

Asia-Pacific Association for International Education

National University of Singapore, 7-9 March 2007

10.30 am Friday 9 March

Ranking Systems: Universities of Choice

Global university rankings: where to from here?

Simon Marginson

Centre for the Study of Higher Education

University of Melbourne, Australia

s.marginson@unimelb.edu.au

ph 613-83448060, 613-93170303

'Principles of division, inextricably logical and sociological, function within and for the purposes of the struggle between social groups; in producing concepts, they produce groups, the very groups which produce the principles and the groups against which they are produced'.

'The whole process constitutes a perfect circle from which the only escape is to objectify it sociologically'

~ Pierre Bourdieu *Distinction: A social critique of the judgment of taste*, transl. R. Nice. Routledge & Kegan Paul, London (1984), pp. 235 & 479

1. Introduction

University rankings are powerful. They compel public attention and shape the behaviour of universities and policy makers. Two decades of the annual *US News and World Report* rankings have closely influenced American practice. Institutions feel compelled to lift their *US News* position and this has encouraged for example the manipulation of student entry to maximize student scores and refusal rates, and the growth of merit-based student aid at the expense of needs-based aid (Kirp 2004). Likewise the Jiao Tong University research rankings focus national government attention on policies designed to concentrate research activity in a small number of universities. At the same time rankings strengthen encourage the flow of doctoral students, elite researchers and the philanthropic and corporate dollar into the top ranked institutions at the expense of the rest. Both the Jiao Tong and the *Times Higher* rankings encourage individual universities to do anything and everything to lift their rankings position, though they use differing criteria and point universities in somewhat divergent directions. These rankings date only from 2003 and 2004 respectively but already they are everywhere in the sector and beyond. They set university reputations.

Rankings function as a meta-performance indicator. They do more than 'reflect' a university's profile and quality. The criteria used to determine a university's position in the ranking system become meta-outputs that every university is duty bound to place on priority. Rankings begin to define what quality means and by shaping university and system behaviours they begin to shape university mission and the balance of activity. In the world according to Shanghai Jiao Tong University rankings, higher education is about scientific research and Nobel Prizes. It is not teaching or community building or solutions to local or global problems. In the world according to the *Times Higher Education Supplement*, higher education is primarily about building reputation as an end in itself, and about international marketing, because it is these metrics that drive the index. It is not about teaching, and not so much about research and scholarship which are only 20 per cent of the *Times* index.

Rankings as a meta-performance indicator have the potential to redefine and reify the core purposes of universities. They shape patterns of activity and priorities for development, as shown by the history of the *US News* rankings in the United States. They cut deeply into the authority of universities over mission and identity.

Rankings can also be capricious and destructive. There is much at stake. In 2004 the oldest public university in Malaysia, the University of Malaya was ranked by the *Times* at 89. The newspapers in Kuala Lumpur celebrated. The Vice-Chancellor ordered huge banners declaring 'UM a world's top 100 university' placed around the city, and on the edge of the campus facing the main freeway to the airport where every foreign visitor to Malaysia would see it. But the next year in 2005 the identity of Chinese and Indian students at the University of Malaya was corrected shifted from international to national, and there were shifts in other parts of the *Times* composite indicator. UM dropped from 89 to 169. It seemed that the University's reputation abroad and at home

was in free fall. The VC was pilloried in the Malaysian media. When his position came up for renewal by the government in March 2006 he was replaced. But it wasn't just the Vice-Chancellor whose reputation had been trashed by the *Times*, it was the University of Malaya, long established and one of the two strongest universities in an emerging knowledge economy with real virtues and strengths. The University of Malaya had dropped 80 places without any decline in its real performance (aside from spending too much on hubristic banners). In the drama of UM's decline there was no positive relationship between performance, competition and outcome. This does not generate useful incentives for better policy and management or better education or research provision. This is simply perverse.

But as in Kuala Lumpur so in every other capital. Higher education retains national and local dimensions but is now also a global system. No single nation can ignore global rankings with the exception of the United States. Precisely because the United States is globally hegemonic in higher education its institutions are solely focused on national not global rankings. Best in the USA automatically means 'best in the world' and what happens in the USA washes over the rest of the world even while most American institutions are indifferent to it. Otherwise, only Europe with its multi-national Bologna process has enough combined critical mass to change the geo-politics of rankings. Significantly it is the Europeans who are pioneering a new and very different approach to comparing university performance, that developed by the German Centre for Higher Education Development (CHE), as will be discussed below.

Not surprisingly global rankings have already promoted an immense number of meetings, papers and analyses. The number is increasing. I think this is the sixteenth conference or seminar paper I have given on this topic in 18 months. Rankings raise questions about the validity and utility of both the process of comparison and data used in that process. Given that all comparisons can only ever focus on some elements, not the whole university, are the elements being used for comparison the right ones to use? Are the hierarchies of institutions contained in league tables accurate and representative of the higher education sector? League tables generate clear-cut winners and losers. Do we want to elevate these winners and downgrade these losers? Are the outcomes fair and regarded as fair across the higher education world? And are the rankings systems useful in terms of outcomes? Do league tables help knowledge economies to develop faster or better? Do they provide data helpful to students?

In many quarters there is a sense that all is not well with rankings. Perhaps this is inevitable given that in league tables there are few winners and many losers, but there is more at stake here than self-interest. Rankings change higher education and it is a question of what kind of global higher education system we want to have. There are widespread desires to modify the downsides of university rankings, which tend to close off options, and to provide better data in relation to teaching quality. In Europe these concerns underpin the wide support for the CHE approach to comparison, and for a typology of institutions with a diverse set of university missions, as in the Carnegie classification in the USA. Then there is the diversity issue. Systems and universities represent a broader range of national and educational traditions than those of

Cambridge Massachusetts and Cambridge UK. How can this broader range of traditions be encompassed? These issues are the cutting edge of discussion.

There is a strong sense of reflexivity in the discussions about rankings, within the rankings community itself, especially the discussions conducted by Jiao Tong University itself, and in Europe at UNESCO and OECD/IMHE meetings. The Jiao Tong Institute of Higher Education is constantly tuning its rankings and invites open collaboration in that. It is a strength of the academically rigorous and globally inclusive Jiao Tong approach. On the whole *The Times* has been less transparent and inclusive, though that process too has been opening up to some extent, which is welcome.

Global rankings have entered a 'second stage' in which systems and approaches are being criticised and alternative approaches are being canvassed. It is not yet clear where this reflexivity is taking universities. We go further down the path mapped by the *US News and World Report*, the *Times Higher* and/or the holistic university research performance measures used by Jiao Tong. Or these could be significantly modified. We could emphasise discipline rankings rather than whole of institution rankings – a direction encouraged by the release of the discipline data from Jiao Tong – or create more plural league tables including lists of specialist institutions. Or we could move in a different direction entirely. There is also the question of who decides the future of rankings: publishing and market research companies, governments, international agencies, universities, social science scholars? Or some mix of the above.

In the spirit of second stage discussion, this paper proceeds as follows. It notes the ubiquitous role of national rankings (section 2), examines the two global rankings systems,¹ Shanghai Jiao Tong and the *Times Higher Education Supplement* (sections 3 and 4) and discusses the Centre for Higher Education Development (CHE) approach to comparison (section 5). It then kicks around the problems of holistic institutional league tables in the light of the diversity of the global sector (section 6). It closes with thoughts about where we go from here (section 7).

2. National rankings

Most countries with large higher education systems have rankings of one kind or another. Countries with rankings devised by newspapers and magazines include China and Hong Kong China, Japan, India, the Ukraine, Romania, Poland, Portugal, Italy, Spain, Germany, Sweden, Switzerland, France, UK, USA and Canada. In Thailand, Malaysia, Pakistan, India, Kazakhstan, Korea, Tunisia, Nigeria, the Netherlands, the UK, Brazil and Argentina rankings have been instigated by ministries of education, grants councils or accreditation agencies. In China, Japan, Australia, Kazakhstan, Slovakia, Romania, Russia, the Ukraine, Germany, Spain, Switzerland, the UK and Canada rankings have been initiated by universities, professional associations or other organisations (Salmi and Saroyan 2006). Usher and Savino (2006) review rankings compiled in Australia (Williams and Van Dyke 2006); Canada; China, with several rankings systems (Liu and Liu 2005), and Hong Kong; Germany; Italy; Poland; Spain; the UK, where four different newspapers have developed rankings; and the USA.

For the most part national rankings consist of a single table but in the USA and Canada higher education institutions have been divided into groups according to mission and other characteristics, creating a set of mini league-tables within which the category of comprehensive research universities has highest status.

Specialist rankings focus on characteristics ranging from research output, to student services, to MBA programs, to *Yahoo Magazine's* ratings of 'connectivity', the university contribution to social diversity and other features. Specialist rankings are not further discussed here but it can be noted that ratings of disciplines and other aspects feed into the larger schemes of Shanghai Jiao Tong University and the German CHE.

In the United States the annual *US News and World Report* survey focuses on aspects of institutions seen to contribute to the quality of teaching and the student experience, rather research and scholarship. The categories of institutions are drawn from the classification in 2000 by the Carnegie Foundation for the Advancement of Teaching (USNWR 2006, p. 81ff). The most important grouping is 'National Universities', 248 universities (162 public and 86 private) with a wide range of fields of study and offering degrees to doctoral level. Most of these institutions are research intensive. Other categories are liberal arts colleges, Master's level universities, Comprehensive Bachelor level universities, business programs, undergraduate engineering programs, and speciality schools mostly in the fine and performing arts.

In National Universities and liberal arts colleges the largest part of the index (25 per cent) is comprised by a survey of 'undergraduate academic excellence' sent to university presidents, provosts and deans of admissions. Two items constitute 20 per cent: student retention and graduation rate; and 'faculty resources' which rewards small classes, high academic salaries, high academic qualifications and a high proportion of staff full-time. Student selectivity at entry, a proxy for market demand, is 15 per cent. The lesser items are spending per student (10 per cent), the proportion of alumni who donate back to the institution, and the graduation rate after controlling for spending and student aptitude (each 5 per cent) (USNWR 2006, pp. 77-79).

3. The Jiao Tong University global ranking

The first world-wide ranking by the Shanghai Jiao Tong University Institute of Higher Education was in 2003. University-wide rankings are issued annually and the first rankings in five broad disciplinary fields were released last month. The sole focus is research. The Jiao Tong group argues that the only data sufficiently reliable for ranking purposes are broadly available and internationally comparable data of measurable research performance (Liu & Cheng 2005, p. 133). It is considered impossible to compare teaching and learning 'owing to the huge differences between universities and the large variety of countries, and because of the technical difficulties inherent in obtaining internationally comparable data'. Further, the Jiao Tong group states that it does not want to employ subjective measures of opinion or data sourced from universities themselves as are used in some national rankings (SJTUIHE 2006).

Table 1. Shanghai Jiao Tong university rankings: weightings

	weighting
Alumni of institution: Nobel Prizes and field medals	10%
Staff of institution: Nobel Prizes and field medals	20%
High citation (HiCi) researchers	20%
Articles in citation indexes in science fields, social science and humanities	20%
Articles in <i>Science</i> and <i>Nature</i>	20%
Research performance (as above) per head of staff	10%
total	100%

Source: SJTUIHE 2006

The bulk of the index is determined by publication and citation, principally in the science-based disciplines with some attention to social sciences and humanities: 20 per cent citation in leading journals; 20 per cent articles in *Science* and *Nature*; and 20 per cent the number of Thomson/ISI 'HiCi' researchers on the basis of citation (ISI, 2006). Another 30 per cent is determined by the winners of Nobel Prizes in the sciences and economics and Fields Medals in mathematics, based on the location of training (10 per cent) and current employment (20 per cent). The remaining 10 per cent is determined by dividing the total derived from the above data by the number of faculty. Jiao Tong research performance is dominated by the English speaking nations, which have 71 per cent of the world's top 100 research universities, and particularly by the United States which has 17 of the top 20 and 54 of the top 100 in 2006.

For the most part the Jiao Tong index is methodologically sound and a valid basis for synchronic global comparisons. It measures only real outputs, rather than subjective assessments of reputation, which may or may not be solidly grounded. Its methods are transparent and the collection has improved over time. Only the Nobel Prize criterion is somewhat controversial, in that Nobel Prizes are submission based. Science and scholarship are not the only determining factors. Politicking can enter the decisions. The Jiao Tong exercise also creates a positive relationship between improved research performance relative to others, and a higher ranking. To this extent it produces data useful for public and institutional policy purposes. It is helpful for national governments to measure the nation's absolute and relative capacity as a knowledge economy in terms of the basic and applied research conducted in the nation's universities, though there remains the question of to what extent that knowledge economy capacity, which is grounded in the global research system, is utilised nationally.

The chief problem with the Jiao Tong data lies not in their validity but their use. For the most part they are understood around the world not as a ranking of university research performance but as a holistic ranking of the universities concerned and a marker of reputation in the emerging global university market (Marginson 2006; 2007). Harvard becomes understood not as number one research site according to Shanghai Jiao Tong University but as number one *university*. This is despite the explicit urging of the Jiao Tong group *not* to interpret the data as holistic rankings; and despite the fact that those data favour certain kinds of institution and disadvantage others.

The Jiao Tong calculations favour universities large and comprehensive enough to amass strong research performance over a broad range of fields while carrying few research inactive staff. They favour universities very strong in the sciences, universities from English language nations because English is the language of research (non English language work is published less and cited less) and universities from the USA because Americans tend to cite Americans. The number of Thomson/ISI 'HighCi' researchers directly and indirectly drives performance in much of the index.

A massive 3614 of the 'HighCi' researchers are in the USA, compared to 224 in Germany, 221 in Japan, 162 in Canada, 138 in France, 101 in Australia, 94 in Switzerland, 55 in Sweden, 20 in China and none in Indonesia (ISI 2006). Harvard and its affiliated institutes have 168 HiCi researchers, more than France or Canada. Stanford has 132 HiCi researchers; UC Berkeley 82 and MIT 74. There are 42 at the University of Cambridge in the UK (ISI 2006).

4. The *Times Higher* global ranking

The first *Times Higher* 'World University Rankings' were published in 2004. In contrast with Shanghai Jiao Tong the explicit aim is to produce a summative, holistic ranking. Half of the *Times Higher* index is comprised by university reputation. As well as the 40 per cent comprised by a reputational survey of academics ('peer review'), another 10 per cent is determined by a reputational survey of 'global employers'. In addition there are two internationalisation indicators: the proportion of students who are international (5 per cent) and the proportion of staff (5 per cent). Another 20 per cent is determined by the student-staff ratio, a quantity measure used as a proxy for teaching 'quality'. The remaining 20 per cent is comprised by research citation performance (THES 2006).

Table 2. *Times Higher* university rankings: weightings

	weighting
'Peer review' survey of academics	40%
Survey of 'global employers'	10%
Proportion of academic faculty who are foreign	5%
Proportion of students who are foreign	5%
Staff-student ration (proxy for 'teaching quality')	20%
Research citations per head of academic faculty	20%
total	100%

Source: THES 2006

Methodologically, the '*Times* tables' are open to criticism. It is not specified who is surveyed or what questions are asked. The survey gathers a response of just 1 per cent from 200,000 e-mails sent worldwide and not all responses are valid and can be used. The responses that do come in tend to be from nations where the *Times* is well-known, so the composition of the pool of responses is heavily weighted in favour of the UK, Australia and some former British colonial locations, for example in Southeast Asia. Despite this composition effect the pool of responses is not re-weighted for global

evenness, which may not be practical when the overall response is 1 per cent. Not surprisingly survey outcomes seem to reflect survey composition, as discussed below.

The student internationalization indicator rewards entrepreneurial volume building but this is not necessarily grounded in the quality of student demand or the quality of programs or services. Teaching quality cannot be adequately assessed using student-staff ratios. Research plays a lesser role in this index than in most understandings of the role of universities. The *Times Higher* rankings reward a university's marketing division better than its researchers.

The *Times* results are too easily to manipulate. By changing the recipients of the two surveys, or how the survey results are crunched (this year the results of three annual surveys, with different recipients, were combined into a single set of results) the results change and often change markedly. This illustrates the more general point that rankings frame competitive market standing as much as they reflect it.

Table 3. Some volatile movements in the *Times Higher* rankings, 2004-2006

University	<i>Times</i> ranking 2004	<i>Times</i> ranking 2005	<i>Times</i> ranking 2006
Fudan, China	195	72	=116
Tsing Hua, China	61	= 62	28
Keio, Japan	-	215	120
Kyushu, Japan	-	222	=128
Osaka, Japan	69	=105	= 70
Tokyo Institute of Technology, Japan	51	99	118
Seoul National, Korea	118	= 93	63
Malaya, Malaysia	89	=169	192
Chulalongkorn, Thailand	-	=121	=161
Royal Melbourne Institute of Technology, Australia	55	82	146
Queensland Technology, Australia	-	118	=192
Massey, New Zealand	108	=188	-
Otago, New Zealand	114	=186	= 79
Munich, Germany	99	55	98
Ecole Pol. Federale de Lusanne, Switzerland	32	34	= 64
Geneva, Switzerland	-	= 88	39
Basel, Switzerland	-	=127	75
Ecole Polytechnique, France	27	10	37
Sciences Po, France	-	69	52
Birmingham, UK	126	=143	=90
School of Oriental & African Studies, UK	44	=103	= 70
Dartmouth, USA	138	117	= 61
Emory, USA	173	141	56
Pittsburgh, USA	-	193	88
Purdue, USA	59	61	127
Vanderbilt, USA	156	=114	53

Source: *THES* 2006 and predecessors

Whether or not the *Times* manipulates its surveys to generate a pre-determined result, over the life of the index three things have become apparent. First, the results are highly volatile. The free fall of the University of Malaya is just one example. There many sharp rises and falls, especially in the second half of the Times top 200 where small differences in metrics generate large rankings effects. Fudan in China has oscillated between 72 and 195. Seoul National between 63 and 118. RMIT in Australia between 55 and 146. Massey in New Zealand, which started at 114, is out of the top 200. Munich has oscillated between 55 and 98, Ecole Polytechnique between 10 and 37, the School of Oriental and African Studies in the UK between 44 and 103. In the USA Emory has risen from 173 to 56, and Purdue fallen from 59 to 127.

Second, the British universities do too well in the *Times* table. They have done better each successive year. The UK has 15 per cent of the GDP of the USA but almost half the number of universities in the *Times* top 100 as the USA: 15 from the UK and 33 from the USA in 2006. Whereas the USA has 54 research universities in the Jiao Tong top 100 the *Times* manages to reduce American world hegemony to just 33. In 2006 Cambridge and Oxford suddenly improved their performance despite Oxford's current problems. The British now have two of the *Times* top three and Cambridge UK has almost closed the gap on Harvard. Yet the British universities are manifestly under-funded and the Harvard faculty is cited at three and a half times the rate of its British counterparts. It does not add up. But the point is that it depends on who fills out the reputational survey and how each survey return is weighted.

Table 4. Australian universities in the *Times Higher* ranking of the top 200, 2004-2006

University	<i>Times</i> ranking 2004	<i>Times</i> ranking 2005	<i>Times</i> ranking 2006
Australian National University	18	23	16
University of Melbourne	22	19	22
University of Sydney	40	=38	=35
Monash University	33	33	38
University of New South Wales	36	40	41
University of Queensland	49	47	45
Macquarie University	68	67	=82
University of Adelaide	56	=80	=105
University of Western Australia	96	=80	=111
Royal Melbourne Institute of Technology	55	82	146
Curtin University of Technology	76	=101	=156
Queensland University of Technology	--	118	=192
University of Wollongong	--	--	196
La Trobe University	142	98	--
University of Newcastle	--	=127	--
University of South Australia	--	=154	--
University of Tasmania	161	=166	--

Source: *THES* 2006 and predecessors
 -- indicates ranking position outside the top 200

Third, the performance of the Australian universities is also inflated. Despite a relatively poor citation rate and moderate staffing ratios they do exceptionally well in the reputational academic survey and internationalisation indicators especially that for students. ANU has been ranked by the 3703 'academic peers' surveyed by the *Times* as ahead of Yale, Princeton, Caltech, Chicago, Penn and UCLA. Melbourne is ranked on the same level as Yale. That sounds like a huge achievement but it is not really plausible. Australia has 13 of the *Times* top 200 and appears as the third strongest system, ahead of Japan, Canada, Germany and the rest of Western Europe. This makes sense in relation to international marketing but not all round performance.

The OECD (2005 and 2006) has taken Australia to task its long-term reduction in public funding of universities. Having done empirical case work recently in research universities in Canada I can testify that aside from the quantity of international students, leading universities in that nation are stronger than their Australian counterparts. Canada has a similar higher education system to Australia but is 50 per cent larger, it enjoys much better total funding and public funding, it has higher participation rates in higher education and stronger research performance in the Jiao Tong rankings, with one university, Toronto, at 24. Australia's highest is ANU at 54. Yet Canada has three universities in the *Times* top 100, Australia has seven.

5. The Centre for Higher Education Development (CHE) comparisons

The Centre for Higher Education Development (CHE), located in Gutersloh in the state of north Rhine-Westphalia in Germany, in conjunction with the German Academic Exchange Service which assists international students and the publisher Die Zeit, has developed a distinctive approach to national rankings. This dispenses with holistic (summative) rank ordering of institutions in league tables. CHE surveys 130,000 students and 16,000 academic faculty in almost 250 higher education institutions, focusing on student experiences, student satisfaction, and academic recommendations on the best places in each field of study. It supplements the surveys with independent sources comprising one third of the total data base. No data are taken from institutions. CHE ranks institutional departments according to each separate indicator of academic and service quality, assigning them to top third, middle third or bottom third of all institutions. It provides neither discipline-based league tables nor institutional league tables. It refuses to integrate the different indicators into a single indicator for each institution because there is no 'one best university' across all areas, and 'minimal differences produced by random fluctuations may be misinterpreted as real differences' in holistic rankings systems (CHE 2006).

The CHE data are made available to prospective students and the public free of charge though an interactive web-enabled database. Any person can interrogate this data base by investigating and rank identified disciplines and administrative services, using their own combination of criteria (CHE 2006), thereby creating weightings and rankings themselves. CHE acknowledges that the definition of 'quality' is purpose-driven and open to variation and passes power over the definition from the ranker to the consumer. It also provides by far the most comprehensive set of comparative data devised so far. A difficulty is that the CHE approach requires sufficient homogeneity

between institutions and their programs and services to enable the comparisons. Arguably it would be necessary to develop several sets of institutions for CHE-style comparisons in more varied environments such as that of the United States.

This process of data collection has been extended from all higher education institutions in Germany to Switzerland and Austria. The Netherlands and Belgium (Flanders) are preparing to join the system, and some Nordic countries may follow. The CHE ranking system is well positioned to develop into a European-wide system, and has attracted attention in other parts of the world (Usher and Savino 2006; see also Van Dyke 2005) including Canada and the USA. The Commission on the Future of Higher Education in the US is working on a comparable concept that would allow consumers to rank colleges based on variables of their choosing (Field 2006).

6. Some problems of global rankings

The present global university rankings bring with them a number of limitations, problems and distortions. These issues will now be reviewed briefly.

(1) Rankings become an end in themselves and protected from critical scrutiny: Rankings are hypnotic and become an end in themselves without regard to exactly what they measure, whether they are solidly grounded or whether their use has constructive effects. The desire for rank ordering overrules all else. Often institutions are rank ordered even where differences in the data are not statistically significant.

(2) National, institutional and program diversity: The two global rankings systems confront a world-wide network of national systems and public and private institutions that is very diverse. The primary forms of diversity are in institutional mission and identity, and in language especially language of instruction. Both forms of diversity have been closely shaped by national and local traditions and cultures.

Within the research university category there are established national systems of high quality in non-English speaking countries in Western Europe, for example Germany, Switzerland, Sweden, Finland, Belgium and the Netherlands; and a distinctive Latin American model has evolved. Each of UNAM in Mexico and the University of Buenos Aires in Argentina provide access to well over 200,000 students on many campus sites and perform a broad range of functions in national and regional development, and social and cultural life. In all of these nations there is significant scholarship in languages other than English that is unrecognised in the rankings process, and the broad functions of UNAM and UBA prevent them from concentrating resources so as to maximise research intensity and reputation in the manner of a Princeton or Caltech. These are great national universities and the model has distinctive strengths and a long international role in higher education but it does not fit the USA/UK template.

Beyond the research university category there are many other kinds of institution with coherent missions – from liberal arts colleges in the USA; to specialist schools in such fields as business, engineering, medicine, agriculture, media and the arts; to specialist research institutes such as those in France and Germany; to the great range of

institutions specifically focused on vocational and professional training without a function in basic research. In some nation, unlike the English-speaking nations, vocational institutions enjoy relatively high levels of prestige and resources, for example the German *Fachhochschulen* and the vocational sectors in Finland, Switzerland, France. Another example is the Indian Institutes of Technology (IITs). In both the Jiao Tong and Times Higher global rankings all forms of specialist institution suffer by comparison with large scale comprehensive research universities. Via its reputational survey and survey of global employers the *Times Higher* values the IITs and research universities emphasizing links to industry but not the *Fachhochschulen*.

It should be possible to understand worldwide higher education as a combination of differing national and local traditions, models and innovations; in which some universities do better than others but no single model as such is necessarily always supreme. There could be a large range of possible global models for the purposes of comparison (Table 5). But currently only two global models are provided and these overlap in significant respects. As noted both the Jiao Tong and *Times Higher* rankings tend to benefit certain kinds of university: research-intensive comprehensive science universities in the Jiao Tong; prestigious and comprehensive universities which tend to be research universities because of the central role of research in determining status, that are also engaged in international marketing, in the *Times Higher*. Only wealthy nations provide scientific infrastructure for basic research at sufficient scale to figure in the Jiao Tong top 50, mostly via government funding. Only a small number of institutions enjoy pre-existing prestige sufficient to get onto the radar of the *Times* reputational surveys.

Table 5. Possible norms of higher education institution for the purposes of global comparison

Ivy League research universities in the USA
Flagship state research universities in the USA
For-profit vocational institutions, as in U Phoenix
British state-affected research universities
European state-centred research universities (various)
German <i>Fachhochschulen</i> and other vocational traditions
State-building university/reform tradition in Latin America
State-led science universities in East Asia
Large open and distance education providers
Commercial 'e-Universities'
Specialist institutions in business, medicine, agriculture, arts, etc.
Specialist research institutes

As noted, governments and institutional leaders have powerful incentives to refashion institutions so as to maximise rankings performance. Yet the Jiao Tong and *Times*

norms constitute only a small fraction of the totality of the global higher education sector; and the narrowing of diversity is likely to significantly reduce the range and depth of the benefits enabled by higher education.

(3) *National and linguistic diversity*: English is one of two languages spoken by one billion people. The other is Putonghua ('Mandarin' Chinese). In addition two pairings of related and mutually intelligible languages are spoken by more than half a billion people: Hindi/ Urdu, and Spanish/ Portuguese. Another three languages are spoken over 200 million people: Russian, Bengali and Arabic. A further four languages have more than 100 million speakers (Table 6). These languages are unlikely to disappear.

In both global rankings surveys institutions in English-speaking countries have advantages. Global research conversations are conducted largely in English. The historical and present imperial roles of their nations ensure that Anglo-American institutions enjoy maximum prestige in the *Times Higher*. Global rankings problematize even the best universities outside the English language systems. The Nordic countries, the Netherlands, Germany and Singapore conduct a growing proportion of doctoral programs and Masters courses in English; even so they enter global competition from behind. Others are further back. If English is used as a foreign language in the nation concerned, rather than as a second language, the position tends to be more difficult.

Table 6. Spoken languages with more than 100 million voices world-wide

language/ language group	number of voices
	millions
English	1000
Putonghua ('Mandarin')	1000
Hindi/Urdu	900
Spanish/Portuguese	450/200
Russian	320
Arabic	250
Bengali	250
Malay-Indonesian	160
Japanese	130
French	125
German	125

Source: Linguasphere Observatory (2006)

Most academic faculty from the English language nations are monolingual but this is not the case in other university systems. Although English is the only global language of research publication it is not the only language in which research is conducted, but most of the research and scholarship originating in languages other than English falls outside the global conversation in English because it is not translated into English. Ideas originating in English tend to be much more broadly spread. Approximately ten times as many books are translated from English to other languages, as are translated from other languages into English and thus made universally accessible (Held et al. 1999). Global rankings tend to reinforce these asymmetries and exclusions.

Increasingly, throughout the world, academic faculty have formal or informal incentives to publish in globally-recognized English language journals. But this norming of English as the single common academic language reduces what can be said.

(4) Partial coverage of purposes and stakeholders: Rankings systems based on institutions, rather than single disciplines, appear to 'evaluate universities as a whole' (van Dyke 2005, p. 106). But no system of rankings covers all purposes of higher education. Any ranking system encompasses the needs of some stakeholders better than others. When rankings systems attempt to cover the generality of purposes and interests, the problem of partial coverage and exclusion is hidden but compounded. Usher & Savino (2006) examine 19 league tables and university rankings systems from around the world. They note that different rankings systems are driven by very different purposes and associated with different notions of what constitutes university quality.

This problem is fundamental because the areas excluded by the Jiao Tong and *Times Higher* rankings include teaching quality. Teaching is difficult or impossible to measure with rigour for comparative national purposes let global comparison; and no ranking or quality assessment system has generated comparative data based on measures of the 'value added' during the educational process. Few comparisons focus on teaching and learning as such. Rather, various proxies for teaching 'quality' are used, such as quantity resource indicators including average student-staff ratios, student selectivity (actually a proxy for reputation not for teaching quality), and research performance. Some might assume that research performance is positively correlated to teaching quality but 'empirical research ... suggests that the correlation between research productivity and undergraduate instruction is very small and teaching and research appear to be more or less independent activities' (Dill & Soo, 2005, p. 507).

(5) The problem of aggregation and weightings: Usher & Savino (2006) also remark on the arbitrary character of the weightings used to construct composite indexes covering different aspects of quality or performance, the means by which ratings agencies construct a total picture of the institutions that are ranked against each other. 'The fact that there may be other legitimate indicators or combinations of indicators is usually passed over in silence. To the reader, the author's judgment is in effect final' (Usher & Savino 2006, p. 3). As Salmi and Saroyan (2006, p. 9) note, 'the weightings vary across league tables and typically reflect the view of the publisher of the table rather than being grounded'. The composite indicators and weightings are untheorised.

(6) Reputational rankings are ill-grounded and circular: In reputation-based rankings known university brands generate 'halo' effects. The *Times Higher* favours universities already well known regardless of merit, tending to recycle existing reputations while blocking newcomer institutions or nations. There is no means of verifying the soundness of subjective judgements of reputation, for example ensuring that they are grounded in actual comparative knowledge. One study of ranking found that one third of those who responded to the survey knew little about the institutions concerned apart from their own. The classical example of these problems is the American survey of students that found Princeton law school was ranked in the top ten law schools in the country. But Princeton did not have a Law school (Frank & Cook 1995, p. 149).

(7) *Rankings produce context-free judgement:* Once institutions are arranged in a single league table the particular historical, economic and cultural factors that shape relative performance (such as national language of use or economic resources per head) become obscured. In this manner rankings readily encourage flawed policy judgement. They foster the illusion of a level playing field in which every university can rise up the table on merit and its place is determined by its own efforts. But all universities have specific conditions of possibility and global higher education is not a level playing field. Institutions are tied to their history, to their local context, to national resourcing, especially for basic research, and to the capacities of their personnel. They have scope to improve but not an absolute freedom to move up the league table.

Consider for example the ranking prospects of universities in Indonesia compared to universities in the USA. The USA has 300 million people, produces over a fifth of global GDP at a per capita rate near highest in the world. It invests 2.9 per cent of GDP in tertiary education, the highest rate of investment, and in 2001 produced almost a third of the world's annual output of scientific papers, 200,870. It has 33.4 per cent of the Jiao Tong top 500 research universities and 54.0 per cent of the Jiao Tong top 100. Indonesia has 220 million people, about three quarters of the population of the USA. Its GDP is the fifteenth largest at 1.4 per cent of world output and GDP per head is at one tenth the level of the USA. It spends about 0.7 per cent of GDP on tertiary education and in 2001 produced 207 scientific papers, one thousand the number of papers of the USA. It has no universities in the Jiao Tong top 100 or top 500 and little early prospect of placing any there. This is not a level playing field. Finding out that American universities outgun Indonesian universities bolsters American power but tells us nothing new. Still less does it enable or encourage improvement in Indonesia.

Table 7. The starting position for global comparison: United States versus Indonesia

	USA	Indonesia
Population 2005	296.5 million	220.6 million
GDP PPP 2005	\$12,409 billion	\$643 billion
Gross National Income per head PPP 2005	\$43,740	\$3730
GDP for tertiary education (USA 2004 Indonesia 2003)	2.9%	0.7%
Scientific papers 2001	200,870	207
Proportion of Jiao Tong top 500 universities 2006	33.4%	0
Proportion of Jiao Tong top 100 universities 2006	54.0%	0

Source: World Bank 2006, OECD 2006. PPP = Purchasing Power Parity.

(8) *Rankings undermine universal improvement:* As noted in relation to the *Times Higher*, when rankings are capricious or are derived from reputation without any necessary link to the quality of outcomes, there is no necessary incentive to improve. Further, while competition for relative position generates universal competition it rarely sustains a universal process of absolute improvement. Because league tables are a zero sum game they generate expectations of failure as well as hopes for success. How can global university rankings, dominated as they are by a handful of wealthy

nations, contribute in positive fashion to capacity building in emerging national systems such as Indonesia? Even where the ranking system is sound in that it measures real performance, as with the Jiao Tong rankings, some institutions will be discouraged; while others will lack the policy autonomy or resource base they need to introduce the internal reforms necessary to lift their relative position (Salmi and Saroyan 2006, p. 23).

(9) Rankings reduce not increase the scope for innovations in strategy, curriculum, pedagogy and research: Rankings encourage university executives to concentrate energy on maximising performance and reputation within the established criteria that comprise the rankings, particularly research performance and student selectivity. They raise the opportunity cost of investment in innovations in curriculum and pedagogy, research or new organisational configurations. Such innovations typically take years to be realised in improved performance and longer to feed into reputation. On the whole rankings discourage locally distinctive missions that broaden diversity in the sector.

7. Conclusions

The current global rankings both reflect and generate competitive pressures in higher education. Arguably the formation of a global market of elite research universities is a product not simply of rankings but of globalisation, with extending networks and intensified cross-border mobility (Marginson 2007). Nevertheless the Jiao Tong rankings enhance worldwide competition for HiCi researchers and the next generation and this will tend or is tending to generate price effects (Marginson & van der Wende 2006); while within particular nations, global rankings and global competition have superimposed another layer of competition above the national hierarchy. This has implications for the local standing of both national systems and their institutions. With the newly established comparability between the global elite and the national elite, the effect is to render venerable national universities less attractive and more vulnerable, viz a viz global standards and models much more present and potent. This will stimulate national universities to improve only if they have resources and autonomy.

Other universities will be marginalised, especially in the vocational sector. Only some universities can do well in global rankings but all are pushed towards imitation regardless of the distance they have to travel to succeed. In other words, once annexed to the formation of a worldwide reputational market, the Jiao Tong and *Times Higher* metrics tend to reproduce both global competition and the competitor institutions in the likeness of those metrics themselves. The formative effect is stronger in the case of the Jiao Tong rankings which are the more credible system. In the absence of moves to shore up diversity by other means, policy focus on global research rankings will tend to weaken the standing of non-research institutions and encourage the evolution of more unitary but vertically differentiated systems.

Responses to the problems inherent in holistic rankings developed by Shanghai Jiao Tong University and the *Times Higher* have taken three forms. First, the development of the CHE rankings in Germany and their gathering momentum in Europe and perhaps beyond. Here the CHE rankings generate a different pattern of effects. Because CHE dispenses with a holistic rank ordering of institutions and because they

focus on specific programs and services these comparisons are more likely to generate a dynamic of continuous improvement across all institutions on a win-win basis.

The second response is to establish a typology of institutions which allows for several groupings within which separate comparisons are conducted. This factors in diversity of mission in the form of a limited number of agreed categories. Such a typology is under active discussion in Europe though problems remain unresolved: for example if institutions themselves determine their mission (which is consistent with autonomy) then what can retard academic drift into higher status categories over time (which would tend to conflate categories and weaken diversity)? And while on one hand a group of different league tables does less damage than a single league table, on the other hand the dynamics within each group will be broadly similar to the dynamics of the single table at present. This is what has happened with the top category of National Universities in the annual *US News and World Report* ranking exercise.

The CHE strategy and the typology strategy both modify the single league table but in different ways. Each solves some problems of league tables but not all. The CHE approach, which is the more innovative and transformative, moves away from holistic rankings based on partial measures and weightings. It overcomes the problems of self-reproducing hierarchy: obviation of key activity such as teaching (the CHE data focus extensively on teaching on a program basis), problems of validity when combining weights, the use of holistic rankings as an end in themselves, recycling reputation without reference to performance, context free judgements, and the downward effects of holistic league tables on incentives to improve and innovate. Where there are limitations with the CHE scheme is that it depends on programs and services being comparable across all institutions in a nation or group of nations. This at best obscures and at worst mitigates against diversity of mission and approach. The typology scheme brings diversity back into the picture, which is important, but fails to tackle the other problems. Perhaps the optimum approach to the collection of comparative data would be to combine the two approaches, by organising institutions into a typology with a number of categories and use CHE-style comparisons within each category. The categories could include diversity of language as well as program orientation.

The third response has been to negotiate consensus on principles for rankings and league tables. The Berlin Principles (UNESCO/IHEP 2006), led by UNESCO-CEPES and the Institute for Higher Education Policy in Washington, constitute modest progress along these lines. The Principles as adopted on 18-20 May 2006 are as follows:

Rankings and league tables should:

A. Purposes and goals of rankings:

1. Be one of a number of diverse approaches to the assessment of higher education inputs, processes and outputs
2. Be clear about their purposes and their target groups
3. Recognize the diversity of institutions and take the different missions and goals of institutions into account

4. Provide clarity about the range of information sources for rankings and the messages each source generates
5. Specify the linguistic, cultural, economic, and historical contexts of the educational systems being ranked

B. Design and weighting of indicators

6. Be transparent regarding the methodology used for creating the rankings
7. Choose indicators according to their relevance and validity
8. Measure outcomes in preference to inputs whenever possible
9. Make the weights assigned to different indicators (if used) prominent and limit changes to them

C. Collection and processing of data

10. Pay due attention to ethical standards and the good practice recommendations articulated in these principles
11. Use audited and verifiable data whenever possible
12. Include data that are collected with proper procedures for scientific data collection
13. Apply measures of quality assurance to ranking processes themselves
14. Apply organizational measures that enhance the credibility of rankings

D. Presentation of ranking results

15. Provide consumers with a clear understanding of all the factors used to develop a ranking, and offer them a choice in how rankings are displayed
16. Be compiled in such a way that eliminates or reduces errors in original data, and be organized and published in a way that errors and faults can be corrected.

My own judgement is that this does not take us very far. It is significant that no-one has voiced disagreement with the Berlin Principles as a package but they have not impacted the approach of either the *Times* or Shanghai Jiao Tong University. The bottom line is that they leave intact the potential for reputational rankings.

Of the Principles as listed, 1-7 and 10-16 are OK in themselves but mostly do not come to grips with the problems of League Tables as outlined in section 6 of this paper. Some such as Principle 5 requiring specification of contexts are worthy but impractical. Others such Principle 7 requiring validity and relevance, Principle 11 requiring audited and verifiable data, and Principle 12 requiring scientific procedures, are all too pious and obvious. No one is going to admit to invalid or irrelevant or unscientific data!

Principle 9 is unacceptable because it entrenches support for composite indexes based on aggregated weightings. Principle 8 is counter-productive because some aspects are better measured through input than output data. For example the *capacity* of a basic research system (though not the quality) is probably better measured by the number of dollars spent on research activity and the number of hours of research labour than by publications or patents. Data based on expenditure or hours are more comprehensive.

I suggest that the following Berlin Principles take us forward on the core problems:

- 1) Be one of a number of diverse approaches to the assessment of higher education inputs, processes and outputs [reduce authority of rankings]

- 2) Be clear about their purposes and their target groups [ground measures explicitly in purposes of higher education]
- 3) Recognize the diversity of institutions and take the different missions and goals of institutions into account [factor in diversity]
- 4) Be transparent regarding the methodology used for creating the rankings [be transparent in the process]
- 5) Be compiled in such a way that eliminates or reduces errors in original data, and be organized and published in a way that errors and faults can be corrected [reflexivity in design and application]

To these I would add the following draft 'Marginson Principles!'

- 6) Not be explicitly grounded in whole or part in judgements about reputation
- 7) Not said to be holistic, i.e. summative in relation to institutions, including ranks that utilise combined indicators grounded in arbitrary weightings
- 8) When comparing research and scholarly capacity or performance, use primarily discipline-based measures rather than whole of institution measures
- 9) Enable the comparative data to be tailored according to variation in purpose or mission
- 10) Be managed by independent agents (neither universities, governments nor commercial publishers) for the purposes of collection, processing and analysis of comparative data, and publication of rankings data on the web and in other forms

A better approach to comparison of university performance is possible but we have to want it badly enough. The development of rigorous discipline indicators by Shanghai Jiao Tong University is very good and important and we should encourage this kind of purpose-based and disaggregated comparison, if we are to make room for more broad-based improvement of the global higher education sector. If we are to ensure that the 'rankings game' is not to become an end in itself. Rankings should be not the end in themselves but the means to better higher education and research for students, communities, nations and the world. University rankings are a good servant but a bad master. We need to master university rankings before they are allowed to master us.

Notes

¹ The Newsweek 'rankings' – not rankings in their own right but a scissors and tape combination of part of the Shanghai rankings, part of the *Times Higher* ranking, plus data on library holdings – are closer to Jiao Tong than the *Times*, being largely grounded in research and publication (Newsweek 2006).

References

- Bourdieu, P. (1984) *Distinction: A social critique of the judgment of taste*, transl. R. Nice. Routledge & Kegan Paul, London
- Center for Higher Education Development, CHE (2006) *Study and Research in Germany*. University rankings, published in association with Die Zeit. Accessed on 16 March 2006 at:
<http://www.daad.de/deutschland/studium/hochschulranking/04708.en.html>
- Dill, D. & Soo, M. (2005). Academic quality, league tables, and public policy: a cross-national analysis of university rankings, *Higher Education*. 49, pp. 495-533.
- Dyke, N. van (2005) Twenty years of university report cards, *Higher Education in Europe*, 30 (2), pp. 103-125
- Field, K. (2006). Another Accountability Idea: a new database that would customize college rankings. *Chronicle of Higher Education*. 31 March 2006.
- Frank, R. & Cook, P. (1995). *The winner-take-all society*. The Free Press, New York.
- Held, D. McGrew, A. Goldblatt, D. & Perraton, J. (1999) *Global Transformations: Politics, economics and culture*. Stanford University Press, Stanford.
- Institute for Scientific Information, Thomson-ISI (2006) Data on highly cited researchers, ISIHighlyCited.com. Accessed 15 August 2006 at:
<http://isihighlycited.com/>
- Kirp, D. (2004) *Shakespeare, Einstein and the Bottom-line: The Marketing of Higher Education*. Harvard University Press, Cambridge MA
- Liu, N. and Cheng, Y. (2005) The academic ranking of world universities, *Higher Education in Europe*, 30 (2), pp. 127-136
- Liu, N. and Liu, L. (2005) University rankings in China, *Higher Education in Europe*, 30 (2), pp. 217-227.
- Linguasphere Observatory (2006) *Linguasphere Table of the World's Major Spoken Languages 1999-2000*. Accessed on 2 April 2006 at
<http://www.linguasphere.org/language.html>
- Marginson, S. (2006) The Anglo-American university at its global high tide, *Minerva* 44, pp. 65-87
- Marginson, S. (2007) Dynamics of national and global competition in higher education, *Higher Education*, 52, pp. 1-39.
- Marginson, S. & van der Wende, M. (forthcoming) *Globalisation and higher education*, draft paper prepared for the OECD Higher Education Futures project
- Newsweek* (2006) The world's most global universities, 21-28 August, pp. 36-69.
- Organization for Economic Cooperation and Development, OECD (2005). *Education at a Glance*. OECD, Paris
- Organization for Economic Cooperation and Development, OECD (2006). *Education at a Glance*. OECD, Paris

- Salmi, J. and Saroyan, A. (2006). *League tables as policy instruments: uses and misuses*. Paper prepared for OECD/IMHE & Hochschulrektorenkonferenz, Germany, Workshop 4-5 December 2006, Bonn. See Organisation for Economic Cooperation and Development, Institutional Management in Higher Education program, at <http://www.oecd.org>
- Shanghai Jiao Tong University Institute of Higher Education, SJTUIHE (2006). *Academic ranking of world universities*. Accessed 1 September 2006 at <http://ed.sjtu.edu.cn/ranking.htm>
- The Times Higher* (2006) World University Rankings, *The Times Higher Education Supplement*. Accessed on 30 November 2006 at: www.thes.co.uk [subscription required]
- Usher, A. & Savino, M. (2006). *A World of Difference: A global survey of university league tables*. Accessed on 2 April 2006 at <http://www.educationalpolicy.org>
- UNESCO/IHEP (2006) International partnership issues groundbreaking principles on ranking of higher education institutions. Accessed on 3 June at http://www.che.de/downloads/Berlin_Principles_IREG_534.pdf
- U.S. News and World Report* (2006). *America's best colleges, 2007 edition*. Washington.
- Williams, R. and Van Dyke, N. (2006). *Rating major disciplines in Australian universities: perceptions and reality*. Melbourne: Melbourne Institute of Applied Economic and Social Research.
- World Bank (2006) *World Bank Data and Statistics*, Accessed 15 August 2006. <http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20399244~menuPK:1192694~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html>