlined as three steps:

1. Data collection including existing data and recently updated compiled data

2. Selecting ranking type and variables

3. Identifying indicators and weightings and then executing analysis. (p. 72)

Among these procedures, Merisotis and Sadlak (as cited in Huang, 2011) alleged that determining the performance indicators and the weightings of indicators were the predominant key factors in the whole process of producing university ranking. In addition to performance indicators, Geuna and Martin (as cited in Huang, 2011) have figured out that "bibliometric and peer review were the two predominant methods of academic evaluations, with bibliometric being quantitative evaluation and peer review being qualitative evaluation" (p. 2). There have been both objective and subjective approaches in university evaluation for a long time (Huang, 2011, p. 3).

Tofallis (2012) has written that "all currently published university rankings combine various measures to produce an overall score using an additive approach" (p. 1). He suggested that the existing normalization procedures lead to contradictory results when applied to same data. Tofallis (2012) has also found that compilers of league tables have switched from one normalization to another occasionally and the choice of normalization clearly makes a substantial difference. Hence, Tofallis has proposed a multiplicative approach as an alternative aggregation scheme to overcome the difficulties associated with the additive approach based on the following three steps:

1. Normalize the data

2. Attach weights to the criteria

3. Add together the weighted values to produce an overall score. (p. 3)

To normalize the data, Tofallis has further identified various ways to make the magnitudes of the values similar across criteria, including:

1. Dividing by the largest value

2. Range normalization

3. z-scores (statistical standardization)

4. Dividing by the sum. (p. 3)

There are two key and closely related issues addressed by compilers for ranking procedures: the one is selecting proxies and indicators, and collecting required data; and the other one is analyzing and standardizing data, namely executing analysis.

2.2.3 The types of university ranking

It is important to recognize differing ranking types involved in the whole procedures of producing ranking.

Dill, Soo, Usher and Savino (as cited in Federkril, 2008) has suggested that "rankings vary in their aims and target groups as well as in terms of what they measure, how they measure it and how they implicitly define quality" (p. 223). Van Vught and Westerheijden (2010) have maintained that rankings could be clarified into five dimensions as below:

1. Primary target groups

2. Producers: public vs. private and not-for-profit vs. private for-profit

3. Level: institutional vs. field-based

4. Scope: national vs. international

5. Focus: education vs. research. (p. 8-9)

Van Vught and Westerheijden (2010) have explained how to distinguish between institutional-level rankings and field-based / field-level rankings and claimed that "institution-based" normally refers to "smaller organizational units, like faculties, schools or departments focusing on a single area of knowledge", while "field-based" probably refers to "academic disciplines like economics and physics, interdisciplinary areas like business studies and nanotechnology and single study programmes or research programmes in a given area" (p. 12).

Further, van Vught and Westerheijden (2010) have reported that field-level rankings are particularly attractive to students or individual researchers since "programmes across institutions may have quite different qualities. Indicators only showing averages for whole institutions mask particularly strong or weak programmes, implying that for these users institutional rankings are irrelevant or even misleading" (p. 12). In this sense, field-based and field-level rankings bring us a depth of judging university performance.

In addition, van Vught and Westerheijden (2010) have stated that institutional-level rankings are mainly "popular with government policy makers and institutional leaders who have a legitimate interest in overall characteristics at the institutional level" (p. 12). From this point of view, it could be concluded that

institutional level rankings are usually geared to policy makers' needs to assess HEIs performance on the whole.

Notwithstanding the fact that traits have been identified for different types of university rankings, categorizing existing university rankings might be impractical. Actually, many university rankings probably do not belong to any single type of university rankings since they are comprehensive. This phenomenon implies that one specific existing university ranking may include many traits of different types of ranking systems mentioned above. For example, ARWU has targeted comparing "the world's top research universities" (Rauhvargers, 2011, p. 24). It is obvious that ARWU is not only a research-based activity, but also an international level project.

There is a broad discussion in terms of national rankings and international rankings. Van Vught and Westerheijden (2010) have figured out two major development trends: the one is "national rankings are expanding to neighboring areas", e.g. CHE ranking includes HEIs from Netherlands, Switzerland, and inter alia Austria now; and the other one is "more focused international rankings are beginning to emerge", e.g. CHE ranking focuses on the European market master and Ph. D. students now, but this ranking only covers a limited number of fields and universities are restricted to international, research-oriented scope across the European countries (p. 13).

2.2.4 Brief summary

A. The concept of ranking

It is interesting to note that none of the quite accurate definitions of university ranking explicitly addresses all of the features which are involved in the variety of university rankings, although many researchers try to define it. Lazaridis (2010) has suggested one typical definition for ranking and pointed out that "ranking a university gives an overall picture of its quality. However, many universities are quite heterogeneous, containing excellent as well as mediocre departments. University assessment fails to give proper credit to these pockets of excellence" (p. 212).

The above viewpoint is too general to reflect the particularity of ranking sufficiently. Indeed, university rankings are full of challenges since each university or institution has its own parti-cular features influenced by the social and cultural environment of

their own country (Shin & Toutkoushian, 2011, p. 2). However, this definition overlooks the original orientation of rankings for providing information, as well as overemphasizing its evaluation function. Moreover, there are more than 17000 universities around the world, but most of WURs concentrate on the top class universities (top 200 or 500) which are a tiny proportion of total HEIs (Rauhvargers, 2011). WURs are designed to explore university excellence within distinguished universities, which agrees with the original motivation of WUR to a certain extent.

Given above the *caveats*, it is believed that rankings aim to provide information of university performance for the stakeholders and compare strengths and weaknesses of elite universities around the world by employing numerical indicators and succinct presentation forms. Although accurate definition of ranking is not existed, researchers and scholars have tried varied approaches to develop it.

B. The types of sample rankings in this study

In line with above statements about the types of ranking, a quick glance of the university rankings sampled in this study are shown in **Table 2**. It appears all rankings aim to rank worldwide universities with a global scope. Meanwhile, "research" is the focus of attention in all of these university rankings although some also involve "education", "innovation", "community outreach", and so on. The importance of research in ranking systems implies the social demand for science and technology development in our times.

It should be noted that the institutional ranking of the U-Multirank project enables a comparison of HEIs on the basis of single dimension of institutional activity, such as education, or internationalization of knowledge transfer, but the scores in these different dimensions will not be combined into an overall score. Field-based ranking will be designed as "a multi-dimensional ranking of a set of study programmes in a specific field or discipline" (Rauhvargers, 2011, p. 55-56).

Table 2 Categories of university rankings

	Target groups	Ranking Producers	Level	Scope	Focus
ARWU	Top research university (1200)	Public, non-profit	Field-based & Institutional-level	Global scale	Research
THE	Top university	Private, for-profit (TSL Education Ltd)	Field-based & Institutional-level	Global scale	Research & education
QS-USNWR	Top university	Private, for-profit	Field-based & Institutional-level	Global scale	Research & education
U-Multirank	Reflecting institutional diversity & balance regional and national issues	Public, non-profit (European Commission, CHERPA network led by the Centre for Higher Education Policy Studies at Twente University, the Netherlands & the Zentrum für Hochschulentwicklung (CHE), Germany	Field-based & Institutional-level	Global scale	Research, education, innovation, employability, internationalization and community outreach

2.3   he impacts of university ranking

University ranking has spread extensively all over the world. In view of its advantages and disadvantages, opinions vary. However, it is undoubted that ranking activities will keep going for a long period, and stakeholders are likely to be influenced by ranking activities anyway whatever support or objection they raise. As Hazelkorn (2008) believed that no matter what influence or impact, positive or negative, pros and cons, "the growing body of academic research and journalist reportage is referenced to contextualize this international experience" (p. 193). Hence, facing and dealing with university ranking is a relatively practical issue. The following section will embark on a review of value and function to contribute to an understanding of the university ranking.

2.3.1   The value and function of WUR

Huang (2011) has offered an overview of the theoretical conclusion of the value and function of university ranking. He has stated several advantages of university ranking. First, it is easy to

compare the performance of universities being assessed for observers; second, it explicitly demonstrates a university's relative attainment in some aspects and can help to diagnose problems as well as provide developmental directions in the near future; third, ranking meets the requirement of public access to university performance information and; finally, carefully designed and systematic quantitative data based rankings provides policymakers with objective information (p. 5).

These statements imply three different functions within ranking systems, which are commonly discussed in the existing literatures, i.e. consumer information, assessment instrument, and stimulating competition (**Fig. 2**).
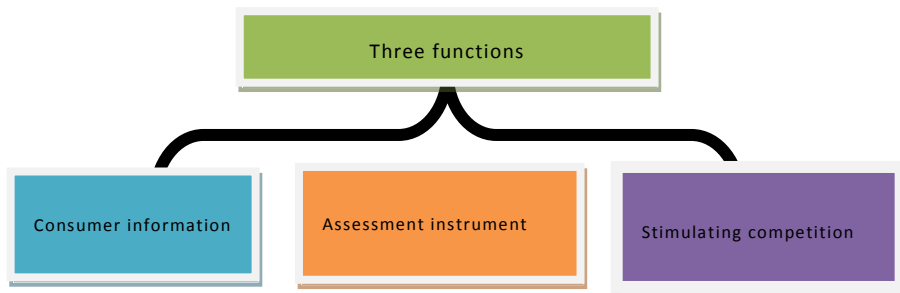
Figure 2 Overview of major functions of university ranking

A. Consumer information.

In realizing the practical value, Buela-Casal et al. (2007) have ensured that university rankings have grown fast in different countries due to "the double purpose of giving information to consumers and functioning as an institutional marketing strategy" (p. 350).

It could be argued that ranking is important and actually valuable as consumer information for multi-stakeholders. In regard to the so-called "consumer information", Boulton (2011) has indicated how rankings influence different stakeholders and summarized many hotly debated problems as below:

> If rankings could be created that accurately reflected the diverse values of universities to their societies, they would in theory be valuable in several ways: to university managers in benchmarking their universities against an international scale and identifying where improvements should be sought; to students and

academics in matching their choices of where to study or work to their aspirations; to public and private bodies seeking links to universities that would further their objectives; and to governments in helping to align their policies for universities to national needs. (p. 76)

Similarly, Hazelkorn (2008) has concluded that four chief target groups think about the league tables:

1. 'Users of the System' which contains students, parents, employers and government

2. HEIs trying to 'Best the System' by re-presenting/configuring their data in the most favorable way or otherwise attempting to influence the input metrics

3. Groups trying to 'Better the System', such as ranking organizations/consortia, governments and supra-governmental organizations, and academics

4. 'Critics of the System', which includes elements of all the above. (p. 194)

As consumer information, readers are reminded that limitations may exist in reality. Boulton (2011) has pointed out that "the delivery of such benefits of ranking system is conditional however on the capacity of rankings to measure the values of universities to their societies" (p. 76) and commented that it is risky to use ranking system with poor measures since they deliver more damage than benefit. Boulton has figured out two major problems existing in ranking systems:

1. That many of the features they seek to measure cannot be measured directly, but depend for their evaluation on indirect proxies; leaving the question of how good are the proxies?

2. That universities now vary greatly in the diverse functions that they are called upon to perform in society, and how therefore can a single, monotonic scale be an accurate measure of institutions that have different roles? (p. 76)

B. Assessment instrument

University rankings are not designed as an "instrument of internal quality assurance" within HEIs, but furnish HEIs with an instrument of external assessment (Federkeil, 2008).

More specifically, Federkeil has contended that rankings as well as league tables are an instrument to generate transparency in a so-called "university jungle" (p. 222). It means that rankings help to support objectivity and credibility in evaluating university performance, and also generate transparency to the public. In short, rankings are an external assessment of university performance and they enable HEIs to reach a high degree of transparency in a competitive context for internal and external stakeholders as Federkeil indicated.

This comes as no surprise given the reliability of assessment in ranking systems. One aspect of reliability of assessment has been examined extensively in the literature is "who are ranking universities"? It is true that "who ranks universities" is a key issue of relevance or ties to the university rankings' validity. A significant point is that university rankings are normally conducted by governments, academics, or eminent magazines, and newspapers. Buela-Casal et al. (2007) have pointed out that most of rankings have been conducted by some private and media-based bodies. As a consequence, the raters could be objective rather than subjective in the process of compiling information and judging university performance, even though there are still some suspicions in the eyes of critics.

As one example, in order to eschew university-provided data and expert reviews, ARWU only uses openly available information, like the number of articles published in *Nature*&*Science*, the number of Nobel Prizes / Field Medals won within faculty and alumni, the number of highly cited researchers in broad categories (Enserink, 2007). In this context, published results are more objective and reliant, at least for the assessment of research dimension, in response to some critics who considered that ARWU exclusively set research performance as benchmarking. It also means that "rankings can help consumers see the value of their investment in higher education and hold institutions accountable for results" (Shin & Toutkoushian, 2011, p. 4).

C. Competition

University ranking has led to sets of systematic rules to judge the quality of collegiate education. It has not only created a visual display of outstanding universities but also promoted academic, scientific and educational competition among universities on the global-scale.

Huang (2012) has pointed out that "the waves of globalization encourage the competition among universities on a global basis" (p. 1). A general impression of ranking, as such, implies higher than something, lower than something or equal to something. "Ranking" could usually be regarded as that somebody or something has a high position in a group or organization. Obviously, a bitter contest among different universities has been initiated by university ranking, and most universities are eager to hold a top position in the consequent league table.

Aguillo et al. (2010) have indicated that the remarkable success of university rankings is due to globalization of the higher education since worldwide universities may compete for human and economic resources. To stay competitive, "higher education institutions are using these rankings as a promotion tool that shows their educational, research or business excellence" (Aguillo et al. 2010, p. 244). University ranking has therefore drawn increasing attention from the public. That is to say, ranking actually help to meet the demand of HEIs internationalization and also stimulate universities' competition and development.

However, one should be aware that such competition stirred by ranking systems can have pernicious effects as well. For example, rankings are somehow in favor of steep and stratified institutional patterns which could undermine horizontal diversity of universities (Teichler, 2011). As Teichler (2011) suggested that problems with rankings were:

1. Encouraging resources at a few outstanding institutions

2. Making newcomers synonymous with losers

3. Reinforcing mechanisms whereby status breeds status; there is the "Matthew effect" in resource allocation, and "reputational recycling"

4. (4) Undermining the pride of institutions which are not top institutions. (p. 66)

As indicated above, there are many problems accompanying university rankings and it becomes an important segment of ranking research. The following part is a broad review of shortcomings of university rankings. While this overview is not exhaustive, it does generalize major aspects which are discussed widely within existing literatures and allow a comprehensive understanding of university ranking for scholars.

2.3.2 The overview of defects

A. The bias within criteria

Boulton (2011) has thought the benefits from university rankings could be achieved through accuracy without doubting the authenticity of a university rating system. That is to say, all these strengths of university rankings are based on the premise that ranking systems can evaluate university quality if a set of reliable criteria and assessment methodologies are provided despite rich diversity among universities and more than one form of excellence. So it comes as no surprise that there are many limitations within rankings systems and we believe that it is worth discussing before looking into three sample rankings. A general overview of shortcomings was formulated as following.

First of all, reliability of methodologies and validity of indicators are two chief issues of criteria facing university rankings and also cause fierce controversy (Huang, 2011). Particularly, the former requires that a rating system can "generate consistent results in replication", while the latter requires that "indicators represent the evaluation criteria and whether the evaluation is properly conducted" (Huang, 2011, p. 6). Huang has claimed that issues of validity are becoming further complicated within multi-dimensional university rankings.

Moreover, as mentioned before, the public requirement of open information concerning university performance makes raters more objective in the process of compiling information and judging university performance. But the problem is that not all the university performance data are open to the public, and some

universities can even reject providing evaluators with requested data (Huang, 2011).

Secondly, university ranking is criticized for overlooking the existing diversity of HEIs, especially overemphasizing research performance with a focus on natural science. Teichler (2011) has summarized the bias systematically with regard to this issue:

1. Rankings miss (or negatively assess) HEIs with other functions than those of research-oriented universities

2. Rankings undermine horizontal diversity and there is a disregard for diverse missions and fitness for purpose

3. Rankings disregard/disadvantage small institutions

4. Rankings discriminate against humanities and social sciences

5. Rankings reinforce dominant paradigms, thereby controlling the choice of theories, methods, and themes

6. Rankings do not sufficiently strike a balance between teaching and research, but often only infer that good research produces good teaching, that input in teaching and research leads to good processes, and that this in turn leads to good outcomes. Available research has often challenged these assumptions and experience has shown that the most successful institutions may not have the best practice. (p. 64)

According to Enserink (2007), the phenomenon of putting an emphasis on research outcomes could be observed within most WURs, partly because WURs are aimed at providing straightforward information for policymakers and also "because education systems and cultural contexts are so vastly different from country to country that solid and meaningful data are hard to come by" (p. 1027). Meanwhile, university performance was influenced by the vast social environments like socio-cultural and politico-economic context actually. Thereby, it is hard to make a fairly unbiased evaluation of university quality even if the relative data are available. Huang (2011) has also stated that it is inappropriate that all universities are subject to the same criteria.

Presumably, it is unreasonable to judge university performance with the same criteria. Considering the diversity, some rating systems like U-Multirank, attempt to evaluate a university with multiple indicators. However, multiple criteria cannot solve the intrinsic problem although this could help to ease it to some extent. Boulton (2011) has pointed out that ranking methodologies have an inherent flaw and structural defect. For example, van Raan, van Leeuwen and Visser (2011) argued that non-English papers decreased rankings and this language effect happened in important ratings by THE, QS, ARWU and the Leiden Ranking. Further, organizing and analyzing university data on a global scale is such a complicated process that reliability and validation have become a key issue and multiple indicators have been popularized (Huang, 2011).

Last but not at least, mobility and reputation are two tricky issues. For instance, Enserink stated that Andrew Fire's 2006 Nobel Prize in Physiology or medicine helped his recent institute, Stanford University in Palo Alto, California, move to second level at ARWU 2006, but actually Fire did his breakthrough work in RNA interference at the Carnegie Institution of Technology in Baltimore, Maryland (Enserink, 2007).

To sum up, ranking systems do not take into account that universities are quite disparate organizations. For example, one of the key factors of university development intermingled within the process of social development is the opportunities to study and explore that universities bring to the students.

B. Challenges within the processes of producing university rankings

Correspondingly, it is well worth investigating the deficiencies of ranking systems. This part will contribute awareness of the limitations of this study.

Longden (2011) has identified some possible challenges in ranking systems by breaking down university rankings into specific procedures:

a) For the reason of creating ranking, altruism is unlikely to be the reason why a publishing company is engaged in providing university ranking, while advertising revenue, purchases of the final edited ranking book and other forms

of endorsement are real reasons for companies to promote the ranking related marketing. In such cases, profit targeted rankings published by companies every year have to be reevaluated carefully.

b) The performance Indicator (PI) has helped to form the visual landscape of HEIs, providing a critical measure to help recognize university quality, and also service a need of HEI development. However, PIs can only be derived from things over which direct control can be exerted leading to achieving an outcome of the ideal measure. Any PI with an emphasis on "easy to collect" thus has inherent limits due to its form as a single index.

c) Concerning data collection, there are three types of data available in the ranking systems, namely: primary data generated by the university itself, survey data generated by the compilers, and data collected from independent third parties. It is still a knotty problem how to obtain more reliable data from the university, unbiased data from surveys and authoritative and comparative data from independent third parties.

d) The construction of the ranking scale is arbitrary and assumed linear without any theoretical analysis when adapting PIs data into a scale.

e) Adding scores from different sources together is simple but it is important to treat different groups of data sets by stretching them so that they conform to common statistical measures when standardizing measures. So it is necessary to check if the compilers apply the process of standardization in preparation of their rankings.

f) In weighting PIs, there is no agreement of the relative contribution of the measures while different measures are added together to generate a single index. It is difficult to determine which contributes more to the overall measure of the university. The weightings of performance proxies and indicators adopted by compilers are idiosyncratic and devoid of a theoretical underpinning.

g) In the final step creating a single index reflecting a university, the significance of the difference in scores is difficult to judge because heavy clustering makes a small difference in scores translate into a substantial difference in ranking so that it remains unclear what is implied behind rankings, for example, the meaningful and useful information behind rank #1 and rank #10.

Further, the rankings are volatile for some universities between years and such instable ranking systems create much debate in academia. The report (CHERI, 2008) to Higher Education Funding Council for England has noted that "compilers are not always clear about their methods for standardizing the individual variables, despite this potentially having a major impact on the rankings" (p. 21).

2.4   Dissecting four selected WURs

To get an in depth understanding of WURs, it is necessary to examine the background of *ARWU*, *QS-USNWR, THE* and *U-multirank* as well as their methodologies.

2.4.1   Three selected single dimensional WURs

According to Rauhvargers (2011), *ARWU* ranking dated from 1998 when *Shanghai JiaoTong University (SJTU)* was selected as one of nine universities in the "985" project by the Chinese government (p. 24). The "985 project" was established in response to a statement by Jiang Zeming (then president of China) on May 4, 1998, and he emphasized that "university should play a critical role in implementing the strategy of invigorating the country through science, technology and education" and "China should have several world-class universities of international standard" (Rauhvargers, p. 24).

At present, *ARWU* actually ranks more than 1000 universities in which the best 500 are published on its website: http://www.shang hairanking.com/.

ARWU adopts six objective indicators for university ranking (**Table 3**), including "the number of alumni and staff winning Nobel Prizes and Fields Medals", "number of highly cited researchers selected by Thomson Scientific", "number of articles published in journals of *Nature* and *Science*", "number of articles indexed in

Science Citation Index - Expanded and Social Sciences Citation Index", and "per capita performance with respect to the size of an institution"

Table 3 Main indicators and weighting indicator in case *ARWU**
           (2011-2012, published for 2012)

Indicators	Weight
1. Highly cited researchers in broad categories	20%
2. Articles published in Nature & Science	20%
3. Articles in Science/Social Science Citation Index	20%
4. Faculty with Nobel Prizes/Field Medals	20%
5. Alumni with Nobel Prizes/Field Medals	10%
6. Per capita academic performance of an institution	10%
Total	100%

*Data cited from http://www.shanghairanking.com/, cited on 10th August, 2012

ARWU has attracted increasing attention from governments, universities and public media, and a survey on higher education published by *The Economist* in 2005 named *ARWU* as 'the most widely used annual ranking of the world's research universities ("A world of opportunity", 2005). Also *ARWU* (produced by Shanghai Jiaotong University) first put forward the "league tables" of university excellence in 2003 and Boulton (2011) pointed out "its advent was perhaps the inevitable consequence of the convergence during the 1990s of liberalisation of international markets, enabled by new communications technologies, and the shift of the global economy towards one based on information and knowledge" (p. 75).

The *THE-QS* University Ranking were published from 2004 to 2009 by the *Times Higher Education Supplement* (*THES*), a weekly British magazine located in London, and *Quacquarelli Symonds Limited*, a company with its head office in North London which specializes in education and study abroad. The partnership of *THES* and *Quacquarelli Symonds Limited* broke up officially in 2010. *Quacquarelli Symonds* continues to use the original ranking methodology and publishes the "*QS World University Rankings*".

Meanwhile, *US News and World Report* (*USNWR*) has started to work with *QS* to publish the best universities worldwide.

USNWR is actually an American news magazine located in Washington, D. C. with a long history. It has been a leader in news for many years along with *Time* (*TIME*, an American news magazine) and *Newsweek* (an American news magazine published in New York). It has also become particularly known for its publication since 1983 of the ranking and annual reports on US universities emphasizing peer review among universities.

USNWR publishes its *US News and World Report's World's Best Universities* based on the *QS* World University Rankings (**Table 4**). *Quacquarelli Symonds* continues to use the original ranking methodology and publishes the "*QS World University Rankings*", while the *Times Higher Education* (*THE*, formerly *THES*) and *Thomson Reuters* have built up a partnership with a new methodology to publish a new world university ranking named as "*Times Higher Education World University Rankings*" (as shown in **Table 5**). Academic peer review and employer review, which are determined by email questionnaire or global survey, amount up to 50% weight in *QS-USNWR* (**Table 4)**. The teaching and research in *THE* methodology are partly determined by subjective measures such as reputation surveys (**Table 5**).

Table 4 Main indicators and weighting indicator in case *QS-USNWR*
         (2011-2012, published for 2012)

Indicators	Weight
1. Academic peer review (email questionnaire)	40%
2. Citations per academic	20%
3. Faculty student ratio	20%
4. Proportion of international academic staff	5%
5. Proportion of international students	5%
6. Employer review (global online survey)	10%
Total	100%

*Data cited from http://www.topuniversities.com, cited on 10th August, 2012

Table 5 Main indicators and weighting indicator in case *THE* (2011-2012, published for 2012)

Indicators	Weight
1. Teaching: the learning environment	30%
2. Research: volume, income and reputation	30%
3. Citations: research influence	32.5%
4. Industry income: innovation	2.5%
5. International mix: staff and students	5%
Total	100%

*Data cited from: http://www.timeshighereducation.co.uk/world-university-rankings/,cited on 10th August, 2012

2.4.2 U-Multirank project

Numerous scholars have debated on issues in university rankings such as "what indicators can accurately measure quality", "which methodology is more useful for the development of ranking systems", "how presenting the information in a ranking-format to increase the transparency of these measures" and so on (Buela-Casal et al., 2007, p. 351).

As mentioned before, the *European Multidimensional University Ranking System* (*U-Multirank*), an EU-found project initiating to recognize and rank university diversely, is becoming widely popular within HEIs. Compared with single dimension ranking activities, *U-Multirank project* aims to create global university rankings with a multidimensional perspective, covering the various missions of HEIs, such as research, education, innovation, employability, internationalization and community outreach (Rauhvargers, 2011).

There are some features of multi-dimensional approaches that are often highlighted within *U-Multirank* research. EU Commission (as cited in Rauhvargers, 2009) has reported that the *U-Multirank* approach is multi-dimensional rather than single-dimensional since it tries to cover all kinds of study fields and various missions of institutions (dimensions) properly and emphasized that "the existing rankings tend to focus on research in 'hard sciences' and ignore the performance of universities in areas like humanities and

social sciences, and aspects such as teaching quality and community outreach" (p. 56).

Van Vught and Westerheijden (2010) have given a further argument allowing for a more explicit comparison between single-dimensional and multi-dimensional approaches. They have stated:

> ➢ Most existing higher education rankings take the form of a league table, a single-dimensional list going from "best" to "worst", assigning ordinal numbers to the entities which relate only to rank and not scales of difference. Other approaches to ranking use:

> ➢ • Multi-dimensional approaches, which do not try to combine education and research rankings, for example, into a single, composite measure and which are often user-driven because they enable an interactive display of data; and/or

> ➢ • Robust group ratings rather than individual rankings, such as in league tables. (p. 5)

U-Multirank heavily emphasizes the inclusiveness for identifying scope. Van Vught and Ziegele (2011) have suggested that "*U-Multirank* must be open to recognized higher education institutions of all types and from all participating countries, irrespective of their membership in associations, networks or conferences" (p. 161).

Correspondingly, *U-Multirank* attempts to map and rank university performance at a global level and with a participative approach. "The *U-Multirank* will cover institutions inside and outside Europe, in particular those in the US, Asia and Australia" (EU Commission, as cited in Rauhvargers, 2011, p. 55). A participative approach requires the ranking compilers to take stakeholders' input seriously including suggestions for design and feedback for results (van Vught, Westerheijden, & Ziegele, 2012, p. 6).

Moreover, *U-Multirank* tried to profile university performance in visual styles. The results at a glance were presented in **Fig. 3** (Federkeil, et al. 2012, p. 173). It has been pointed out that "graphic presentations may help to convey insights into the institutional results 'at a glance' with the performance of the institution as a whole presented without being aggregated into one composite indicator" (Federkeil, et al. 2012, p. 173).

However, it is important that the university samples of the pilot study "reflect as much institutional diversity as possible; and making sure that the sample was regionally and nationally balanced (van Vught & Ziegele, 2011, p. 97). According to the viewpoints of van Vught and Ziegele, *U-Multirank* project focuses on producing institutional ranking and a field-based ranking in business studies and two fields of engineering, i.e. electrical and mechanical.
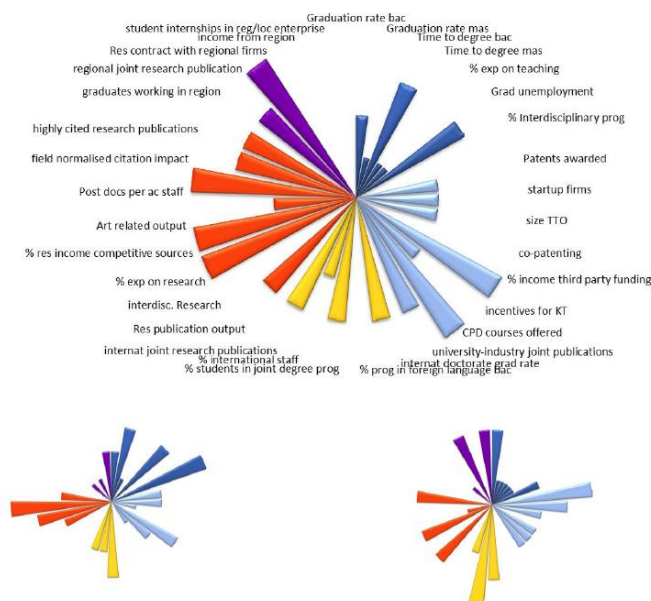
Figure 3 Institutional sunburst charts by *U-Multirank*

2.5 Concluding remarks

The notion underpinning the stated discourse is that university ranking should be considered as information source for all related stakeholders as well as the general public. Ranking activities will undoubtedly keep going, and stakeholders would be continuously influenced by such activities although it is still a long way to develop unbiased rankings in face of globalization. It is meaningful to determine the underlying factors of university excellence in the new era; while global university rankings could serve as an effective tool to research global universities at top tier.

3 Research methodology

3.1 3.1 Objective of this research

This study aims to determine the relationships and similarities among international university rankings by analyzing four well known WURs, i.e. *THE*, *ARWU*, *QS-USNWR*, and *U-Multrirank*, and then to try to discuss and understand university excellence as well as explore the traits of world-class university. The key research question is what the salient features of top university are in our time. Two specific questions were raised:

1. What is the relationship between selected WURs?

2. What characteristics of top universities can be derived from WURs?

3.2 Research methods and design

The first step in this study was to analyze the results of WURs and methodological issues underlying it. The next step was to discuss the methodologies of selected WURs and further explore the traits of outstanding universities. In particular, the first step was to attest the effectiveness of WUR as a tool for studying university excellence, as well as provide theoretical underpinnings for this study; the second step was to try to find the essence of extraordinary universities on top of the ranking list and draw a profile of university excellence.

Keeping research questions and goals in mind, multiple-case studies were designed and initiated. Comparative approach involving both qualitative and quantitative methods was used.

Cohen, Manion and Morrison (as cited in Brötzmann, 2010) have suggested that case study is usually applied to generalize conclusions from one case or more cases to other similar cases by dissecting a specific example or more examples of something deeply (p. 16). Case study method covers both single and multiple-case studies (Yin, 2009, p. 19). By definition, case study method is an appropriate way to recognize university ranking in depth. Specifically, this study utilized

case study methodology with multiple cases to extract a single set of "cross-case" conclusions (Yin, 2009, p. 20).

The emphasis of this study is put on "understanding and comparison" of representative WURs, which benefits further discussions on university excellence. Therefore, comparison approach involving quantitative and qualitative methods was used in the processes of analysis. With regard to the comparative approach of a multiple-case study, it is supposed to perform "intensive analysis of a few cases rather than more superficial statistic analysis of many cases" (Enli, 2010, p. 4).

Yin (2009) has mapped a pathway that depicts important logics of multi-case study, as shown in Fig. 4 This model was utilized for conducting this study.
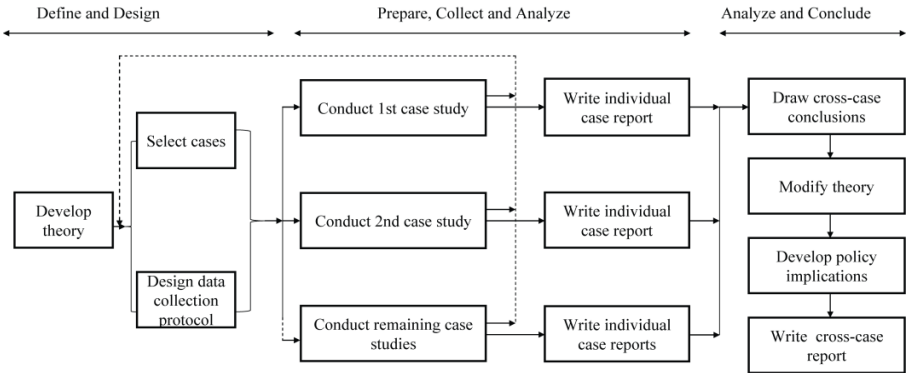
Figure 4 Case study method (Yin, 2009, p. 57)

Accordingly, the basic framework of research methodology in this study was created by combining the model of multiple-case study with actual research context. This study is framed schematically in **Fig.5** Four representative WURs, i.e., *Times Higher Education (THE),Academic Ranking of World Universities (ARWU), QS* together with *US News and World Report (QS-USNWR)* and *U-Multriank* project, were selected as four cases to discuss university excellence. Then the top 10 and the top 50 universities ranked by *THE, ARWU* and *QS-USNWR* were extracted and agreement among rankings of top-tier universities was evaluated by both qualitative and quantitative comparisons. Meanwhile, methodologies adopted by these rankings were also assessed. Based on these findings and summary, traits of top-tier universities were demonstrated and discussed.

University excellence

WURs

Selected cases
(ARWU,THE,
USNWR, U-Multriank)

Results

Extracting top universities

Methodology

Overlapping indicators

Findings

Summary

Discussion & conclusion

Figure 5 A framework of the research design

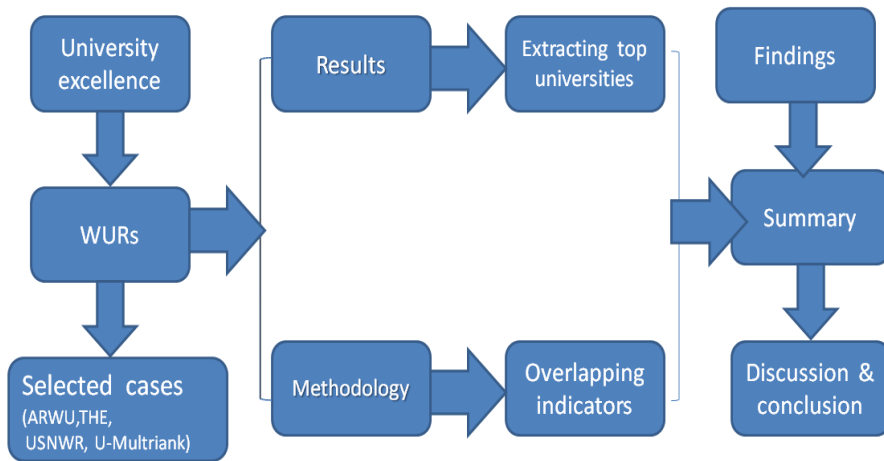## 3.3 Selection of WURs

As a general guideline, identification of the units, i.e. the research cases, is related to the original research questions (Yin, 2009, p. 30). Yin further indicates that "each case must be carefully selected so that it either can predict similar results (a literal replication) or predicts contrasting results but for anticipatable reasons (a theoretical replication)" (p. 54).

As stated in Literature Reviewsection, *ARWU* is the ancestor of global university ranking; *U-Multirank* represents cutting-edge development in the field of WURs; *THE* and *QS-USNWR* have been popular in many countries and experienced changes with regard to their methodologies because of re-organization. Most important of all, all of these rankings devote to comparing top universities in terms of the strengths and weakness all over the world. *ARWU*, *THE*, *QS-USNWR* and *U-Multirank* are the suitable and representative research cases in this study. We therefore elected these four WURs for discussing the traits of outstanding university by deciphering results and methodologies of WURs.

3.4 Data collection

The data was mainly obtained through reviewing existing information available to the public. Yin believes that "this type of information can take many forms and should be the object of explicit data collection plans" (2009, p. 101). In this study, the information related to WURs

was derived from websites of university rankings, academic books, reports and journals. Specifically, methodologies and league tables of selected rankings were obtained from official websites of rankings and only the data of rankings for 2011-2012 were collected and used for in depth analysis.

3.5 Data analysis

To perform data analysis, both qualitative and quantitative approaches were utilized in this study. For qualitative analysis, raw datasets were grouped and categorized and key features were indentified through formatted tables and figures. For quantitative analysis, relationship between WURs was evaluated by calculating correlation coefficients with some software, e.g. SPSS, Origin and Excel. For instance, nonparametric correlations, i.e. Spearman coefficients were computed between the ranking results of different WURs. The data processes also took used comparative and graphical approaches to visualize the data.

3.6 Research limitations

There are some limitations in the research due to the data availability restrictions. The information concerning WURs only relies on the secondary data source. Some original data for making university rankings are not available at present. In addition, only four WURs are considered in this study, so that the collected data of rankings are limited to some extent.

4 Findings

University ranking is explicitly manifesting university performance in a combination of parameters (academic achievement, reputation, etc.). Both the implemented indicators and the presentations of results construct the essential underpinnings of university ranking, as well as the key used for exploring university excellence in this study. This section strives to find out some latent rules of WURs, on the basis of reviewing ranking results and related methodologies. Since the *U-Multirank* project is still in process and no results get published, only three representative WURs will be analyzed in depth.

4.1 Extracting top universities from *THE, ARWU and QS-USNWR*

Most WURs put forward a list of top universities year by year according to compiling criteria. Top 50 universities in 2011-2012 of *THE, ARWU and QS-USNWR* are shown in **Table 6**.

Apparently, there is a similarity between these ranking results or league tables. Top 10 universities were quite similar within all of the league tables while some disagreements were observed by comparing top 50 universities in one ranking with another.

Table 6 Top 50 universities in *THE, ARWU* and *QS-USNWR* (2011-2012, published for 2012)

Ranking	ARWU	QS-USNWR	THE
1	Harvard U.	U. Cambridge	Caltech
2	Stanford U.	Harvard U.	Harvard U.
3	MIT	MIT	Stanford U. (2)
4	UC-Berkeley	Yale U.	U. Oxford
5	U. Cambridge	U. Oxford	Princeton U.
6	Caltech	Imperial College London	U. Cambridge
7	Princeton U.	University College London	MIT
8	Columbia U.	U. Chicago	Imperial College London
9	U. Chicago	U. Pennsylvania	U. Chicago

10	U. Oxford	Columbia U.	UC-Berkeley
11	Yale U.	Stanford U.	Yale U.
12	UC-Los Angeles	Caltech	Columbia U.
13	Cornell U.	Princeton U.	UC-Los Angeles
14	U. Pennsylvania	U. Michigan	Johns Hopkins U.
15	*UC-San Diego*	Cornell U.	ETH Zurich
16	*U. Washington*	Johns Hopkins U.	U. Pennsylvania
17	UC-San Francisco	*McGill University*	University College London
18	Johns Hopkins U.	ETH Zurich	U. Michigan
19	U. Wisconsin-Madison	Duke U.	U. Toronto
20	University College London	*U. Edinburgh*	Cornell U.
21	The University of Tokyo	UC-Berkeley	Carnegie Mellon U.
22	U. Michigan-Ann Arbor	*U. Hong Kong*	*U. British Columbia*
23	ETH Zurich (Swiss Federal Institute of Technology)	U. Toronto	Duke U. (22)
24	The Imperial College of Science, Technology and Medicine	Northwestern U.	Georgia Institute of Technology
25	*U. Illinois at Urbana-Champaign*	The University of Tokyo	*U. Washington*
26	U. Toronto	Australian National U.	Northwestern U.
27	Kyoto U.	Kings College London (U. London)	U. Wisconsin-Madison
28	*U. Minnesota, Twin Cities*	National University of Singapore	*McGill University*
29	New York U.	The University of Manchester	*University of Texas at Austin*
30	Northwestern U.	U. Bristol	University of Tokyo
31	*Washington U. in St. Louis.*	The University of Melbourne	*U. Illinois at Urbana-Champaign*
32	U. Colorado at Boulder	Kyoto U.	*Karolinska Institute*
33	Rockefeller U.	Ecole Normale Superieure, Paris	*UC-San Diego*
34	*UC-Santa Barbara*	UC-Los Angeles	*U. Hong Kong*
35	Duke U.	*Ecole Polytechnique Fédérale de Lausanne*	*UC-Santa Barbara*
36	*The University of Texas at Austin* (35)	Ecole Polytechnique	*U. Edinburgh*
37	*U. British Columbia*	The Chinese University of Hong Kong	U. Melbourne

38	**U. Manchester**	The University of Sydney	Australian National U.
39	U. Maryland, College Park (38)	Brown University	*UC-Davis*
40	U. Paris Sud (Paris 11)	The Hong Kong University of Science and Technology	National University of Singapore
41	Pierre and Marie Curie U.-Paris 6	**U. Wisconsin-Madison**	*Washington U. in St Louis*
42	*U. North Carolina at Chapel Hill*	Seoul National U.	*U. Minnesota*
43	U. Copenhagen	Carnegie Mellon U.	*U. North Carolina at Chapel Hill*
44	*Karolinska Institute*	**New York U.**	**New York U.**
45	Pennsylvania State U.-University Park	Osaka U.	Ludwig-Maximilians-Universität München
46	U. Southern California	*Peking U.*	*École Polytechnique Fédérale de Lausanne*
47	Technical University Munich	Tsinghua U. (46)	London School of Economics and Political Science
48	*UC-Davis*	The University of Queensland (46)	**U. Manchester**
49	UC-Irvine (48)	The University of New South Wales	Brown U.
50	Utrecht U. (48)	The University of Warwick	**Peking U. (49)**

*Data were cited from the official websites of ARWU, QS and THE (cited on 10th August, 2012). Universities listed in all three rankings are in bold face; while universities listed in any two university rankings are in italics. Same universities may be named different in rankings, for example, The Imperial College of Science, Technology and Medicine in ARWU is named asImperial College London in QS-USNWR and THE.

It is somewhat surprising how similar all the three university rankings are especially there is a similarity between the *ARWU* and *THE* rankings for the top 10 universities. It is as high as 90% (9/10) (**Fig. 6**). But the orders of top 10 universities are divergent between the league tables of *ARWU* and *THE* (**Fig. 6**), i.e., the same universities are in different positions at the two stated ranking lists. Further comparison between *ARWU* and *THE* was performed by mapping top 50 universities in their league tables (**Fig. 6**). Specifically, top 50 universities of rankings *THE* and *ARWU* published for 2012 were 74% or 37/50 in common. These results indicated good agreement between rankings *THE* and *ARWU*. However, only three universities take the same positions in *THE* as in *ARWU* (**Fig. 6**), e.g. *Stanford U.*, *Yale U.*, and *U. Chicago*.
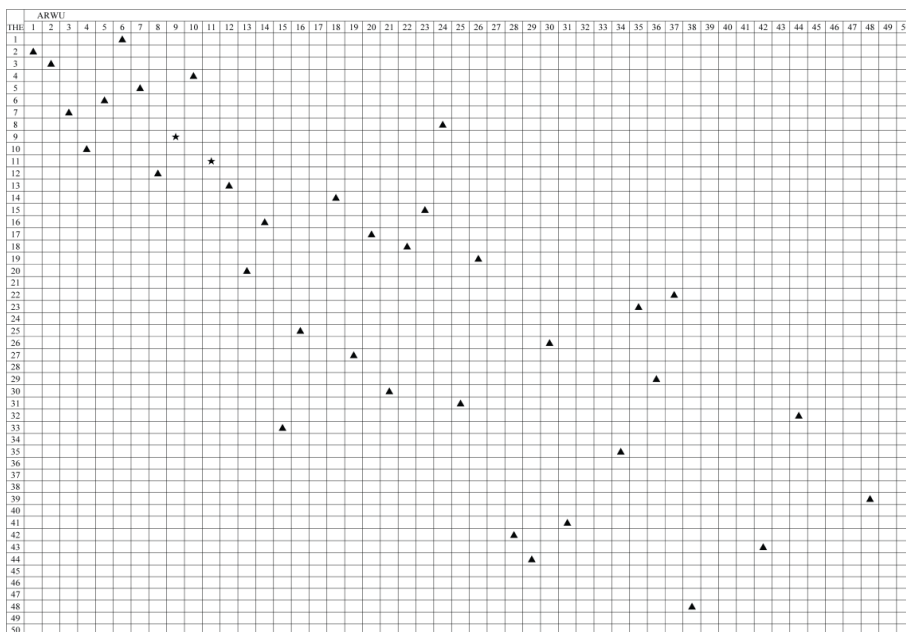
*The five-pointed star represents perfect agreement while the triangle represents good agreement between rankings THE and ARWU on top universities. Some universities may be the same orders in the rankings but numbered differently for mapping, for example, UC Davis, UC-Irvine and Utrecht U. were all the 48th in ARWU, but UC-Irvine and Utrecht U. were numbered as the 49th and the 50th.

Figure 6 Mapping* of top 50 universities within rankings *THE* and *ARWU*

Six universities are shared by *ARWU* and *QS-USNWR* as well as by *THE* and *QS-USNWR*, indicating a similarity of 60% (6/10) for top 10 universities. The lower similarities indicate some disagreement between rankings *THE* and *QS-USNWR* and between *ARWU* and *QS-USNWR* based on their ranking outcomes of top 10 universities.

Further, all these rankings *ARWU*, *THE* and *QS-USNWR* share five universities in their top 10 lists and it means they are in good agreement on these universities, *i.e. Harvard U., MIT, U. Cambridge, U. Chicago, and U. Oxford* though their orders are varying among league tables. In addition, *University College London, Yale U. and U. Pennsylvania* are only ranked in *QS-USNWR* as top 10 universities.

In order to carry out an in-depth analysis, top 25 universities published for 2012 are extracted from *ARWU*, *QS-USNWR* and *THE*. Those 25 universities are then compared to top 50 universities in the other two. Similarities between rankings were evaluated by

pairing 25 universities in one ranking with the top 50 universities in the other two according to **Table 6**. The results were summarized in **Fig. 7**.
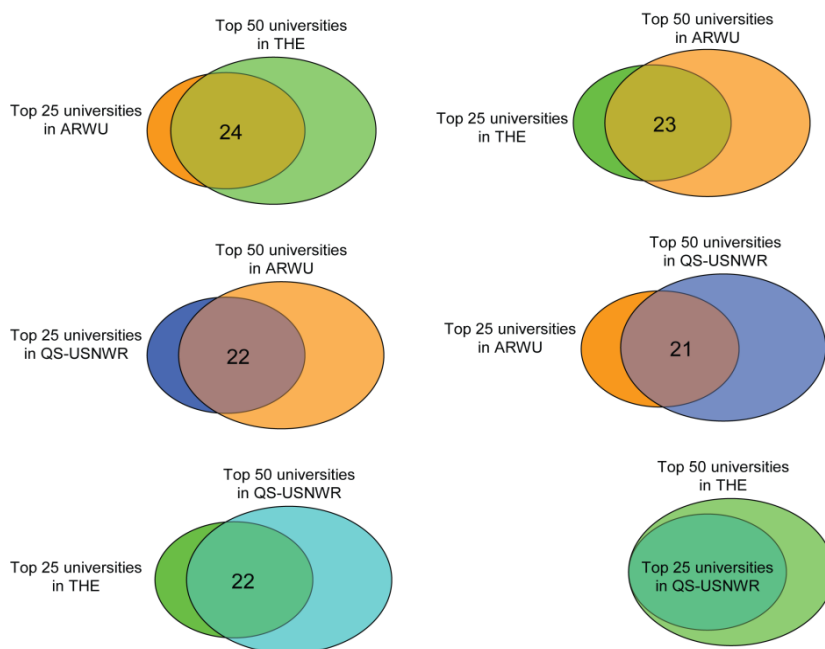
*Comparing top 25 universities in one ranking with top 50 universities in the other two.

Figure 7 Comparisons* between league tables of *ARWU*, *QS-USNWR* and *THE*

Outcomes of *ARWU* and *THE* were first examined. 23 of top 25 universities in *THE* were ranked as top 50 universities in *ARWU*. Only the 21st (*Carnegie Mellon U.*), and the 24th (*Georgia Institute of Technology*) universities in *THE* were not in the top 50 in *ARWU*. In a similar vein, 24 of top 25 universities in *ARWU* were ranked as top 50 universities in *THE*. Only the 17th (*UC-San Francisco*) university in *ARWU* was not in top 50 in *THE.*

Outcomes of *ARWU* and *QS-USNWR* were also examined. For *QS-USNWR*, there are 22 of top 25 universities were ranked as top 50 universities in *ARWU*, and the 17th (*McGill University*), the 20th (*U. Edinburgh*) and the 22nd (*U. Hong Kong*) were out of top 50 in *ARWU*. For *ARWU*, there are 21 of top 25 universities were ranked as top 50 universities in *QS-USNWR*, and the 15th (*UC-San Diego*), the 16th (*U. Washington*), the 17th *(UC-San Francisco*) and the 25th (*U. Illinois at Urbana-Champaign*) were not in top 50 in *QS-USNWR*.

Finally, outcomes of *THE* and *QS-USNWR* were examined. For *THE*, there are 22 of top 25 universities were ranked as top 50 universities in *QS-USNWR*, and the 22nd (*U. British Columbia*), the 24th (*Georgia Institute of Technology*) and the 25th (*U. Washington*) were out of top 50 in *QS-USNWR*. For *QS-USNWR*, in contrast, all the top 25 universities were ranked as top 50 universities in *THE*.

For quantitative analysis, top 25 universities in one ranking were mapped to the other two and Spearman correlation coefficients were computed between any two rankings. The Spearman correlation coefficients (*r*) between any two rankings were shown in **Table 7**. In most cases, top 25 universities in one ranking list were paired to their counterparts in any other two ranking lists. However, there were two exceptions. *University of Hong Kong*, one of top 25 universities in *QS-USNWR* was excluded in correlation computation for ARWU/QS-USNWR since it was an outlier which was listed as the 250th in ARWU. *University of California, San Francisco* one of top 25 universities in *ARWU* was excluded in correlation computation for QS-USNWR/ARWU and THE/ARWU since it was an outlier which was out of full ranking lists of *QS-USNWR* and *THE*.

Table 7 The Spearman correlation coefficients between any two rankings of *THE*, *QS-USNWR* and *ARWU*.

X axis	Y axis	Spearman correlation coefficient (*r*)	Sig. (2-tailed)	N
THE	ARWU	0.82	<0.001	25
	QS-USNWR	0.72	<0.001	25
QS-USNWR	THE	0.72	<0.001	25
	ARWU	0.59	0.003	24[ζ]
ARWU[§]	QS-USNWR	0.49	0.015	24[§]
	THE	0.75	<0.001	24[§]

[ζ]*University of Hong Kong* (one of top 25 universities in QS-USNWR but listed as the 250th in ARWU) was excluded in correlation computation.

[§]*University of California,* San Francisco (one of top 25 universities in ARWU but out of full lists of QS-USNWR and THE) was excluded in correlation computation.
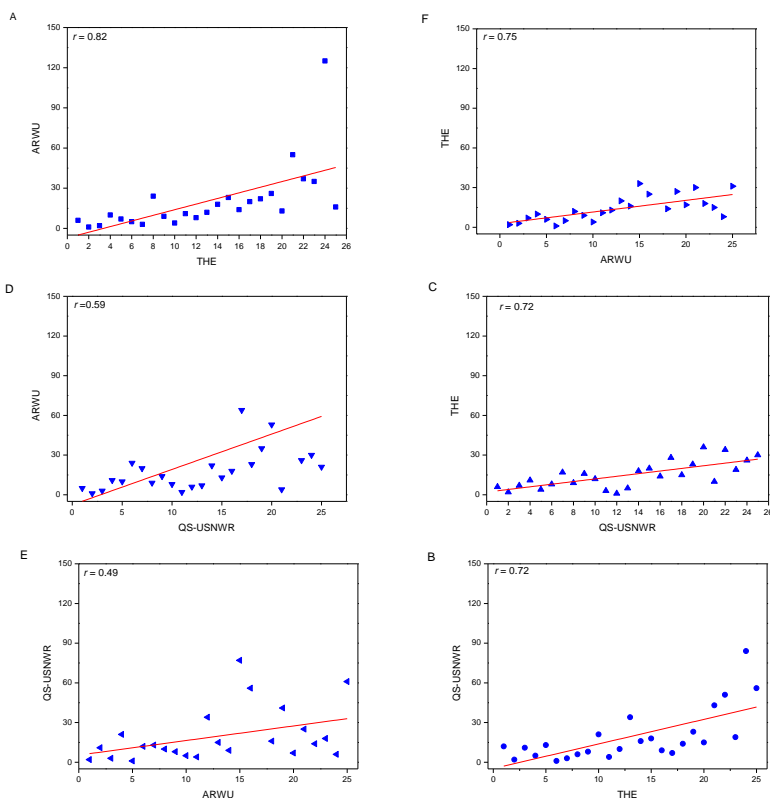
Top 25 universities in *THE* and their corresponding ranks in *ARWU* (A) and *QS-USNWR* (B), respectively. Top 25 universities ranked in *QS-USNWR* and their corresponding orders in *THE* (C) and *ARWU* (D), respectively. Top 25 universities ranked in *ARWU* and their corresponding orders in rankings *QS-USNWR* (E) and *THE* (F), respectively. Universities which were ranked in one ranking but out of ranking range of other rankings were excluded, for example, *University of California, San Francisco* which was ranked as 17th in *ARWU* was not listed in *QS-USNWR* and *THE.*

Figure 8 Analysis of correlations between rankings

The correlation between *THE* and *ARWU* was highest. The correlation coefficient (r_s) was as high as 0.82 (**Fig. 8 A**) by matching the top 25 universities in *THE* with top universities in *ARWU* while the correlation coefficient (r_s) was decreased to 0.75 by matching the 25 universities in *ARWU* with the top universities in *THE* (**Fig. 8 F**). The relationship between *QS-USNW*R and *ARWU* was complicated. The obvious difference between two correlation coefficients (r_s), i.e. 0.59 for ARWU/QS-USNWR and 0.49 for QS-USNWR/ARWU attested that *QS-USNWR* was somehow less related to *ARWU* based on their ranking outcomes (**Fig. 8 D** and **Fig. 8 E**). But the difference between correlations could also be

caused by the fact that *University of California, San Francisco* listed as the 17th in *ARWU* was out of the top university list of *THE* and *QS-USNWR* However, the highly similar coefficients (r_s), i.e. 0.72 for both QS-USNWR/THE and THE/QS-USNWR indicated good positive correlation between rankings *THE* and *QS-USNWR* (**Fig. 8 B** and **Fig. 8 C**).

4.2 Deciphering ranking methodologies

This part is about ranking methodology issues in WUR. It concentrates on addressing the question that what are the significant traits of WUR methodologies. The methodologies of four WURs selected in the study were studied in depth.

Methodology as a principal component is a big challenge in making university rankings. It's actually a core part of university ranking that arouses intensive debates. University rating systems are making some changes both nuanced and sophisticated in response to intense critiques every year. *ARWU* tries to avoid all of these problems by eschewing university-provided data and expert reviews. Instead, it uses only publicly available data. The methodology for ranking is growing increasingly complex and comprehensive. Overall, the ranking methodology including its indicators and related weights is switching to be more transparent.

Ranking methodologies were assessed in terms of their indicators and weights as shown in **Table 8**. There is an awareness of the fact that *ARWU* ranks worldwide universities based on only publicly available and objective data while both *QS-USNWR* and *THE* include reputation as a component of indicators by questionnaires or surveys.

It remains unclear if reputation should be taken into account when ranking universities in the scope of academic and/or comprehensively. *ARWU* completely excludes reputation while *THE* weights reputation including both teaching and research as 33% of total ranking scores (**Table 8**), however, there is positive correlation (r_s = 0.82 for ARWU/THE and 0.75 for THE/ARWU) between the two stated rankings on top 25 universities world widely (**Fig.8**). Another question is that what percentage of reputation conducted by questionnaires or surveys should be worth in the total ranking score. Actually, there is a big difference between weights of reputation in *QS-USNWR* and *THE* (**Table 8**) though both of the rankings emphasize

the importance of reputation when ranking universities. Finally, if reputation is supposed to be included as a component of indicators for ranking scores, the questionnaires and surveys are not quite appropriate to assess university reputation though these are existing approaches in practice.

Table 8 Evaluation of ranking methodologies in terms of their indicators and weights

	Indicators and weights	Objectiveness* of approaches	Impact of reputation	Data source and reliability	Limitations
THE	Specified	Medium (67%)	Medium	Publicly available and private questionnaires; medium	Ranking score is affected by both teaching and research reputation surveys significantly
ARWU	Specified	High (100%)	Low	Publicly available; high	Ranking score is affected by publications in Nature & Science significantly
QS-USNWR	Specified	Low (50%)	High	Publicly available and private questionnaires; medium	Ranking score is highly dependent on academic peer review and employer review significantly
U-Multirank	In development; complicated and multi-dimensional				

*Objectiveness refers to the percentage of weights based on objective data.

It is surprising that different rankings come to similar consequences, e.g. similar top 25 universities in league tables. However, it is clear that *QS-USNWR*, *THE* and *ARWU* share some common indicators by dissecting their ranking methodologies. For better clarity, an overview of different methodologies adopted by three selected single-dimensional rankings was figured out as shown in **Fig.9**. Indicators referring to teaching and research are quite similar in *QS-USNWR*, *THE* and *ARWU* although some indicators have different weights in these rankings.
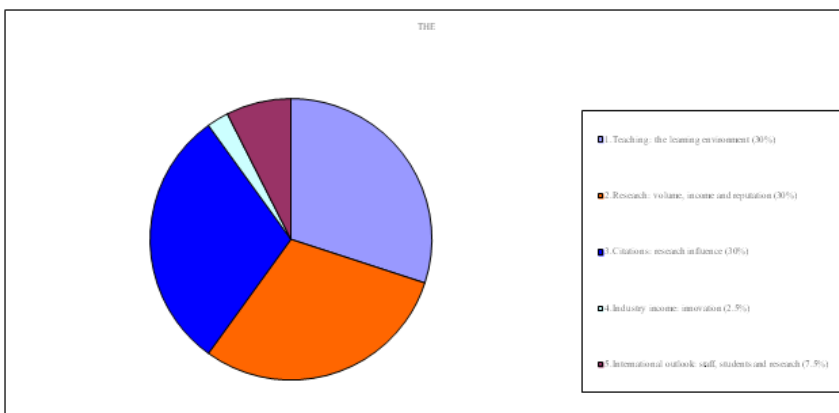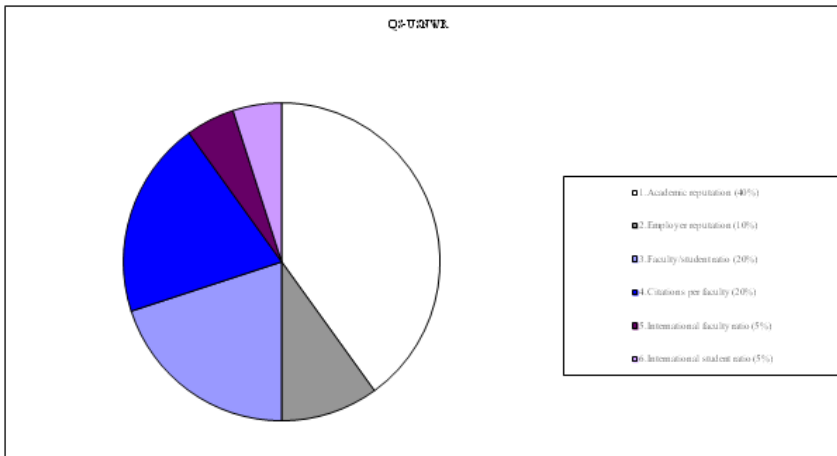
ARWU

1.Alumni of an institution winning Nobel Prizes and Fields Medals (10%)

2.Staff of an institution winning Nobel Prizes and Fields Medals (20%)

3.Highly cited researchers in 21 broad subject categories (20%)

4.Papers published in Nature and Science (20%)

5.Papers indexed in Science Citation Index-expanded and Social Science Citation Index (20%)

6.Per capita academic performance of an institution (10%)

QS-USNWR

1.Academic reputation (40%)

2.Employer reputation (10%)

3.Faculty/student ratio (20%)

4.Citations per faculty (20%)

5.International faculty ratio (5%)

6.International student ratio (5%)

THE

1.Teaching: the learning environment (30%)

2.Research: volume, income and reputation (30%)

3.Citations: research influence (30%)

4.Industry income: innovation (2.5%)

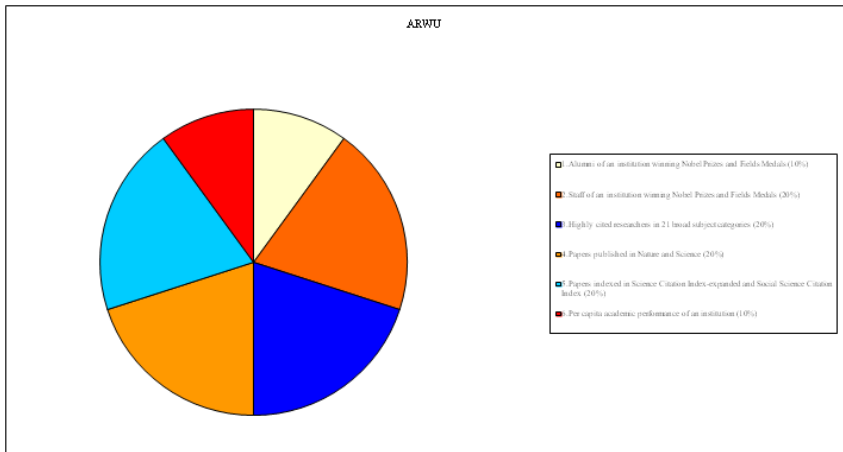5.International outlook: staff, students and research (7.5%)

Figure 9 Visualizing ndicatorsand their weightings in *ARWU*, *QS-USNWR* and *THE*

5 Discussion

University ranking system and league table have been part of higher education for decades. This study provides us results for look into global university ranking systems from a perspective of university orientation towards excellence.

5.1 Overlapping between university rankings

University rankings provide us the opportunity to approach world-wide universities although much debate is created among the public in respect to their league tables. The main similarities and difference among these methodologies of three selected rankings were summarized and shown in **Fig. 10**.
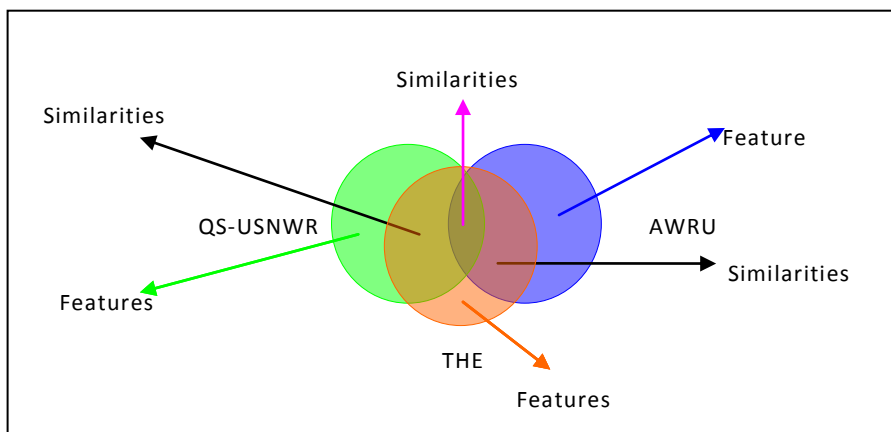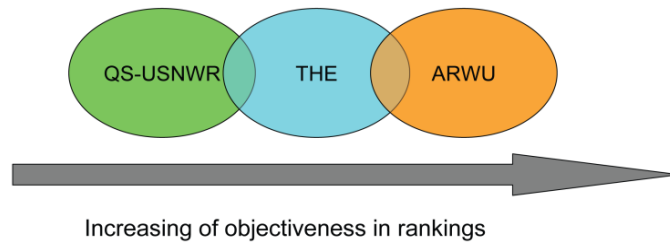

Details were shown in Fig..9.

**Figure 10 Similarities and difference between ranking methodologies**

It was found that any two rankings, e.g. *QS-USNWR* and *ARWU*, *QS-USNWR* and *THE*, and *ARWU* and *THE*, had some indicators in common. Meanwhile, any of *QS-USNWR* and *THE*, and *ARWU* had its own feature, i.e. unique indicators. In addition, same or similar indicators were adopted by all these three rankings.

In addition, ranking in *QS-USNWR* is the least related to objective data/performance (objectiveness) but in *ARWU* is the most related to objectiveness, while ranking in *THE* is moderately related to objectiveness (**Fig. 11A**). Further, it appears that the correlation between any two stated rankings depends on the difference between their objectiveness (**Fig. 12B**).
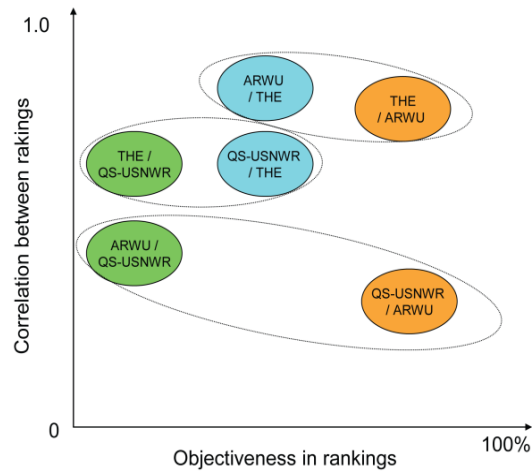
Increasing objectiveness in rakings (A) and relationship between ranking correlation and ranking objectiveness (B). Correlation and objectiveness plotting was based on Table 8.

Figure 11 Overview of interplay among selected rankings

All of these three rankings consider objective performance of universities as a determinant in generating league tables. In other words, the objectiveness of any rankings, i.e. weights based on objective measures, is ≥ 50% (**Table 8**). Objective data utilized by rankings are more reliable and repeatable. Since these rankings adopted some same or similar indicators which were reflected by objective data, the positive correlations between rankings could be connected to their objectivenesses (objective data used for rankings). Further, pure reputation survey for ranking universities has been largely abandoned even by the National Research Council (NRC) of the United States (Mervis, 2010) though it's well known that the first assessment of U.S. research doctorate programs was

done by NRC in 1982 and then repeated in 1995 by ranking departments based on their reputations (Mervis, 2003). One can also conclude that the most controversial but popular ranking exercise should not be dependent on reputations *per se* since reputations suffer from many structural flaws, such as time lag, a halo effect and uncertainty of measuring (Mervis, 2010).

5.2   Characterizing world-class universities

To characterize top universities, the top universities in selected rankings will be categorized into region, economy, and language related groups. It's clear that the rankings *ARWU*, *THE* and *QS-USNWR* have a lot in common based on their league tables for top 25 universities (**Table 6 and Table 9**).

Table 9 Categorizing top 25 universities into region, economy, and language related groups

	Geographic continent			Economic development		Official language	
	Europe	North America	Other continents	Developed* countries/regions	Developing countries/regions	English	Non-English
ARWU	4	19	2	25	0	23	1 (Japanese); 1 (trilingual)
THE	5	20	0	25	0	24	1 (trilingual)
QS-US NWR	6	17	2	25	0	22	1 (Japanese); 1 (bilingual); 1 (trilingual)

*Data are cited from United Nations Development Programme.

Firstly, all of the top 25 universities are from developed countries/regions regardless of *ARWU*, *THE* or *QS-USNWR* and it appears top universities, representing the higher education quality to some extent, are grounded on economic development. In addition, by looking into top university in a wider context historically, best universities usually followed the growth of most advanced nations in the world. For instance, *Cambridge University* and *Oxford University* were the best universities of the world when UK placed its position at the top of the world.

Secondly, most of top 25 universities, e.g. 22 of top 25 universities in *ARWU*, 24 of top 25 universities in *THE* and 22 of top 25 universities in *QS-USNWR* are from English-speaking countries/regions, indicating the advantage of English as an international language for modern university development. Van Raan et al. (2011) argued that non-English papers decreased rankings of universities from non-English countries and this language effect happened in important rankings such as *ARWU*, *QS-USNWR* and *THE*. Marginson and van der Wende (2007) believe that in a global higher education environment the educational resources are unequally distributed and "the English language and institutions from the Anglo-American nations (especially the United States) are often dominant" (p. 307). Since English is gradually becoming the dominant international language, top universities are probably consequences of long-term development.

Finally, the abundance of financial/economic resources alone might be not sufficient to propel universities forward. Although based on economic growth some Asia universities have been recently listed as top 100 universities by rankings, e.g. Peking University from China, some oil-producing countries which are among richest countries in the world are still out of ranking lists.

5.3 Featuring WURs in global higher education

Marginson and van der Wende (2007) have pointed out that "world university rankings especially the global ranking of research performance, higher education itself has entered an era of open global competition between nations and between individual HEIs as global actors in their own right" (p. 307) and national higher education systems and institutions are increasingly judged by their positions at global scale. Hazelkorn (2011) has stated that "around the world, rankings consciousness has risen sharply and, arguably inevitably, in response to globalization and the pursuit of new knowledge as the basis of economic growth, and the drive for increased public accountability and transparency" (p. 6). The increasing concerns over rankings could be the development direction of university excellence in the context of globalization. Globalization is influencing global higher education by the processes of international interchange of views, culture, products for teaching and research as well as students and faculties/staff.

To examine the interactions between WURs and university and even higher education development, it is necessary to look over the status of higher education.

Higher education has switched to its current style for at least one decade in USA and most of other developed and developing countries. Simpson (1998) discussed the overwhelming nature of higher education in our complex society. He further suggested that there were five forces that lead to the condition of being overwhelmed and pointed out that the overpopulation and rapid rate of growth is one major reason to tax our institutions of higher learning and the rapid expansion of knowledge also contributed a lot to the sense of overwhelmedness in higher education.

Moreover, Simpson (1998) argued that the effect of mass media was another factor and "competition among the mass media has turned us into such a rating oriented society that even colleges and universities have become obsessed with opinion polls, image building, and rankings" (p. 268). In addition, the amount of paperwork generated by electronic means and the massive number of courses and programmes furthered the condition of overswhelmedness in higher education. Based on these statements, one can draw the conclusion that higher education development has limits as well as the volume and pace in our society. So it is important to balance the university development and make it under control and fit for society. In light of this point, it is a necessity to assess university quality and development nationally and internationally, for instance, rankings. Hazelkorn (2011) believes that rankings are unavoidable outcomes "for the intensification of global competition, around which, higher education as both the progenitor of human capital and knowledge has become the fulcrum around which geo-political battles for a greater share of the global market are being fought" (p. 15).

Meanwhile, WURs contribute to the policy making directly in higher education through the effect of mass media in our time. Hazelkorn (2011) has pointed out that "HEIs are knowledge intensive industries behaving as other actors/firms in a competitive environment; to survive and thrive, many institutions are making changes to institutional strategy or adapting their behavior to fit the 'norm' promulgated by rankings" (p. 15). To date, most of the ranking systems has been established for years and updated annually. These

rankings have directly or indirectly influenced policy makers in department of education and university committee. In contrast, the development of global higher education has been reflected in these rankings to some extent.

Further, the traits of university excellence could be profiled by analyzing top universities in league tables of WURs and their related indicators (**Fig. 8**, **Fig. 9**, **and Table 6**). Top universities in our times should not only cover the functions of traditional universities, e.g. teaching students and research, but also develop new functions such as internationalization promoter, multi-culture protector, and education for sustainability.

5.4   The *pros* and *cons* of university rankings

University ranking and league table have generated constant impacts on HEIs despite rising concerns about methodological and technical issues. In this study, university ranking is clarified as a new perspective and effective tool to recognize top universities. It is obvious that global higher education can benefit from the university rankings. Morphew and Swanson (2011) have figured out that "both the products and technology of higher education are nebulous and hard to measure, rankings provide a seemingly objective input into any discussion or assessment of what constitutes quality in higher education" (p. 186). Usher and Savino have suggested that "another notable aspect of league tables is that they are, for the most part, produced by commercial publishing enterprises. In part, this reflects the fact that rankings share some characteristics with 'consumer guides' to various products" (p. 5). Actually, these are the major advantages of university rankings as interpreted in the literature review part. Findings from this study revealed that different WURs had much in common by examining their league tables and methodologies. Top universities could then be recognized accordingly. These facts confirm that it is worth researching top universities by studying WURs. Hazelkorn (2008) has pointed out that "while universities, policymakers and stakeholderscriticize and lampoon league tables and rankings, few can afford to ignorethem — and most have incorporated them in some fashion into their strategicthinking if not their planning" (p. 213).

However, university rankings do have numerous flaws. University rankings somehow insist that everything can be measured and quantified in relation to university quality and performance. Miley (2012) has commented that rankings are overvalued since "we deceive ourselves by insisting thateverything can be measured and quantifiedand that economic accountability should bethe most important criterion driving society" (p. 1042). In addition, rankings have generated constant pressure on higher education globally and drive universities develop to unified style. Arimoto (2011) has pointed out that "emergence of worldwide academic ranking is establishing a hierarchy of higher education institutions as unified pyramidal structure around the world, in which the west-centered structure focused on the USA and the UK is prevailing" (p. 254). Further, ranking lists are somehow doubtful. Some universities were regarded as top 25 universities by one ranking but underestimated seriously by another one. As evidenced by Table 7 and Fig. 8, there was obvious disagreement between rankings on specific universities. For example, *University of California, San Francisco* which was ranked as the 17th in *ARWU* was not listed in *QS-USNWR* and *THE*, indicating its position was out of lists of top universities. In addition, *University of Hong Kong* was ranked as the 22nd in *QS-USNWR* but the 250th in *ARWU.* The underlying reasons for such disagreement are complex. One major reason could be the incomplete and premature rules and standards in ranking field. Another possible reason could be the methodological issue in *QS-USNWR*. Huang (2012) has observed that 50% of the indicators used in *QS-USNWR* were based on reputation measured by peer reviews and "the indicatorsused in *QS* Rankings might lack validity. In theprocess of calculating return questionnaire for universityreputation, QS Rankings failed to control the number andqualification of questionnaire, thus leading to a selectionbias" (p. 72). Morphew and Swanson (2011) have noted that "rankings are a game everyone plays, but a game with constantly shifting rules that no one can control" (p. 189). In general, research performance was overemphasized in university rankings. Van Vught et al. (2012) have pointed out that the development of *U-Multirank* could help to overcome one of the major flaws of existing rankings that put too much emphasis on research performance and neglect many other important roles of universities.

6 Conclusions

Characterization of world-class universities as well as their underlying components is a top priority for higher education development world widely in the new era. Worldwide university ranking (WUR) has started assessing university performance globally in light of increasing demands while generating fierce debates in academia as well as the general public with a focus on its usefulness and accuracy.

This study revealed the social background of university ranking and the development of WUR by examining its birth, present situation and future direction. The conception of ranking was studied in comparison with "evaluation" and "clarification" and ranking was found uniquely functional as information provider to all related stakeholders with regard to HEIs. Different types of WURs were categorized according to their primary target groups, producers (public or private organization), level (institutional or field-based), scope (transnational or international) and focus (education centered or research centered). Methodologies of WURs involving indicators and related weightings for evaluating university performance as well as key steps in producing rankings were deciphered to understand WUR in depth.

In addition, the impacts of WUR were analyzed by determining both advantages and disadvantages. Consumer information, assessment instrument, and stimulating competition were suggested as three major functions of ranking system. It could be argued that ranking is important and actually valuable for multi-stakeholders. WURs were also criticized fiercely for the issues of their methodology reliability and indicator validity as well as for overlooking the existing diversity of HEIs, especially overemphasizing research performance with a focus on natural science. In addition to the stated bias in criteria, there is a big challenge in the process of producing rankings since raters have not always been sure about their methods for standardizing all the components and variables. However, global university rankings could serve as an effective tool to research global universities at

top tier for determining the underlying factors of university excellence in the new era. It has been pointed out that stakeholders would be continuously influenced by such activities as it is still a long way to develop unbiased rankings in the wake of globalization.

This study analyzed university rankings further and provided some new findings accordingly.

Firstly, representative university rankings, i.e. *QS-USNWR, THE*, and *ARWU* were elected based on their features and increasing impact. Both league tables and ranking methodologies of these rankings were well examined.

Secondly, top 10 as well as top 50 universities ranked by three selected single-dimensional rankings were extracted and agreement between rankings on top tier universities was evaluated by both qualitative and quantitative comparisons. It has been found that different WURs lead to similar consequences.

Thirdly, methodologies of selected rankings were studied in detail and objective performance of university was determined as a key factor among three WURs.

Fourthly, universities from native English-speaking and developed countries were found more advantageous in the top lists of WURs. However, the abundance of financial/economic resources alone might be not sufficient to propel universities forward.

Finally, traits for top universities were featured in our times by not only covering the functions of traditional universities, e.g. teaching students and research, but also developing new functions such as internationalization promoter, multi-culture protector, and education for sustainability.

In conclusion, world university ranking, emerging along with the higher education development globally could help to recognize world-class universities and reveal key components of university excellence in the new era. Ignoring peers' performance by looking down upon all rankings is an illusion of reality in open society. However, the existing university rankings should not been abused by decision-makers or any other stakeholders since not everything can be measured and quantified especially in education field. In addition, different university rankings are usually published for specific primary target groups, e.g. *ARWU* mainly for academia and

researchers and *THE* probably for undergraduate students as reference. Moreover, inspiring young generations with virtues, e.g. sustainability is also an important principle driving society in addition to economic accountability, e.g. cost-effectiveness. The initiating and ongoing *U-Multirank* project might help to develop WUR into more reliable and objective system in the future.

To further this study in the next step, two potential directions are proposed. On the one hand, the viewpoints and attitudes of varied stakeholders towards rankings could be sampled and analyzed through well organized interviews. On the other hand, the special case of top universities such as Hong Kong University could be studied in detail by examining different ranking criteria and its development historically. The suggested study would improve our understanding of connections between university development and rankings and help to reveal the university excellence in our times.

7 References

A world of opportunity. (2005, September 8). *The Economist*, p. 14–16.

Academic Ranking of World Universities. (2011). ARWU. Retrieved May 5, 2011 fromhttp://www.shanghairanking.com/

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

Albers, S. (2009). Misleading rankings of research in business. *German Economic Review*, *10*, 352–363. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0475.2009. 00483.x/pdf

Arimoto, A. (2011). Reacting to academic ranking: knowledge production, faculty productivity form an international perspective. In J. C. Shin, R. K. Toutkoushian & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 229–258). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-1116-7

Boulton, G. (2011). University rankings: diversity, excellence and the European initiative. *Procedia Social and Behavioral Sciences, 13,* 74–82. doi:10.1016/j.sbspro.2011.03.006

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

Brötzmann, C. W. (2010). Educating for sustainable development: A case study of the SIT study abroad Iceland 2009 program. Retrieved from http://skemman.is/en/item/view/1946/5606

Buela-Casal, G., Gutierrez-Martinez, O., Bermudez-Sanchez, M., & Vadillo-Munoz, O. (2007). Comparative study of international academic rankings of universities. *Scientometrics, 71,* 349–365. doi: 10.1007/s11192-007-1653-8.

Butler, D. (2010). University rankings smarten up. *Nature*, *464*,16–17. doi:10.1038/464016a

Centre for Higher Education Research and Information, Open University, Hobsons Research (CHERI). (2008). Counting what is measured or measuring what counts? League tables and their impact on higher education institutions in England. Retrieved June 15, 2012 from http://www.hefce.ac.uk/pubs/ year/ 2008/200814/

Charon, A., & Wauters, J. P. (2008). University ranking: a new tool for the evaluation of higher education in Europe. *Nephrol Dial Transplant*, *23*, 62–64. doi:10.1093/ndt/gfm279

Cheng, Y., & Liu, N. C. (2008). Examining major rankings according to the Berlin Principles. *Higher Education in Europe*, *33*, 201–208. doi: 10.1080/03797720802253686

Deem, R., Mok, H. K., & Lucas, L. (2008). Transforming higher education in whose image? Exploring the concept of the 'world-class' university in Europe and Asia. *Higher Education Policy*, *21*, 83–97.

Dehon, C., McCathie, A., & Verardi, V. (2010). Uncovering excellence in academic rankings: a closer look at the Shanghai ranking. *Scientometrics, 83,* 515–524*.* doi: 10.1007/s11192-009-0076-0

Dill, D.D. (2006). *Convergence and diversity: The role and influence of university rankings*. Retrieved January 20, 2012 from http://www.unc.edu/ppaq/docs/Kassel_Paper_Final_US.pdf

Enli, G. S. (2010). Comparative analysis and case studies. Retrieved August 10, 2012 from http://www.uio.no/studier/emner/ hf/imk/ MEVIT4800/h10/Method_4800_9.11.10.pdf

Enserink, M. (2007). Who ranks the university rankers? *Science*, *317*, 1026–1028. doi: 10.1126/science.317.58 41.1026

European Commission. (2010). EUR24187–Assessing Europe's university-based research - Expert group on assessmentof university-based research. Luxembourg: Publications Office of the European Union. doi: 10.2777/80193

Federkeil, G. (2008). Rankings and quality assurance in higher education*. Higher Education in European, 33,* 219–231*.* doi:10.1080/03797720802254023

Federkeil, G., File, J., Kaiser F., van Vught, F. A., & Ziegele, F. (2012). An interactive multidimensional ranking web tool. In F. van Vught & F. Ziegele (Eds.), *Multidimensional ranking: The design and development of U-Multirank* (pp. 167–178). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-3005-2_7

Guri-Rosenblit, S., Sebkova, H., & Teichler, U. (2007). Massification and diversity of higher education systems: Interplay of complex dimensions. *Higher Education Policy*, *20*, 373–389. doi:10.1057/palgrave.hep.8300158

Hazelkorn, E. (2008). Learning to live with league tables and ranking: The experience of institutional leaders. *Higher Education Policy, 21,* 193–215. doi: 10.1057/hep.2008.1

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

Hazelkorn, E. (2011). *Globalization and the reputation race in rankings and the reshaping of higher education: The battle for world class excellence*. London: Palgrave MacMillan.

Huang, M. H. (2011). A comparison of three major academic rankings for world universities: From a research evaluation perspective. *Journal of Library and Information Studies, 9*, 1–25. Retrieved from http://jlis.lis.ntu.edu.tw/article/v9-1-1.pdf

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

IREG Observatory on Academic Ranking and Excellence. (2006). Berlin Principles on Ranking of Higher Education Institutions. Retrieved September 19, 2012 from http://www.ireg-obser vatory.org/index.php?option=com_content&task=view&id=41 &Itemid=48

IREG Observatory on Academic Ranking and Excellence. (2011). IREG Ranking Audit Takes Off. Retrieved September 19, 2012 from http://www.ireg-observatory.org/index.php?option=com _content&task=view&id=187&Itemid=162

IREG Observatory on Academic Ranking and Excellence. (2011). IREG Ranking Audit Manual. Retrieved September 19, 2012 from http://www.ireg-observatory.org/index.php?option=com _content&task=view&id=187&Itemid=162

Lazaridis, T. (2010). Ranking university departments using the mean h-index. *Scientometrics, 82*, 211–216. doi: 10.1007/s11192-009-0048-4.

Longden, B. (2011). Ranking indicators and weights. In J. C. Shin, R. K. Toutkoushian, & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp.73–104). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-1116-7

Lukman, R., Krajnc, D., & Glavic, P. (2010). University ranking using research, education and environment indicators. *Journal of Cleaner Production, 18*, 619–628. Retrieved from http://www .sciencedirect.com/science/article/pii/S0959652609003047

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

Mervis, J. (2003). Who's No. 1? Academy hopes new rankings will say more. *Science, 302*, 1883–1885. doi: 10.1126/science.302.5652.1883

Mervis, J. (2010). Academy rankings tell you a lot, but not who's No. 1 in any field. *Science, 330*, 18–19. doi:10.1126/ science.330.6000.18

Miley, G. (2012). Saudi university policy: Overvalued rankings. *Science*, *335*, 1041–1042. doi: 10.1126/science. 335.6072.1041-b

Morphew, C. C., & Swanson, C. (2011). On the efficacy of raising your university's rankings. In J. C. Shin, R. K. Toutkoushian & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 185–199). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-1116-7

Movement in the ranks. (2011, November 16). *Nature,* p. 435. doi:10.1038/nj7373-435a

Rauhvargers, A. (2011). Global university rankings and their impact. Belgium: European University Association asbl.

Sadlak, J., & Liu, N. C. (2009). Introduction to the topic "world-class": Aspirations and reality checks. In J. Sadlak & N. Liu (Eds.), *The world-class university as part of a new higher education paradigm: From institutional qualities to systemic excellence* (pp. 13–19). Bucharest, Shanghai, Cluj-Napoca: UNESCO European Centre for Higher Education, Shanghai Jiao Tong University, Cluj University Press.

Salmi, J. (2009). *The challenge of establishing world-class universities*. Washington. DC: The World Bank.

Salmi, J. (2011). The road academic excellence: lessons of experience. In P. G. Altbach & J. Salmi (Eds.), *Academic excellence: The making of world-class research universities* (pp. 323–346). Washington, DC: The World Bank.

Shin, J. C., & Harman, G. (2009). New challenges for higher education: Global and Asia-Pacific perspectives.*Asia Pacific Educ. Rev*, *10*, 1–13. doi: 10.1007/s12564-009-9011-6

Shin, J. C., & Toutkoushian, R. K. (2011). The past, present, and future of university rankings. In J. C. Shin, R. K. Toutkoushian & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 1–16). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-1116-7

Simpson, R. D. (1998). The overwhelming nature of higher education. *Innovative Higher Education, 22*, 267–269.

Skulason, P. (2009). L'Université et l'éthique de la connaissance. *Philosophia Scientiæ,13,* 177–199. doi: 10.4000/philoso phiascientiae.81

Swedish National Agency for Higher Education. (2009). Ranking of universities and higher education institutions for student information purposes? Stockholm: Swedish National Agency for Higher Education. Retrieved 16 January, 2012 from http://www.hsv.se/download/18.3127d9ec12496ad89237ffe85 7/0927R.pdf

Taylor, P., & Braddock, R. (2007). International university ranking systems and the idea of university excellence. *Journal of Higher Education Policy and Management*, *29*, 245–260. doi: 10.1080/13600800701457855

Teichler, U. (2011). Social context and systemic consequence of university rankings: A meta-analysis of the ranking literature. In J. C. Shin, R. K. Toutkoushian & U. Teichler (Eds.), *University rankings: Theoretical basis, methodology and impacts on global higher education* (pp. 55–69). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/ 978-94 -007-1116-7

The QS World University Rankings. (2011). QS Top universities. Retrieved May 1, 2011 from http://www.topuniversities.com/

THE World University Rankings. (2011). Times Higher Education (THE). Retrieved May 5, 2011 from http://www.time shigh eredu cation.co.uk/world-university-rankings/

Tofallis, C. (2012). A different approach to university rankings. *Higher Education, 63,* 1–18. doi:10.100//s 10734-011-9417-z

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

Yang, R., & Welch, A. (2012). A world-class university in China? The case of Tsinghua.*Higher Education*, *63*, 645–666. doi: 10.1007/s10734-011-9465-4

Yin, R. K. (2009). Case study research: Design and methods (Applied social research methods). Thousand Oaks, CA: SAGE Publications, Inc.

van Raan, T., van Leeuwen, T., & Visser, M. (2011). Non-English speakers decrease rankings. *Nature, 469*, 34. doi: 10.1038/469034a

van Vught, F. A., & Westerheijden, D. F. (2010). Multidimensional ranking: a new transparency tool for higher education and research. *Higher Education Management and Policy*, *22*, 31–56.

van Vught, F., Westerheijden, D. F., & Ziegele, F. (2012). Introduction: Towards a new ranking approach in higher education and research. In F. van Vught & F. Ziegele (Eds.), *Multidimensional ranking: The design and development of U-Multirank* (pp. 1–7). Dordrecht, Heidelberg, London & New York: Springer Press. doi:10.1007/978-94-007-3005-2_7

van Vught, F., & Ziegele, F. (2011). U-Multirank: Design and testing the feasibility of a multidimensional global University ranking. Final report to Consortium for Higher Education and Research Performance Assessment (CHERPA-Network).Retrieved from http://ec. euro pa.eu/ education/higher-education/doc/multirank_en.pdf

Williams, R. (2008).Methodology, meaning and usefulness of rankings. *Australian Universies' Review*, *50*, 51–58. Retrieved from http://eric.ed.gov/PDFS/EJ824688.pdf