



Rankings are the sorcerer's new apprentice

Michael Taylor^{1,*}, Pandelis Perakakis^{2,3}, Varvara Trachana⁴, Stelios Gialis^{5,6}

¹Institute for Environmental Research and Sustainable Development (IERSD), National Observatory of Athens (NOA), Metaxa & Vas. Pavlou, Penteli, 15236 Athens, Greece

²Department of Psychology, University of Granada, Campus Cartuja, 18071 Granada, Spain

³Laboratory of Experimental Economics, Universitat Jaume I, 12071 Castellón, Spain

⁴Laboratory of Animal Physiology, School of Biology, Aristotle University of Thessaloniki, University Campus, 54124 Thessaloniki, Greece

⁵Department of Geography, University of Georgia, 210 Field St., Athens, Georgia 30601, USA

⁶Hellenic Open University, Par. Aristotelous 18, 26331 Patras, Greece

ABSTRACT: Global university rankings are a powerful force shaping higher education policy worldwide. Several different ranking systems exist, but they all suffer from the same mathematical shortcoming – their ranking index is constructed from a list of arbitrary indicators combined using subjective weightings. Yet, different ranking systems consistently point to a cohort of mostly US and UK privately-funded universities as being the 'best'. Moreover, the status of these nations as leaders in global higher education is reinforced each year with the exclusion of world-class universities from other countries from the top 200. Rankings correlate neither with Nobel Prize winners, nor with the contribution of national research output to the most highly cited publications. They misrepresent the social sciences and are strongly biased towards English language sources. Furthermore, teaching performance, pedagogy and student-centred issues, such as tuition fees and contact time, are absent from the vast majority of ranking systems. We performed a critical and comparative analysis of 6 of the most popular global university ranking systems to help elucidate these issues and to identify some pertinent trends. As a case study, we analysed the ranking trajectory of Greek universities as an extreme example of some of the contradictions inherent in ranking systems. We also probed various socio-economic and psychological mechanisms at work in an attempt to better understand what lies behind the fixation on rankings, despite their lack of validity. We close with a protocol to help end-users of rankings find their way back onto more meaningful paths towards assessment of the quality of higher education.

KEY WORDS: Global university rankings · Higher education · Multi-parametric indices · Greek crisis · Education system · Policy · Perception · Value-based judgement

—Resale or republication not permitted without written consent of the publisher—

INTRODUCTION

In 1797, Goethe wrote a poem called 'Der Zauberlehrling' – the sorcerer's apprentice. In the poem, an old sorcerer leaves his apprentice to do chores in his workshop. Tired of mundanely fetching water, the apprentice enchants a broom to fetch a pail of water using magic in which he is not yet fully trained. Unable to stop the broom from bringing more and

more water, the apprentice splits it into 2 with an axe, but then each piece becomes a new broom. With every swing of the axe, 2 brooms become 4, and 4 become 8. The apprentice is soon in a state of panic, unable to stop the tide. Just when the workshop is on the verge of being flooded, the sorcerer returns. The poem ends with the sorcerer chiding the apprentice and reminding him that 'powerful spirits should only be called by the master himself'. With the passage of

time, the ‘workshop’ has become a ‘cyberlab’—a digital space teeming with data and information. Apprentices have come and gone, and the sorcerer has intervened to ensure that knowledge continues to stably accumulate. But the latest apprentices, ranking agencies, are particularly careless. Their lack of attention to mathematical and statistical protocols has lured them into unfamiliar territory where their numbers and charts are wreaking havoc. The sorcerer may appear elusive but is still exerting a presence—some say on the semantic web, others say in the truth or wisdom hiding behind the 1s and 0s; perhaps the sorcerer is the embodiment of quality itself.

Global university rankings influence not only educational policy decisions, but also the allocation of funds due to the wealth creation potential associated with highly skilled labour. In parallel, the status of academics working in universities is also being ranked via other single parameter indices like citation impact. While it is now widely accepted that the journal impact factor and its variants are unsound metrics of citation impact and should not be relied upon to measure academic status (Seglen 1997, Bollen et al. 2005, 2009), university rankings are only recently being subjected to criticism. Invented by the magazine *US News & World Report* in 1983 as a way to boost sales, global university rankings are now coming under fire from scholars, investigators at international organisations like UNESCO (Marope 2013) and the OECD (2006), and universities themselves. In one significant public relations exercise, a set of guidelines known as ‘the Berlin Principles’ (IREG 2006) were adopted. However, while the Berlin Principles were developed in an attempt to provide quality assurance on ranking systems, they are now being used to audit and approve rankings by their own lobby groups. Criticisms of global rankings have led to regional and national ranking systems being developed to help increase the visibility of local world-class universities that are being excluded. In this paper, we analyse how and why this state of affairs has come about. We discuss why global university rankings have led to a new ordering of the diverse taxonomy of academic institutions and faculties, and how they have single-handedly created a monopoly ruled by Anglo-American universities that is disjointed from the real global distribution of world-class universities, faculties and even courses. We describe what mechanisms we believe are at play and what role they have played in helping create the current hierarchy. We also discuss the effect that the omnipotence of rankings is having on

higher education at the national and local level, and we delve into the underlying problem of why the construction of indices from a multitude of weighted factors is so problematic. ‘Not everything that counts can be counted; not everything that can be counted counts’ (Cameron 1963).¹

The layout of this paper is as follows. First, we compare and contrast 6 of the most popular global university ranking systems in order to provide an overview of what they measure. With published rankings for 2012 as a context, we then analyse the national and continental composition of the top 200 universities as well as the top 50 universities by subject area/faculty, to highlight and explain some important new trends. The next section is devoted to a case study of the 10 yr ranking trajectory of universities in one country, Greece, chosen to investigate the impact of serious economic restructuring on higher education. This is followed by a discussion of some of the possible mechanisms at work, wherein we identify the problems associated with the current approach for constructing multi-parametric ranking indices. Finally, we draw some conclusions and present a protocol that we hope will help end-users of rankings avoid many pitfalls.

RANKMART

The publishing of global university rankings, conducted by academics, magazines, newspapers, websites and education ministries, has become a multi-million dollar business. Rankings are acquiring a prominent role in the policies of university administrators and national governments (Holmes 2012), whose decisions are increasingly based on a ‘favourite’ ranking (Hazelkorn 2013). Rankings are serious business; just how serious, is exemplified by Brazil, Russia, India and China (collectively, the ‘BRIC’ countries). Brazil’s Science without Borders Programme for 100 000 students comprises a gigantic £1.3 billion (US\$2 billion) fund that draws heavily on the *Times Higher Education Ranking* to select host institutions. Russia’s Global Education Programme has set aside an astonishing 5 billion Roubles (US\$152 million) for study-abroad scholarships for tens of thousands of students who must attend a university in the top 200 of a world ranking. India’s Universities Grants Commission recently laid down a

¹See <http://quoteinvestigator.com/2010/05/26/everything-counts-einstein/> for a detailed discussion.

new set of rules to ensure that only top 500 universities are entitled to run joint degree or twinning courses with Indian partners (Baty 2012). On 4 May 1998, *Project 985* was initiated by the Chinese government in order to advance its higher education system. Nine universities were singled out and named 'the C9 league' in analogy to the US 'Ivy League'. They were allocated funding totalling 11.4 billion Yuan Renminbi (US\$1.86 billion) and as of 2010 made a spectacular entry into global university rankings. Of the C9 League, Peking University and Tsinghua University came in at 44th and 48th position, respectively, in the *QS World Universities Ranking* in 2012. Incidentally, these 2 universities have spawned 63 and 53 billionaires according to the Forbes magazine, and China is currently the only country making ranking lists for colleges and universities according to the number of billionaires they have produced.

Rankings can be either national or global in scope. National university rankings tend to focus more on education as they cater primarily to aspiring students in helping them choose where to study. The *US News and World Report* for example, ranks universities in the US using the following percentage-weighted factors: student retention rate (20%), spending per student (10%), alumni donations (5%), graduation rate (5%), student selectivity (15%), faculty resources (20%) and peer assessment (25%). In the UK, *The Guardian* newspaper publishes *The University Guide*, which has many similar indicators, and also includes a factor for graduate job prospects (weighted 17%). Global university rankings differ greatly from national rankings, and place emphasis on research output. Global university ranking agencies justify not focussing on education with claims such as, 'education systems and cultural contexts are so vastly different from country to country' (Enserink 2007, p. 1027) — an admission that these things are difficult to compare. But isn't this precisely one of the factors that we would hope a university ranking would reflect? We shall see that conundrums such as this are abundant in relation to global university rankings.

A logical assessment of global university rankings requires an understanding of (1) what they measure, (2) how they are constructed, (3) how objective they are and (4) how they can be improved, if at all. Since 2003, 6 ranking systems have come to dominate:

(1) the Chinese Jiao Tong University's 2003–2013 Academic Ranking of World Universities ('ARWU Ranking' or 'Shanghai Ranking'; www.arwu.org/)

(2) the British 2010–2013 Times Higher Education World University Rankings (or 'THE Ranking';

www.timeshighereducation.co.uk/world-university-rankings/)

(3) the Taiwanese National Technical University's 2007–2012 Performance Ranking of Scientific Papers for world universities published by the Higher Education Evaluation and Accreditation Council of Taiwan ('HEEACT Ranking'; <http://nturanking.lis.ntu.edu.tw/Default.aspx>)

(4) the US 2004–2013 Quacquarelli-Symonds Top Universities Ranking published by the US News and World Report (or 'QS Ranking'; www.topuniversities.com/)

(5) the Spanish 2004–2013 Webometrics Ranking of World Universities produced by the Cybermetrics Lab, a unit of the Spanish National Research Council (CSIC) (or 'Ranking Web'; <http://webometrics.info>)

(6) the Dutch 2012–2013 Centre for Science and Technology Studies Ranking ('CSTS Ranking' or 'Leiden Ranking'; www.leidenranking.com).

All of these systems aim to rank universities in a multi-dimensional way based on a set of indicators. In Table 1, we tour the 'supermarket aisle' of global university rankings we call *Rankmart*.

We constructed *Rankmart* by collecting free online information provided from the 6 ranking systems above for the year 2012. We focused on the indicators they use and how they are weighted with percentages to produce the final ranking. In addition to ranking universities, the ranking systems also rank specific programmes, departments and schools/faculties. Note that the Leiden Ranking uses a counting method to calculate its ranks instead of combining weighted indicators (Waltman et al. 2012); as such, no percentage weightings are provided. We also have grouped the indicators into the following general categories: (1) teaching, (2) international outlook, (3) research, (4) impact and (5) prestige, in accordance with the rankings' own general classification described in the methodology sections of their websites. This, in itself, presents a problem that has to do with mis-categorisation. For example, it is difficult to understand why the QS ranking system assigns the results of 'reputation' surveys to 'prestige', or why the THE ranking system assigns the survey results to 'teaching' or 'research' categories. In addition, there is obviously great diversity in the methodologies used by the different ranking systems. What this reflects is a lack of consensus over what exactly constitutes a 'good' university. Before delving into macroscopic effects caused by rankings in the next section, it is worth considering exactly what ranking systems claim to measure, and whether there is any validity in the methodology used.

Table 1. *Rankmaria* a comparison of the indicators used by 6 different global university ranking schemes to measure the 'quality' of universities. Values within column headings are universities analysed and universities ranked (**bold**). Indicators are grouped by type (Teaching, International Outlook, Research, Impact and Prestige) according to the ranking systems' own classification. Percentages refer to the weighting of each given indicator used by each ranking system. S&SSCI is the combined science and social science citation index published by Thomson-Reuters. The h-index is the Hirsch index or Hirsch number

Indicator	THE (2012/2013) (%)		ARWU (2011) (%)		HEEACT (2011) (%)		QS (2012/2013) (z-scores)		Webometrics (log-normal)		Leiden (2013) (%)	
	2000	500+	3000	500	3500	500	2000	400	17000	6000	500	%
Teaching 1	Lecturers to student ratio	4.50	Alumni Nobel prizes & Fields medals	10.00			Faculty to student ratio	20.00				
Teaching 2	Academic reputation survey	15.00										
Teaching 3	PhD to BSc ratio	2.25										
Teaching 4	PhDs per discipline	6.00										
Teaching 5	Institutional income to staff ratio	2.25										
International outlook 1	International to domestic student ratio	2.50					International to domestic student ratio	5.00				Proportion of international collaborative publications
International outlook 2	International to domestic staff ratio	2.50					International to domestic staff ratio	5.00				Mean geographical collaboration distance (km)
International outlook 3	Proportion of publications with >1 international co-author	2.50										
Research 1	Academic reputational survey	18.00	Nature & Science articles	20.00	Articles in last 11 years in S&SSCI	10.00						Proportion of inter-institutional collaborative publications
Research 2	Research income (scaled)	6.00	Articles in S&SSCI	20.00	Articles in current year in S&SSCI	10.00						Proportion of collaborative publications with industry
Research 3	Articles per discipline	6.00										
Impact 1	Citation impact (normalised average citation per article)	30.00	Highly cited researchers	20.00	Citations in last 11 years in S&SSCI	10.00	SCIVERSE SCOPUS Citations per faculty	20.00	Webpages in the main web domain	16.66	Proportion of 10% of highly cited articles	
Impact 2					Citations in last 2 years in S&SSCI	10.00			Deposited self-archived files	16.66	Mean citation score	
Impact 3					Average citations in last 11 years in S&SSCI	10.00			Articles in top 10% of highly cited articles	16.66	Mean citation score (field normalised)	
Impact 4									External in-links to University	50.00		
Prestige 1	Industry income	2.50	Faculty Nobel prizes & Fields medals	20.00	h-index of the last 2 years	20.00	Academic reputation survey	40.00				
Prestige 2			Academic performance with respect to size of institution	10.00	Number of highly cited articles in last 11 years	15.00	Employer reputation survey	10.00				
Prestige 3					Number of articles in current year in high impact journals	15.00						

'The Berlin Quarrel' and methodological faux pas

Global university rankings aim to capture a complex reality in a small set of numbers. This imposes serious limitations. As we shall see, the problem is systemic. It is a methodological problem related to the way the rankings are constructed. In this paper, we will not therefore assess the pros and cons of the various rankings systems. This has been done to a very high standard by academic experts in the field and we refer the interested reader to their latest analyses (Aguillo et al. 2010, Hazelkorn 2013, Marope 2013, Rauhvargers 2013). Instead, here we focus on trying to understand the problems associated with multi-parametric ranking systems in general.

No analysis would be complete without taking a look at how it all began. The first global university ranking on the scene was the ARWU Ranking, which was introduced in 2003. Being the oldest, it has taken a severe beating and has possibly been over-criticized with respect to other systems. A tirade of articles has been published that reveal how the weights of its 6 constituent indicators are completely arbitrary (Van Raan 2005, Florian 2007, Ioannidis et al. 2007, Billaut et al. 2009, Saisana et al. 2011). Of course, this is true for all ranking systems that combine weighted parameters. What makes the ARWU Ranking unique is that its indicators are based on Nobel Prizes and Fields Medals as well as publications in *Nature* and *Science*. One problem is that prestigious academic prizes reflect past rather than current performance and hence disadvantage younger universities (Enserink 2007). Another problem is that they are inherently biased against universities and faculties specialising in fields where such prizes do not apply. The fact that Nobel Prizes bagged long ago also count has led to some interesting dilemmas. One such conundrum is 'the Berlin Quarrel', a spat between the Free University of West Berlin and the Humboldt University on the other side of the former Berlin Wall over who should take the credit for Albert Einstein's Nobel Prizes. Both universities claim to be heirs of the University of Berlin where Einstein worked. In 2003, the ARWU Ranking assigned the pre-war Einstein Nobel Prizes to the Free University, causing it to come in at a respectable 95th place on the world stage. But the next year, following a flood of protests from Humboldt University, the 2004 ARWU Ranking assigned the Nobel Prizes to Humboldt instead, causing it to take the 95th place. The disenfranchised Free University crashed out of the top 200. Following this controversy, the ARWU Ranking solved the problem by removing both universities from their rankings (Enserink 2007).

Out of the 736 Nobel Prizes awarded up until January 2003, some 670 (91%) went to people from high-income countries (the majority to the USA), 3.8% to the former Soviet Union countries, and just 5.2% to all other emerging and developing nations (Bloom 2005). Furthermore, ranking systems like the ARWU Ranking rely on *Nature* and *Science* as benchmarks of research impact and do not take into account other high-quality publication routes available in fields outside of science (Enserink 2007). The ARWU Ranking is also biased towards large universities, since the academic performance per capita indicator depends on staff numbers that vary considerably between universities and countries (Zitt & Filliatreau 2007). The ARWU Ranking calculations favour large universities that are very strong in the sciences, and from English-language nations (mainly the US and the UK). This is because non-English language work is published less and cited less (Marginson 2006a). A second source of such country bias is that the ARWU Ranking is driven strongly by the number of Thomson/ISI 'HighCi' (highly-cited) researchers. For example, there are 3614 'HighCi' researchers in the US, only 224 in Germany and 221 in Japan, and just 20 in China (Marginson 2006a). It is no surprise that in the same year global university rankings were born; UNESCO made a direct link between higher education and wealth production: 'At no time in human history was the welfare of nations so closely linked to the quality and outreach of their higher education systems and institutions' (UNESCO 2003).

As *Rankmart* shows, different ranking systems use a diverse range of indicators to measure different aspects of higher education. The choice of indicators is decided by the promoters of each system, with each indicator acting as a proxy for a real object because often no direct measurement is available (e.g. there is no agreed way to measure the quality of teaching and learning). Each indicator is also considered independently of the others, although in reality there is some co-linearity and interaction between many if not all of the indicators. For example, older well-established private universities are more likely to have higher faculty:student ratios and per student expenditures compared with newer public institutions or institutions in developing countries. The indicators are usually assigned a percentage of the total score, with research indicators in particular being given the highest weights. A final score that claims to measure academic excellence is then obtained by aggregating the contributing indicators. The scores are then ranked sequentially.

With such high stakes riding on university rankings, it is not surprising that they have caught the attention of academics interested in checking the validity of their underlying assumptions. To perform a systemic analysis, we refer to *Rankmart* again and group indicators by what they claim to measure. This allows us to assess the methodological problems, not of each ranking system, but of the things they claim to measure. Our grouping of methodological problems is as follows: (1) research prestige, (2) reputation analysis via surveys and data provision, (3) research citation impact, (4) educational demographics, (5) income, (6) web presence.

Methodological problem 1: Research prestige measures are biased

This is only relevant to the ARWU Ranking, which uses faculty and alumni Nobel Prizes and Fields Medals as indicators weighted at 20% and 10% of its overall rank, respectively. A further 20% comes from the number of *Nature* and *Science* articles produced by an institution. In the previous section, we discussed how huge biases and methodological problems are associated with using science prizes and publications in the mainstream science press. On this basis, 50% of the ARWU Ranking is invalidated.

Methodological problem 2: Reputational surveys and data sources are biased

Rankings use information from 4 main sources: (1) government or ministerial databases, (2) proprietary citation databases such as Thompson-Reuters Web of Knowledge and ISI's Science and Social Science Citation Index (S&SSCI), Elsevier's SCOPUS and Google Scholar, (3) institutional data from universities and their departments and (4) reputation surveys (by students, academic peers and/or employers; Hazelkorn 2013). However, the data from all of these sources are strongly inhomogeneous. For example, not all countries have the same policy instruments in place to guarantee public access and transparency to test the validity of governmental data. Proprietary citation databases are not comprehensive in their coverage of journals, in particular open-access journals, and are biased towards life and natural sciences and English language publications (Taylor et al. 2008). The data provided by institutions, while generally public, transparent and verifiable, suffer from what Gladwell (2011, p. 8) called 'the self-fulfilling

prophecy' effect: many rankings rely on universities themselves to provide key data – which is like making a deal with the devil.

There are many documented cases of universities cheating. For example, in the *US News & World Report* rankings, universities started encouraging more applications just so they could reject more students, hence boosting their score on the 'student selectivity' indicator (weighted 15%; Enserink 2007). Systems like the THE Ranking are therefore subject to a biasing positive feedback loop that inflates the prestige of well-known universities (Ioannidis et al. 2007, Bookstein et al. 2010, Saisana et al. 2011).

Reputational surveys, a key component of many rankings, are often opaque and contain strong sampling errors due to geographical and linguistic bias (Van Raan 2005). Another problem is that reputation surveys favour large and older, well-established universities (Enserink 2007) and that 'reputational rankings recycle reputation'. Also, it is not specified who is surveyed or what questions are asked (Marginson 2006b). Reputation surveys also protect known reputations. One study of rankings found that one-third of those who responded to the survey knew little about the institutions concerned apart from their own (Marginson 2006a). The classical example is the American survey of students that placed Princeton law school in the Top 10 law schools in the country. But Princeton did not have a Law school (Frank & Cook 1995). Reputational ratings are simply inferences from broad, readily observable features of an institution's identity, such as its history, its prominence in the media or the elegance of its architecture. They are prejudices (Gladwell 2011). And where do these kinds of reputational prejudices come from? Rankings are heavily weighted by reputational surveys which are in turn heavily influenced by rankings. It is a self-fulfilling prophecy (Gladwell 2011). The only time that reputation ratings can work is when they are one-dimensional. For example, it makes sense to ask professors within a field to rate other professors in their field, because they read each other's work, attend the same conferences and hire one another's graduate students. In this case, they have real knowledge on which to base an opinion. Expert opinion is more than a proxy. In the same vein, the extent to which students chose one institute over another to enhance their job prospects based on the views of corporate recruiters is also a valid one-dimensional measure.

Finally, there are also important differences in the length of the online-accessible data record of each institution and indeed each country due to variation

in age and digitization capability. The volume of available data also varies enormously according to language. For example, until recently, reputation surveys were conducted in only a few languages. The THE Ranking only now has starting issuing them in 9 languages. Issues such as these prevent normalisation to remove such biases. National rankings, logically, are more homogeneous. Institutional data and in particular reputational surveys provided by universities themselves are subject to huge bias as there is an obvious conflict of interest. Who can check allegations of 'gaming' or data manipulation? It has been suggested that proxies can avoid such problems, for example, with research citation impact replacing academic surveys, student entry levels replacing student selectivity; faculty to student ratios replacing education performance, and institutional budgets replacing infrastructure quality. However, even in this case, what is the guarantee that such proxies are independent, valid measures? And, more importantly, the ghost of weightings comes back to haunt even proxies for the real thing.

For example, the THE Ranking assigns an astonishing 40% to the opinions of more than 3700 academics from around the globe, and the judgement of recruiters at international companies is worth another 10%, i.e. together, they make up half of the total ranking. However, when researchers from the Centre for Science and Technology Studies (CWTS) at Leiden University in the Netherlands compared the reviewers' judgements with their own analysis based on counting citations to measure scholarly impact, they found no correlation whatsoever.

Two ranking systems place large emphasis on reputational surveys: (1) QS (faculty reputation from a survey of peers = 40%, institutional reputation from a survey of employers = 10%); (2) THE (faculty reputation from a survey of peers = 18%, faculty teaching reputation from a survey of peers = 15%). The methodological problems described above mean that 50% of the QS Ranking and 33% of the THE Ranking are invalidated.

Methodological problem 3: Research citation impact is biased

The ranking systems assess research citation impact by the volume of published articles and the citations they have received. All ranking systems use data provided by Thomson-Reuters' S&SCI reports with the exception of the QS Ranking that uses data provided by Elsevier's SCOPUS Sciverse. Some

rankings use the total number of articles published by an institution (ARWU: 20%), normalised by discipline (THE: 6%), the fraction of institutional articles having at least one international co-author (THE: 2.5%, Leiden), or the number of institutional articles published in the current year (HEEACT: 10%) or the last 11 yr (HEEACT: 10%). Citations are used by all ranking systems in various ways:

- HEEACT (last 11 yr = 10%, last 2 yr = 10%, 11 yr average = 10%, 2 yr h-index = 20%, 'HighCi' articles in last 11 yr = 15%, current year high impact factor journal articles = 15%)
- ARWU (number of 'HighCi' researchers = 20%, per capita academic performance = 10%)
- QS (citations per faculty = 20%)
- THE (average citation impact = 30%)
- WEBOMETRICS ('HighCi' articles in top 10% = 16.667%)
- Leiden ('HighCi' articles in top 10%, mean citation score and normalised by field, fraction of articles with >1 inter-institutional collaboration, fraction of articles with >1 industry collaboration).

However, serious methodological problems underpin both counts of articles and citations. Journal impact factors have been found to be invalid measures (Seglen 1997, Taylor et al. 2008, Stergiou & Lessenich 2013 this Theme Section).

Citation data themselves are considered only to be an approximately accurate measure of impact for biomedicine and medical science research, and they are less reliable for natural sciences and much less reliable for the arts, humanities and social science disciplines (Hazelkorn 2013). While some ranking systems normalise citation impact for differences in citation behaviour between academic fields, the exact normalisation procedure is not documented (Waltman et al. 2012). It has also been demonstrated that different faculties of a university may differ substantially in their levels of performance and hence it is not advisable to draw conclusions about subject areas/faculty performance based on the overall performance of a university (López-Illescas et al. 2011). Subject area bias also creeps in for other reasons. The balance of power shifts between life and natural science faculties and arts, humanities and social science faculties. Since citations are based predominantly on publications in English language journals, and they rarely acknowledge vernacular language research results, especially for papers in social sciences and humanities. The knock-on effect is disastrous as faculty funding is diverted from humanities and social sciences to natural and life sciences to boost university rankings (Ishikawa 2009).

It is therefore obvious that another significant source of bias is due to linguistic preference in publications. For example, 2 recent publications have shown a systematic bias in the use of citation data analysis for the evaluation of national science systems (Van Leeuwen et al. 2001) with a strong negative bias in particular against France and Germany (Van Raan et al. 2011). The emergence of English as a global academic language has handed a major rankings advantage to universities from nations whose first language is English. Furthermore, many good scholarly ideas conceived in other linguistic frames are not being recognized globally (Marginson 2006a). Approximately 10 times as many books are translated from English to other languages as are translated from other languages into English (Held et al. 1999). Here, the global reliance on English is driven not by the demographics of language itself but by the weight of the Anglo-American bloc within global higher education, the world economy, cultural industries and the Internet (Marginson 2006c). In terms of demographics, English is only one of the languages spoken by a billion people; the other is Pū-tōnghua ('Mandarin' Chinese). In addition, 2 pairings of languages are spoken by more than half a billion people: Hindi/Urdu and Spanish/Portuguese (Marginson 2006a). Global rankings must account for the bias due to the effect of English (Marginson 2006a). An interesting measure introduced by the Leiden Ranking is the mean geographical collaboration distance in km. We are not clear how this relates to quality but are interested nonetheless in what it is able to proxy for. These examples give a flavour of some of the difficulties encountered when trying to perform quantitative assessments of qualitative entities like quality or internationality for that matter (Buela-Casal et al. 2006).

It is also wrong to assume that citations reflect quality in general. For example, there are a number of interesting cases of highly-cited articles that create huge bias. Some issues:

(1) Huge authorship lists can include several researchers from the same institute; for example, 2932 authors co-authored the ATLAS collaboration paper announcing the discovery of the Higgs boson (Van Noorden 2012). The geographical distribution of authors in very high-profile articles like this can create a large biasing effect at the departmental, institutional and country level.

(2) Retracted articles do not mean that citations also get subtracted; for example, the anaesthesiologist Yoshitaka Fujii is thought to have fabricated results in 172 publications (Van Noorden 2012). This raises the questions: Should citations to these articles be deducted from institutes where he worked?

(3) Incorrect results can generate citations that later become irrelevant to quality; e.g. by March 2012 as a result of a Twitter open commentary, researchers overturned the 2011 suggestion that neutrinos might travel faster than light. In another case, the 2010 claim that a bacterium can use arsenic in its DNA was also refuted (Van Noorden 2012). Recent efforts like the Reproducibility Initiative that encourages independent labs to replicate high-profile research support the view that this is a crucial issue. Furthermore, it is vital so that science can self-correct. It also requires that results are both visible and openly accessible. In this regard, scholars have started launching open-access journals such as *eLife* and *PeerJ* and are pressuring for the issuing of self-archiving mandates at both the institutional and the national level.

Finally, citations are reliant on the visibility and open access of the hosting publication (Taylor et al. 2008). A particularly dangerous development is that Thomson-Reuters is the sole source of citation data for the majority of the ranking systems. The idea that a single organisation is shaping rankings—and hence policy-making and higher education practices around the world—is not an attractive one (Holmes 2012). A very dark recent development reveals the implication of rankings in ethical and ethnographical decision-making:

In 2010, politicians in Denmark suggested using graduation from one of the top 20 universities as a criterion for immigration to the country. The Netherlands has gone even further. To be considered a 'highly skilled migrant' you need a masters degree or doctorate from a recognised Dutch institution of higher education listed in the Central Register of Higher Education Study Programmes (CROHO) or a masters degree or doctorate from a non-Dutch institution of higher education which is ranked in the top 150 establishments in either the Times Higher Education 2007 list or the Academic Ranking of World Universities 2007 issued by Jiao Ton Shanghai University [sic] in 2007 (Holmes 2012).

The methodological problems described above mean that 80% of the HEEACT Ranking, 30% of the ARWU Ranking, 20% of the QS Ranking, 30% of the THE Ranking, 16.667% of the Webometrics Ranking and potentially the entire Leiden Ranking are invalidated.

Methodological problem 4: Proxies do not imply quality teaching

Two ranking systems include weighted indicators related to educational demographics. They classify

these as measures of 'teaching'. The THE Ranking includes the number of PhDs per discipline (6%) and a number of ratios: lecturers:student (4.5%), PhDs:BScs (2.25%), international:domestic students (2.5%) and international:domestic staff (2.5%)—a sizeable 17.75% of their score. The QS Ranking also includes a number of common ratios: lecturers:student (20%), international:domestic students (2.5%) and international:domestic staff (2.5%)—again, a sizeable 25% of their score. A serious problem is that teaching quality cannot be adequately assessed using student:staff ratios (Marginson 2006a). Furthermore, all ranking systems make the mistake that high-impact staff equates with high-quality teaching. The latter problem is also present in the Webometrics Ranking. Until a quantitative link between citation metrics and/or alt-metrics (download and web-usage statistics) and quality is established and validated, such ratings cannot be relied upon (Waltman et al. 2012).

We have seen that rankings focus disproportionately on research compared to proxies for teaching. This is due to the fact that citation impact data, which are claimed to be a good proxy of research quality, are available (to some extent). Furthermore, rankings fall into the trap of making a huge and erroneous assumption, namely that research quality is an indicator of higher educational performance in general. This simply ignores the fuller extent of higher education activity that includes teaching and learning, the student experience and student 'added value' such as personal development. Moreover, Rankmart shows that research indicators account for 100% of indicators used by the ARWU Ranking, 100% of the HEEACT Ranking, 62.5% of the THE Ranking, 33% of the Webometrics Ranking and 20% of the QS Ranking.

Methodological problem 5: Income measures are biased

Only one ranking system has indicators that relate to income. The THE Ranking assigns 6% to research income, 2.5% to industry income and 2.25% to the ratio of institutional income:staff. The problem is that a key ingredient, viz. the cost of education including tuition fee payments, is not included. Since research and industry income are not likely to be distributed evenly across all subject areas and faculties, such measures will be strongly biased, effectively invalidating 10.75% of the THE Ranking's total score.

Methodological problem 6: Web presence is biased

Ranking systems tend to downgrade smaller specialist institutions, institutions whose work is primarily local and institutions whose work is primarily technical and vocational (this includes even high-quality technical institutes in Finland and Switzerland and some French and German technical universities; Marginson 2006a). Webometrics is the only ranking system that uses alt-metrics to try to make inferences about the 'size' and web presence of universities. For example, it counts the number of webpages in the main web domain (16.667%), the number of deposited self-archived files (16.667%) and importantly, the number of external in-links to each university (50%). Alt-metrics are discussed more later on. For now, we point out that web-based statistics disadvantage developing countries that have poor internet connectivity (Ortega & Aguillo 2009). University and country-level statistical inferences are therefore highly unsound (Saisana et al. 2011).

Time and again, we see that a recurring problem is that of bias. This invalidates large portions of a ranking system's score. We also see that the choice of indicators being used to construct all 6 rankings portrays a single institutional model: that of an English-language research-intensive university, tailored to science-strong courses, and without any guidance on the quality of teaching (Marginson & Van der Wende 2007). With this in mind, we consider some of the missing ingredients.

Missing ingredients

In a display of 'after-thinking', the website for the QS Top Universities Ranking has a link to a page 'for parents' (www.topuniversities.com/parents) that states, 'Sending your child to study abroad can be a nerve-wracking experience. While we can't promise to make it easy, we do our best by providing you with all the information you need to understand what's out there so you can make an informed decision together'. Sounds good? Read on. Under the section 'Finance' on the same webpage, there is information and advice about finding 'best value' universities. Here, 5 important factors 'to consider' include: (1) return on investment (ROI), (2) contact time, (3) campus facilities, (4) cost of living, and (5) course length.

We agree. However, not a single measure or proxy is included in any of the ranking systems for these very important criteria. Instead, parents are directed to Bloomberg Business Week's university league

table of the US colleges with the best ROI and to a Parthenon Group UK publication that ranks the top 30 best-value universities based on graduates' average salaries. Outside the US and the UK, no data are provided. No measure is offered to reflect the time students can expect to have with teaching staff, the number of lectures and the class size. Indeed, critics of rising tuition fees for domestic students have pointed out that since the fee increased from £1000 to £3000 and now stands at £9000 per year in the UK, students have not benefitted from more contact time with teaching staff. This is based on a survey conducted by the Higher Education Policy Institute (HEPI), which also found that students with the lowest number of timetabled classes per week were the most likely to be unhappy with the quality of their course. The QS Ranking website states in regard to campus facilities, 'The quality of the learning environment will make a big difference to your university experience, and could even have an impact on the level of degree you graduate with'.

Yet, none of this is captured by any indicator. With regard to cost of living, parents are directed to Mercer's Cost of Living report; again, no indication of how this compares among institutions. Most shocking of all is what is said in relation to the length of courses: 'There are schemes that allow students to pay per-year or per-semester and the choice of fast-track degrees offered by some universities which allow students to graduate in as little as two years. Completing your degree more quickly could save you a significant amount—in both fees and living expenses'.

What about the quality of this form of condensed education? Once again, no incorporation of this information is available via either a proxy or an indicator. For graduate students, the QS Ranking page for parents points them to the QS World University Rankings by Subject and reminds them that there are differential fees. There is advice on how to apply for a scholarship, support services you can expect from your university as an international student, and even a country guide—which reads like a copy-paste from WikiTravel rather than a serious metric analysis. The 5 steps to help you choose are: (1) use the university ranking, (2) refer to QS stars (research quality, teaching quality, graduate employability, specialist subject, internationalization, infrastructure, community engagement and innovation), (3) read expert commentary (follow articles written by people who know a lot about higher education, and are able to offer advice and speculate about future developments, follow tweets from experts in the sector on Twitter, or keep up-to-date with research on whether

or not universities are set to increase their fees or cut budgets), (4) attend university fairs and (5) research into the location (climate, principal and secondary languages, available music/sports/social scenes—all very important—and totally absent from ranking indicators. The QS Ranking is not alone. None of the ranking systems consider any of these issues that are very much at the forefront of the minds of students and parents.

As we have seen, ranking algorithms rely heavily on proxies for quality, but the proxies for educational quality are poor. In the content of teaching, the proxies are especially poor.

Do professors who get paid more money really take their teaching roles more seriously? And why does it matter whether a professor has the highest degree in his or her field? Salaries and degree attainment are known to be predictors of research productivity. But studies show that being oriented toward research has very little to do with being good at teaching (Gladwell 2011, p. 6).

The notion that research performance is positively correlated to teaching quality is false. Empirical research suggests that the correlation between research productivity and undergraduate instruction is very low (Marginson 2006a). Achievements in teaching, learning, professional preparation and scholarship within the framework of national culture are not captured by global university rankings. There is also a lack of information on student pre-graduation and the alumni experience in general.

State funding is another key missing ingredient from the rankings, especially in the 'Age of Austerity'. While China boosted its spending on science by nearly 12.5%, Canada has slashed spending on the environment and has shut down a string of research programmes including the renowned Experimental Lakes Area, a collection of 58 remote freshwater lakes in Ontario used to study pollutants for more than 40 yr. Even NASA's planetary scientists have held a bake sale to highlight their field's dwindling support (<http://www.space.com/16062-bake-sale-nasa-planetary-science.html>). Spain's 2013 budget proposal will reduce research funds for a fourth consecutive year and follows a 25% cut in 2012 (Van Noorden 2012). India scaled down its historic funding growth to more cautious inflation-level increases for 2012–2013 (Van Noorden 2012), but most concerning of all is that since 2007 Greece has stopped publishing figures on state funding of higher education. In 2007, university and research funding in Greece stood at a miserable 0.6% of the GDP, way below the EU average of 1.9% (Trachana 2013).

Rankings also ignore the social and economic impact of knowledge and technology transfer, the contribution of regional or civic engagement or 'third mission' activities to communities and student learning outcomes, despite these aspects being a major policy objective for many governments and university administrators (Hazelkorn 2013).

Systemic problems are rife and rankings are posing many more problems than solutions offered. Students and parents have high expectations of them: they want to know whether they really measure quality (Hazelkorn 2013). Do they raise standards by encouraging competition or do they undermine the broader mission of universities to provide higher education? How useful are they in choosing where to study? Are they an appropriate guide for job prospects or for employers to use when recruiting?

The hidden hand

Rankings have a compelling power. They rework the seemingly opaque and complex inner workings of academic institutions into a single ordered list of numbers that is instantly understandable. A mystical algorithm does the thinking for you and out pops a ranking with the best on the top and the worst at the bottom. However, given that ranking indicators can never hope to capture the whole spectrum of activity of institutions, how can we be sure that the ones chosen are representative? And what about the weighting used? Who decides and how? Nobody knows. The ranking agencies, like a sorcerer's apprentice, tinker continuously with their ranking formulas — sometimes in response to criticism but more often in the hope of stumbling across a mixture of ingredients that could go into a magical potion for assessing quality. The pervading influence of rankings means that they have a strong restructuring effect on higher education policy. Moreover, this effect varies according to the ranking system used, and its impact is different for different nations, different pedagogical cultures and for different types of institution. When algorithm designers at ranking agencies tweak their indicators and weightings, the new rankings produced can transform the reputations of universities and entire national higher education systems overnight. Marginson (2006a, p. 1) recounts the story of one such fatality:

In 2004 the oldest public university in Malaysia, the University of Malaya [UM] was ranked by the THE-QS Ranking at 89th position. The newspapers in Kuala Lumpur celebrated. The Vice-Chancellor ordered huge banners declaring 'UM a world 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 Times changed the definition of Chinese and Indian students at the University of Malaya from international to national, and the University's position in the reputational surveys which comprise 50% of the overall ranking, also went down. The result was that UM dropped from 89th to 169th position. The University's reputation abroad and at home was in free fall. The Vice-Chancellor was pilloried in the Malaysian media. When his position came up for renewal by the government in March 2006 he was replaced.

The University of Malaya had dropped 80 places without any decline in its real performance. This is not an isolated case that is unique to global university rankings. When financial agencies like Standards & Poor's, Moody's or Fitch downgrade the credit rating of a country, the values of shares on the national stock-market plummet and real people and the real economy suffer without realising that a hidden hand is pulling the strings. Global university rankings have the power to affect not just the prestige of universities but also the cross-border flow of graduate students and faculty, as well as the pattern of investments by global corporations in 'intensive' research fields. All over the world, governments and university administrators are considering research policies that will raise their ranking. Ranking systems clearly influence behaviour. The problem is that once entrenched, such proxies for things as intangible as quality are hard to undo (Marginson 2006a). Do we really want a future where university ranking agencies, like financial agencies, have so much power over the fate of global higher education? If not, we have some serious thinking to do if we are to stop it. The sorcerer's apprentice has a dark streak. In England, for example, (the state that first applied quantitative methods to evaluate teaching) the valuation of educational work was found to have a high human cost associated with it because of the transfer of financial resources towards administering such an evaluation (Gewirtz et al. 1995), and also because of the constant pressure placed on teachers to respond to the requirements of external audits, leading to professional stress, insecurity and even diseases (Woods & Jeffrey 1998, Morley 2003).

'Manufacturing prestige'

An extremely concerning consequence of world university rankings is their impact on national higher education culture. For example, Ishikawa (2009) studied how the omnipotence of rankings is affecting

non-Western, non-English language world-class universities such as those in Japan. Japan had a previously self-sustained, national language-based higher education model with a stratification mechanism to select and produce future leaders and professionals. Japan's higher education system had existed outside the realm of Western higher education power domains, and Western university degrees were of little relevance to the upward mobility in the existing national social ladder. Ishikawa pointed out that the competition brought about by rankings is leveraging unprecedented and unwanted changes in Japan's higher education system. Foreign students coming to Japan, far from benefiting from a Japanese perspective, are likely to find that they will be educated in English and that their studies are simply a mirror-image of those offered at highly ranked US universities, but on Japanese soil. Ishikawa (2009) also described in some detail the process by which Japanese universities were approached by rankers and the legal and logistic problems that domestic universities faced in trying to procure the data requested. Worse still, the published rankings that followed immediately had negative effects on the local higher education ethos and caused discontent. For example, Japanese universities were being placed in low-ranking positions (>100th place) despite the fact that Japan occupies the second highest position only after the US in terms of the volume of academic papers listed in the Thomson Scientific database (King 2007). 'The prevalence of world university rankings suggests the emerging hegemony in higher education in the world today' (Ishikawa 2009, p. 170).

What is striking is that, 'there have been few concerted efforts to discredit the rankings process, which appears to have secured public credibility' (Marginson & Van der Wende 2007, p. 309). In an excellent portrayal of the emerging empire of knowledge, Altbach (2007) talked of a mania to identify world-class universities: universities at the top of a prestige and quality hierarchy. Rankings have created an image of what are the world's best and strongest universities. Out of the top 30 universities on the 2007 lists, the combined number of American and British universities amount to 26/30 and 22/30 for the ARWU Ranking and the THE Ranking, respectively. Universities that usually occupy the top 10 positions include the so-called 'Big Three' (Harvard, Yale and Princeton) as well as 'prestige' or 'elite' colleges in the US (Karabel 2005) and 'Oxbridge' in the UK. They present a powerful image of being on the top of the world and thus function as 'global models' to emulate (Ishikawa 2009). Alliance with such top-tiered uni-

versities is actively sought after by non-American, non-European universities aspiring to cultivate an image as being among this global elite—what Ishikawa calls 'manufacturing prestige' (Ishikawa 2009). Some transnational 'alliances' include joint or double degree programs, accredited offshore institutions and for-profit institutions whose degrees have paid-for validation by principally US and UK institutes, for example: the MIT–National University of Singapore alliance and a recent high-profile deal between Stanford University and UC Berkeley with the King Abdullah University of Science and Technology in Saudi Arabia (Schevitz 2008). Curriculum development is based on the model of top universities with English as the medium of instruction. Linguistic and cultural autonomy is being eroded and leading to extinction of local knowledge. Global rankings demonstrate the reality of a global hierarchy in higher education. They portray the powerful image of the world's top-class universities in a way that drowns out the most competitive domestic counterparts. Students from wealthy backgrounds, who previously would have chosen the leading universities at home, appear to now be going overseas, courted by the promise of future success in the global marketplace, superior academic training and the cultivation of personal connections (Ishikawa 2009). 'As the gap between winners and losers in America grows ever wider—as it has since the early 1970s—the desire to gain every possible edge has only grown stronger' and 'the acquisition of education credentials is increasingly recognized as a major vehicle for the transmission of privilege from parents to child' (Karabel 2005, p. 3).

Playing the ranking game has caused Japanese universities to increase their intake of foreign students, thereby further reducing the access to higher education of the potential domestic pool. Such changes are having important macroscopic effects on the global higher education landscape. In the next section, we identify and try to explain some of these trends.

MACROSCOPIC TRENDS

In an attempt to contribute to the debate on rankings, we collected data for the year 2012 from 6 global university ranking systems that provide free, online data. We analysed both their Top 200 ranking of universities distributed by country, and also their Top 50 ranking by subject area/faculty, in order to compare their level of agreement and to try to identify underlying trends and mechanisms that may explain them.

Table 2. Contribution of each continent to the 'Top' 200 universities in 2012 using 6 common global university ranking systems. Country codes are as per the ISO-3166.1 system

Top 200 by continent	THE 2012	ARWU 2012
Africa	1 (ZA)	0
Asia	20 (CN=6, JP=5, KP=4, IL=3, SG=2)	20 (JP=9, CN=5, IL=4, KP=1, SG=1)
Europe	78 (GB=31, NL=12, DE=11, FR=8, CH=7, SE=5, BE=4)	67 (GB=19, DE=14, NL=8, FR=8, CH=6, SE=5, BE=4, IT=4, DK=3)
North America	84 (US=76, CA=8)	92 (US=85, CA=7)
Oceania	9 (AU=8, NZ=1)	8 (AU=7, NZ=1)
South America	1 (BR)	3 (BR=1, MX=1, RA=1)
Top 200 by continent	HEEACT 2012	QS 2012
Africa	0	1 (ZA)
Asia	23 (JP=8, CN=8, IL=3, KP=2, SG=2)	33 (CN=14, JP=10, KP=6, SG=2, IL=1)
Europe	72 (GB=19, DE=14, NL=9, CH=7, FR=6, IT=6, SE=5, BE=2, DK=2, ES=2)	73 (GB=30, NL=11, DE=11, CH=7, SE=5, BE=5, FR=4)
North America	93 (US=84, CA=9)	63 (US=54, CA=9)
Oceania	7 (AU=7)	9 (AU=8, NZ=1)
South America	2 (BR=1, MX=1)	2 (BR=1, MX=1)
Top 200 by continent	Webometrics 2012	Leiden 2012
Africa	0	0
Asia	17 (CN=12, JP=2, IL=1, SG=1, RU=1)	16 (CN=11, KP=2, SG=2, IL=1)
Europe	57 (DE=13, GB=11, NL=7, SE=6, IT=5, ES=5, CH=4, BE=3, DK=3)	83 (GB=29, DE=15, NL=12, FR=10, CH=7, BE=4, DK=4, IE=2)
North America	101 (US=87, CA=14)	81 (US=77, CA=4)
Oceania	7 (AU=6, NZ=1)	5 (AU=5)
South America	5 (BR=4, MX=1)	0

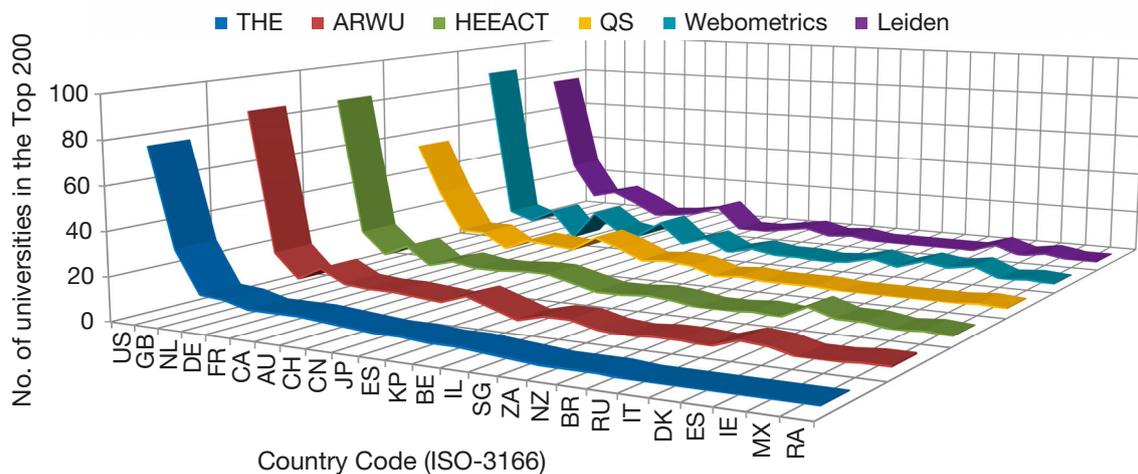


Fig. 1. Distribution of 'Top 200' universities by country, using freely available online data for the year 2012 from 6 different world university ranking schemes

National, continental and subject area composition

Using 2012 as a reference year, we analysed the Top 200 ranking offered by 6 world university ranking systems. Table 2 shows the number of universities per nation contributing to the number of Top 200 universities in each continent using these systems.

Note that we have adopted the ISO-3166.1 country code naming system (https://en.wikipedia.org/wiki/ISO_3166-1). The number of universities that contributed to the Top 200 by country is plotted in Fig. 1.

Again, with 2012 as a reference year, we also analysed the number of universities per nation that contributed to the Top 50 universities by subject area/

Table 3. Comparison of the contribution of each faculty area to the 'Top' 50 universities in 2012 using 4 global university ranking systems for which data were freely and publicly available. Country codes are as per the ISO-3166.1 system

Top 50 by faculty	THE 2012	ARWU 2012
Arts & Humanities	45 (US=25, GB=10, DE=4, CA=3, AU=3)	0
Clinical Medicine, Health & Pharmacy	45 (US=24, GB=7, CA=5, AU=5, NL=2, JP=2)	47 (US=32, GB=7, DE=1, CA=3, CH=2, AU=1, JP=1)
Engineering & Technology/ Computer Sciences	45 (US=24, GB=5, AU=3, KP=3, JP=3, CN=3, CA=2, CH=2)	48 (US=31, GB=4, CN=5, JP=3, CA=2, KP=1, CH=2)
Life & Agriculture Sciences/ Agriculture	45 (US=24, GB=6, CA=4, AU=3, CH=3, JP=2, NL=1, SG=1, SE=1)	47 (US=32, GB=7, CA=3, AU=3, CH=2)
Physical Sciences/Natural Sciences & Mathematics	48 (US=27, GB=6, FR=4, DE=3, CA=2, AU=2, CH=2, JP=2)	48 (US=33, GB=4, JP=3, IL=2, CH=2, FR=2, DE=1, CA=1)
Social Sciences	48 (US=29, GB=8, AU=5, CA=3, CN=3)	50 (US=42, GB=5, CA=3)
Top 50 by faculty	HEEACT 2012	Leiden 2012
Arts & Humanities	0	0
Clinical Medicine, Health & Pharmacy	45 (US=32, GB=4, CA=4, AU=2, DE=1, SE=1, DK=1)	50 (US=36, GB=8, CH=2, IL=1, DE=1, FR=1, IE=1)
Engineering & Technology/ Computer Sciences	45 (US=24, CN=10, GB=3, JP=3, CH=2, CA=1, NL=1, BE=1)	50 (US=30, CN=9, GB=2, CH=2, TR=2, IT=1, DE=1, IR=1, ES=1, AU=1)
Life & Agriculture Sciences/ Agriculture	45 (US=31, GB=6, CA=3, JP=3, NL=1, CH=1, AU=1)	50 (US=34, GB=11, CH=1, FR=1, IL=1, KP=1, AU=1)
Physical Sciences/ Natural Sciences & Mathematics	46 (US=29, JP=5, CN=4, GB=4, FR=2, CH=2)	50 (US=34, GB=3, NL=3, CH=3, DK=2, CN=1, IL=1, SG=1, IE=1)
Social Sciences	47 (US=37, GB=4, NL=4, CA=2)	50 (US=39, GB=6, NL=2, CA=1, SE=1, CH=1)

faculty. The results of our analysis for 4 world university ranking systems making these data available are shown in Table 3. Gathering the data into continental components, a very large skew in the data is apparent for all rating schemes as shown in Fig. 2.

Categorisation of the Top 50 universities by subject area/faculty reveals a shocking finding: Africa and South America are not represented at all. Furthermore, despite their strong economies, Australia and New Zealand ('Oceania') hardly make an impact in the Top 50 of each subject area/faculty. The same is true for the whole of Asia in all subject areas/faculties apart from the field of Engineering and Technology/Computer Sciences, where their impact is representative according to all 4 rating systems. While Europe is seen to be on an equal footing in this subject area/faculty, it clearly dominates over Asia in all of the other subject areas/faculties. The most dominant player is North America, which is massively over-represented in all subject areas/faculties.

It seems that rankings are a reflection of the unequal academic division of labour, the latter resulting from the uneven economic and geographical conditions under which both production of knowledge and education occurs across different countries and regions. The academic inequalities that rankings depict can be better explained when critical theories of unequal socio-spatial development, such as the cumulative-causation effect, are used (Amsler &

Bolsmann 2012). The cumulative-causation theory explains why development 'needs' under-development, in our case how highly-ranked institutions 'need' low-ranked ones (Myrdal 1957). In sharp contrast to neoclassical models of development, cumulative-causation theory shows that economic growth increased rather than closed the gap between wealthier and less-privileged/poor socio-spatial entities. Thus, the majority of less powerful and low-ranked universities follow the economic 'fate' of the countries/regions they belong to and lag even further behind their prevailing and developed counterparts over time. This is partly substantiated by the Greek case studied below.

Another explanation for the enormous inequality of continental representation across all subject areas/faculties lies, as we explain below, in 2 psycho-social mechanisms at work in the selection process of how students together with their parents choose a university.

Mechanisms at work

Mechanism I: The 'blink mentality'

Typing 'Top 10' into Google.com returns 402 million ranked webpages. This is a phenomenal number and reflects a strong social tendency at work in the

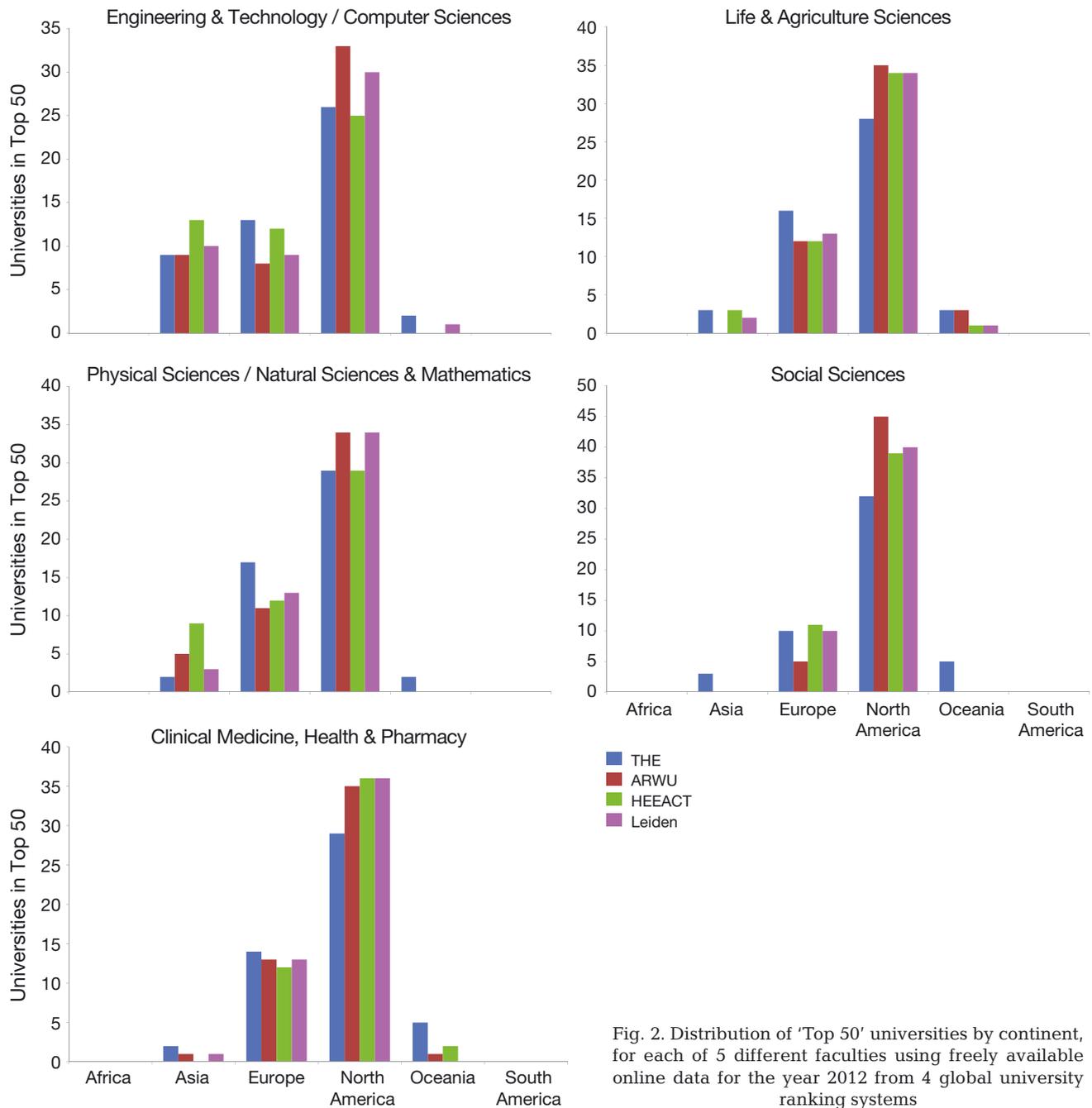


Fig. 2. Distribution of 'Top 50' universities by continent, for each of 5 different faculties using freely available online data for the year 2012 from 4 global university ranking systems

social web. For would-be students of higher education, there are webpages presenting 'top 10 degree subjects for getting a job' and even the 'top 10 universities for joining the super-rich'. Students and their parents act on a range of criteria that influence to some degree their final choice of higher education establishment—tuition fees, the local cost of living, ease of relocation (including returning home during holidays), academic reputation, education quality and post-graduation employment potential. Global

university rankings (and the general global perception they have engendered) measure very little of this overall decision-making process, and yet, they are one of the most important opinion-forming yardsticks that constrain people's final choice. It is worth asking ourselves: Why have we become so fixated and dependent on rankings in our daily lives? According to the book *Blink: the power of thinking without thinking* (Gladwell 2007), one reason is that instantaneous and unconscious decision-making was, and

still is, part of our survival instinct. Being able to react to potential threats and to exploit potential opportunities in the blink of an eye is one reason our species has not yet disappeared off the face of the Earth. Online rankings feed this survival instinct by helping surfers of large volumes of web content to process information rapidly and to make snap decisions in each clickstream.

In the same way that a photographer's choice of frame and moment 'boxes' reality in space and time, our habit of considering only the first 10 or so ranked entries in an ordered list precludes a thorough assessment of all available items. For example, how often do we delve more than 5 pages deep in a Google search before clicking on a result? And how often do we stop to consider what criteria are being used to rank the search results and whether or not they are valid? Without often realising it, our dependence on rankings and the speedy convenience they offer makes us vulnerable to a process which Herman & Chomsky (1988) called 'manufacturing consent'. The ranker, through their choice of ranking system, guides our perception of what is best and what is not.

Mechanism II: 'Preferential attachment'

Acting on perceptions acquired from rankings, parents and their children then become agents in the network of universities, preferentially attaching to those establishments they consider to be the 'best'. If the 25 different countries contributing universities to the Top 200 in Fig. 1 had an equal share, then each would contribute 8 universities and Fig. 1 would be a straight horizontal line. Variance would then be expected to be superimposed on this horizontal line by

small differences in the way the different systems calculate their ranking, and due to small differences year on year—a 'normal' (in the Gaussian sense) distribution. Instead, what is seen is a highly skewed (to the left) rank-frequency distribution. Such distributions are commonplace and arise from the action of preferential attachment in complex networks (Newman 2001, Jeong et al. 2003). Statisticians have even created competition models of the dynamics of university ranking changes year on year (Grewal et al. 2008) and found that with greater than 90% probability, a university's rank will be within 4 units of its rank the previous year. Not only are 'the rich getting richer', their privilege is maintained and stable.

Ranking tables have a compelling popularity, regardless of questions of validity, and of the uses to which the data are put, and of the effects on the organisation of global higher education. Rankings are easily recalled as lists and have quickly become part of 'common-sense' knowledge. Given that global university rankings are a potent device for framing higher education on a global scale, the development of internationally agreed-on principles for good practice is going to be crucial. It is vital that rankings systems are crafted so as to serve the purposes of higher education (Marginson & Van der Wende 2007).

'Greece is a special case'

Last year in an interview on the role of the 'Troika' (the International Monetary Fund, IMF; the European Central Bank, ECB; and the European Commission, EC) in helping struggling economies in the southern periphery of Europe, the head of the IMF, Christine Lagarde, was quoted as saying, 'Greece is a special case'. When it comes to ranking of its uni-

Table 4. Available data for the 'Top' 6 ranked Greek universities. Entires refer to position rank. Year on year statistics for the period 2007–2012 are only available online for the QS Ranking. 2012 data for the ARWU ranking were estimated from online data

University	THE	ARWU	HEEACT	QS					Webometrics	Leiden	
	2012	2012	2012	2012	2011	2010	2009	2008	2007	2012	2012
Aristotle University of Thessaloniki		≈360	347	451–500	451–500		401–450	401–500	430	194	
University of Crete			468	451–500	451–500		401–500	401–500	437	460	
National and Kapodistrian University of Athens		≈300	214	501–550	387		177	200	248	268	
National Technical University of Athens				551–600	551–600		401–450	401–450	356	417	
Athens University of Economy and Business				601+	601+		501–600	501+	472	1269	
University of Patras				601+	551–600		501–550	501+	457	556	

versities, we will see that *Rankmart* paints a paradoxical picture of Greece's higher education institutions over the period 2007–2012. We collected the available online data in Table 4.

The QS Ranking shows that all universities (apart from the National and Kapodistrian University of Athens) slid down to the bottom end of the ranking scale over the period 2007–2012. Despite substantially improving its position in the QS Ranking from 2007–2009, even the prestigious National and Kapodistrian University of Athens suffered the same fate in the years that followed. Somewhat surprisingly, the 2012 rankings provided by HEEACT, Webometrics and ARWU show that all universities with the exception of the University of Crete and the Athens University of Economics and Business are performing better than the QS ranking suggests. *Rankmart* suggests that the difference may be attributable to the fact that the QS ranking is reputational survey-dominated whereas the other ranking systems are citation-dominated. It is possible that, in the absence of expert knowledge, opinion-providers are associating rising government debt and escalating socio-political instability with an expected decline in higher education quality and are downgrading Greek universities as a result. The QS ranking is therefore prone to a perception bias. However, while it is true that Greek higher education is facing extreme pressures (Trachana 2013), when assessed by citation-based metrics, their ranks appear to be slightly cushioned against this biasing effect. It should also be borne in mind that time-lag factors on the scale of 1 to 2 yr frequently exist between the appearance of publications and the citations they accrue. This creates uncertainty in the analysis of data (particularly post-2012), but it is expected that the somewhat longer-term trends described above for the period 2007–2012 are indicative of the expected trend also in the post-2012 period.

In the context of citation impact, Van Noorden (2012) compiled data for the contribution of each country to the top 1% of the highest-cited articles, and ranked the results. Greece is extremely highly ranked in this world list (13th position), ahead of much stronger economies such as Canada and France. In particular, while the US required 311 975 publications to account for its 1.19% share of the most highly-cited articles, Greece achieved an almost equal share (1.13%) with just 9281 published articles. In terms of citation impact, Greece is actually faring surprisingly well in research performance. This points toward a possible explanation for the large deviation in the results of the different ranking

systems in the case of Greece. While funding of its 'real economy' has totally collapsed as a result of the impact of the 2008 Eurozone crisis and the diversion of expenditure towards debt repayment, the QS ranking reflects a perception that Greek universities are in decline, while at the same time the ARWU, HEEACT and Webometrics rankings are reflecting research accomplishment. In most nations, the level of basic research capacity is largely determined by the level of public and philanthropic investment (Marginson 2006a), but in Greece it appears that research quality is being sustained by sheer effort and the international collaborations of Greek academics. Our analysis of the available ranking information for Greek universities suggests that

- There is a very large deviation in the ranking of Greek universities according to the different ranking systems
- There is suggestion of a strong perception bias in the QS ranking
- There is a contradiction between observable trends in the ARWU ranking and QS rankings over the period 2007–2012, which is related to the maintenance of a stable source of citations.

The contradiction between the ARWU and QS ranking of Greek universities echoes the results of a recent comparison of QS and ARWU rankings on the global scale performed by Jöns & Hoyler (2013). Their study revealed that differences in the ranking criteria used produce partial and biased, but also very different and uneven, geographies of higher education on the global scale. The authors suggested that the emergence of these 2 rankings in particular reflects a radical shift in the geopolitics and geo-economics of higher education from the national to the global scale that is prioritising academic practices and discourses in particular places and fields of research. The authors also suggested that these rankings are having a significant public impact in the context of the on-going neo-liberalization of higher education and reveal a wider tension in the global knowledge-based economy:

An examination of different geographical scales and individual ranking criteria provided further evidence that both league tables [world rankings] produce highly partial geographies of global higher education that are to some extent reflective of wider economic and socio-cultural inequalities but also convey a very narrow view of science and scholarship, namely one that can be captured by Anglophone neoliberal audit cultures (Jöns & Hoyler 2013, p. 56).

The case of Greece is no exception to this picture and presents uneven results that depend on which ranking system is applied. While not obvious, there

are signs that successive Greek governments are responding to rankings. In recent years, emphasis has been placed on securing the role of Greek science in European projects in particular. In parallel, there is a trend towards liberalization of private higher education institutions. While the funding tap has effectively been switched off to state universities, more private colleges and for-profit educational institutions have started to appear. There are currently 25 such institutions, but at the moment they are manufacturing prestige by paying for the validation of their degrees by accredited universities and institutions in the US and the UK. On 9 July 2012, the Troika pressed Greek authorities to progress towards full recognition of degrees granted by private colleges operating as franchises of foreign universities in Greece, part of a plan to liberalize the country's higher education sector. This is confirmed by the findings reported by the 2013 annual report 'Greek higher education institutions in the world rankings' (Interuniversity Higher Education Policy Network 2013), which used data from the Webometrics Ranking over the years 2011–2013. The report revealed that Greek private education institutions are starting to be noticed overseas, as evidenced by the fact that they are being evaluated by world rankings like Webometrics in spite of constitutional and other bans (Interuniversity Higher Education Policy Network 2013). For example, out of the 64 Greek institutions or structures appearing in the 2013 Webometrics Ranking, 23 are universities, 16 are Technological Educational Institutions (TEIs), 20 are private structures, 2 are military academies, 2 are schools of music, and 1 is mysteriously classified as 'other'. Despite making up 25/64 (or 39%) of all Greek higher education institutions, only 3 of these private education institutions can be found in the top third of the 21 250 institutions appearing in the 2013 Webometrics Ranking list (Interuniversity Higher Education Policy Network 2013). The authors of the Interuniversity Higher Education Policy Network (2013) report concluded that private higher education institutions in Greece are of low quality compared to their public counterparts:

at the base of this evaluation, the view that these [private education institutions] are low-level educational structures is substantiated, with a tendency towards further deterioration and in every case they have a much worse ranking than the public higher education institutions (Interuniversity Higher Education Policy Network, 2013, p. 17).

Interestingly, the report also compared indicators for Greek higher education institutes with other EU member states of similar GDP and found that Greek

higher education institutions are by no means inferior to those in European countries with GDP comparable to that of Greece. However, the report also recognized that the distance between the better and 'less good' institutions is growing (Interuniversity Higher Education Policy Network 2013), suggesting a change away from a systemic (social democratic) approach towards a neoliberal outlook based on individual institutions, i.e. towards a logic of 'excellence':

The new laws for higher education supported the need for its existence [TEI 'excellence'] and laid the foundations for its legalization not so much in the development and improvement of a satisfactory institution, but in the more generalized discrediting of the Greek public university (Interuniversity Higher Education Policy Network 2013, p. 5).

Summarizing this section, with the passing of each subsequent year from 2007 to 2013, Greek universities appear to follow the cumulative-causation effect described in the section above entitled 'National, continental and subject area composition'. Ranking systems that are reputational survey-based like the QS ranking are imitating the credit agencies' perception of Greece's economy and are driving a devaluation of the status of Greece's state universities. National policy makers are then using these rankings to justify implementing a neoliberal agenda supporting the drive to large-scale privatisation of Greece's universities and research institutes. The paradox is that Greece's research output from public institutions is maintaining its high international quality impact as reflected by citation-based rankings. Ironically, if cost was a strongly weighted indicator in ranking systems, Greece would have a glowing place, since its undergraduate education, and a significant part of its post-graduate education, is still free.

THINKING IT THROUGH

The causes

We have seen how rankings have become popular, hierarchical and entrenched, but that in their current form, they are irreproducible and also invalid measures of the 'quality' of universities. We have argued that one key mechanism behind their popularity is 'blink mentality'. With the advent of the social web where instant one-dimensional ordered lists like rankings have become a simple device for helping end-users make numerous instant decisions, the blink mentality comes to the fore as they process large volumes of data. We also claim that the main

mechanism responsible for the rise of an elite cohort of universities is 'preferential attachment'. This was acutely present in the distribution of the top 200 universities by country. The nature of the blink mentality necessitates that end-users of rankings train their eyes on the top part of the ranking lists, and this selection effect dramatically changes the dynamics of their interaction with the entire network of universities. They are drawn to and tend to reinforce the top universities by preferentially attaching to an elite core. The popularity of rankings and, in particular, the way they are easily interpreted by non-expert policy makers, has created an image that they are the standard against which higher education quality should be benchmarked and assessed. This process has entrenched them on the global scale with sometimes disastrous consequences as we have described. However, we have also seen that all of the ranking systems are invalid to a large degree. There is no consensus on which variables should be used as direct measures or even as proxy indicators. There is also no consensus on how many of them there should be. Ranking systems that aggregate and combine the variables using arbitrary or subjectively-chosen weights cannot be relied upon. The results of the ranking process are inconsistent even when methodological tweaking is kept constant (e.g. by comparing universities during a given year). The results are also not reproducible due to a lack of transparency in the sharing of source data. Another very important problem involves the end-users of rankings. They see rankings as a holistic measure of universities rather than a measure of their research performance in certain fields:

Holistic rankings become an end in themselves without regard to exactly what they measure or whether they contribute to institutional and system improvement. The desire for rank ordering overrules all else. Once a rankings system is understood as holistic and reputational in form it fosters its own illusion (Marginson 2006a, p. 8).

The effects

Global rankings have generated international competitive pressures in higher education. On this they have also superimposed a layer of competition on established national higher education systems (Marginson 2006a). In some countries, like Japan and Greece, the effect has been to make national universities less attractive and also more vulnerable to their own people. Their best clients, the best students, are

increasingly likely to cross the border and slip from their grasp. This is especially damaging in emerging nations that lack capacity in research and communications technologies. While only certain kinds of universities are able to do well in global rankings, other institutions are being pushed towards imitation regardless of their rank.

Effect 1: Globalisation of higher education

Global higher education is changing rapidly for 4 key reasons. Firstly, policy-makers have understood that higher education is a principal source of wealth creation, and that skilled citizens are healthier, more prosperous and have the knowledge to contribute to society throughout their lives. As a result there has been a rapid growth in the global research population and the academic social web. This in turn has led to a large increase in mobility and global knowledge production and dissemination (Hazelkorn 2013). Secondly, participation in 'world-class' research depends on the ability of nations to develop, attract and retain talent. However, despite global population growth, the availability of skilled labour is actually declining. In 2005, young people represented 13.7% of the population in developed countries but their share is expected to fall to 10.5% by 2050 (Bremner et al. 2009). In response, governments around the world are introducing policies to attract the most talented migrants and internationally mobile students, especially postgraduate students in science and technology (Hazelkorn 2013). Thirdly, the quality of individual higher education institutions as a whole (teaching and learning excellence, research and knowledge creation, commercialisation and knowledge transfer, graduate employability) has become a strong indicator of how well a nation can compete in the global economy (Hazelkorn 2013). This has produced a bias in favour of demonstrating value-for-money and investor confidence. Last, but by no means least, students (and their parents) who are paying for higher education as consumers, are selecting institutions and education programmes according to their perception of return on investment, i.e. by balancing the cost of tuition fees and the cost-of-living with career and salary-earning potential. Rankings feed their needs and the needs of many other different stakeholders by apparently measuring the quality of the educational product. Employers use them to rank new recruits by their perceived level and capability. Policy makers and university administrators use them as proxies of performance, quality,

status and impact on the national economy. For the public, they provide an at-a-glance check of the performance and productivity of higher education institutes funded by tax payers' money.

Effect 2: Authoritarian irresponsibility

Lobbying of decision makers by ranking agencies has brought them financial dividends but has damaged the global higher education landscape by creating the perception of 'good' and 'bad' universities. The rankers have let out a beast. They have supported and have reproduced an expanding scale of educational inequality, and are now trying to warn people that the rankings should be handled with care; but the responses by rankers to criticisms of their ranking systems verges on schizophrenia. In one breath they tell us that, 'no university ranking can ever be exhaustive or objective' (Baty 2012), and in another breath that, '[the rankings] can play a role in helping governments to select potential partners for their home institutions and determine where to invest their scholarships' (Baty 2012). Why don't governments cut out the ranking middle-men and simply ask the experts—their own academics? Governments also are contradicting themselves:

We are blessed in Britain to have a world-leading higher education sector. Our universities are a great source of strength for the country and their role—in an increasingly knowledge-based economy—is becoming more and more central to our future prosperity. Universities are also becoming increasingly central to our future social prospects (Social Mobility and Child Poverty Commission 2013, p. 2).

The UK government's Social Mobility and Child Poverty Commission (2013) then goes on to show how the proportion of state-educated pupils attending elite 'Russell Group' universities (the club of UK research-intensive universities that includes Oxford and Cambridge, as well as the London School of Economics and Imperial College) has declined since 2002, and that the UK's 'most academically selective universities are becoming less socially representative' (Wintour & Adams 2013). This is despite 'the growing evidence base that students from less advantaged backgrounds tend to outperform other students with similar A-level grades on their degree' (Wintour & Adams 2013). So which is it? Enabling able students to fulfil their potential is central even to economic efficiency. We agree with the man behind the THE Ranking when he says that,

The authority that comes with the dominance of a given ranking system, brings with it great responsibility... All global university ranking tables are inherently crude as they reduce universities and all their diverse missions and strengths to a single, composite score... One of the great strengths of global higher education is its extraordinarily rich diversity and this can never be captured by any global ranking, which judges all institutions against a single set of criteria. (Baty 2012)

This is the best argument we have heard yet for immediately abandoning his ranking the THE Ranking. There is, as we have seen, absolutely no mathematical basis for arbitrary choices of indicator weights and, as we have also seen, there is not even a consensus over which indicators should be included. In fact, the field of multi-parametric indexing is still in a state of early evolution while researchers work towards a meaningful protocol for their general calculation and evaluation. As such, policy makers have jumped the gun. They should have trusted the advice of academic experts and waited until a representative, reproducible and reliable (the 3 'R's) methodology can be designed. Instead they have chosen to side with an opportunistic ranking lobby. This is power and irresponsibility.

Effect 3: Academic mutinies

In 2007, there was a mutiny by academics in the US. They boycotted the most influential university ranking system in use in the country, the *US News and World Report*, claiming that the ranking is based on a dubious methodology and on spurious data (Butler 2007). Presidents of more than 60 liberal arts colleges refused to participate in providing data for the report, writing that the rankings, 'imply a false precision and authority' and 'say nothing or very little about whether students are actually learning at particular colleges or universities' (Enserink 2007, p. 1026). In 2006, 26 Canadian universities revolted against a similar exercise by Maclean's magazine. Soon after, academics in highly influential universities including MIT, CalTech and Harvard began rebelling against another ranking, the impact factor, and started a challenge to the conventional publishing system that they saw as costly, slow, opaque and prone to bias, and which had erected barriers to access to potential readers (Taylor et al. 2008). They boycotted key culprit journals and began issuing mandates in institutes for academics to self-archive and peer-review each other's articles. The mutiny is still gathering momentum and the fall-out is unforeseeable. If one considers for a moment the near

absence in the top rankings of nations such as Germany, Japan and South Korea that drive world-class economies, it is easy to ask: Who and what are the rankings really representing? There is growing potential for a global mutiny in higher education caused by rankings. The critics take aim not only at the rankings' methodology but also at their undue influence. For instance, some UK employers use them in hiring decisions (Enserink 2007). MIT, Cal-Tech and Harvard did not only rebel against the ranking schemes; they also established *OpenCourseWare*, a platform for providing free, online open access to teaching materials and resources produced by their faculty staff. In one fell swoop they made a huge contribution to knowledge conservation. As we will discuss below, 'cloud' education is vital for the sustainability of higher education in countries suffering from 'brain drain'.

Effect 4: Brain drain and the threat to global knowledge conservation

The brain drain of highly skilled professionals fleeing the crisis in southern Europe is a brain gain for the countries to which they migrate. This has a significant impact on higher education and research at both ends, as universities and research centres must adjust to the new conditions brought about and have to strive to retain potential creators of knowledge and therefore wealth. Most skilled expatriates have the willingness and capacity to contribute to the development of their new host country. The gap they leave behind is sorely felt at home and the new beneficiaries of their talent have a social responsibility to support information and distributed computing technologies that enable distance cooperation. In this way, a digitally literate generation of young people will be able to continue to benefit from the knowledge acquired by their peers. Knowledge can be conserved via virtual classrooms and virtual laboratories, or by remote access to rare or expensive resources of colleagues in host countries and will help keep afloat small, low-budget universities at home by providing them access to the infrastructure and quality of large, foreign ones. Such approaches are needed if a segregation of nations into high and low knowledge capital is to be avoided. Brain drain without knowledge conservation will simply lead to intellectual and cultural extinction of nations (Wilson et al. 2006). The impact of this on the global scale, while not yet fully understood, is irreversible.

Effect 5: Ranking agencies are under pressure

Well aware of their influence, and also the criticisms, the rankers themselves are starting to acknowledge that their charts are not the last word. The ARWU Ranking website has a pop-up window warning that: 'there are still many methodological and technical problems' and urges 'caution' when using the results. In a self-critique of the THE Ranking, Baty (2012) explained that it 'should be handled with care'. In response to the critics, some rankers are also continuously tinkering with their formulas. But this opens them up to another criticism, namely that this invalidates comparisons of the results of their ranking system over different years. They are losing their grip. In another attempt to boost their credibility, big players in the ranking game (UNESCO European Centre for Higher Education in Bucharest and the Institute for Higher Education Policy in Washington, DC) founded the International Rankings Expert Group (IREG) in 2004. During their second meeting, convened in Berlin from 18 to 20 May 2006, they set forth principles of quality and good practice in higher education institute rankings called the 'Berlin Principles' (IREG 2006). They stress factors such as the importance of transparency, relevant indicators and the use of verifiable data. The principles are very general in nature and focus on the 'biggest common denominators' between ranking systems. Sounds good? It turns out that it was simply a public relations exercise. Most of the rankers are still not yet compliant with their own agreed principles (Enserink 2007). In a worrying recent turn, the IREG Executive Committee at its meeting in Warsaw on 15 May 2013 decided to grant QS the rights to use the 'IREG Approved' label in relation to the following 3 rankings: QS World University Rankings, QS University Rankings: Asia and QS University Rankings: Latin America. With all of the systematic problems inherent to *Rankmart's* ranking systems still to be addressed, it is not possible to believe the results of an auditing and approval process. With powerful ranking companies using their own lobbies to approve themselves, this is a dangerous precedent. What is needed, in addition to a truly independent commission free from self-interests, is a clear statement of the purpose of the rankings, as measuring excellence is clearly not the same as measuring the quality of teaching, cost, graduate employability or innovation potential (Saisana et al. 2011). 'It is important to secure "clean" rankings, transparent, free of self-interest, and methodologically coherent - that create incentives to broad-based improvement' (Marginson & Van der Wende 2007, p. 306).

In league tables and ranking systems, ranks are often presented as if they had been calculated under conditions of certainty. Media and stakeholders take these measures at face value, as if they were unequivocal, all-purpose yardsticks of quality. To the consumers of composite indicators, the numbers seem crisp and convincing. Some may argue that rankings are here to stay and that it is therefore worth the time and effort to get them right. Signs are that this is incorrect. Rankings are coming under mounting pressure. As their negative effects take root, resistance to them will grow, and new and better modes will be developed. This is what happened with journal impact factors. Resistance to them gave birth to open access, open source, open repositories, *OpenCourseWare* and institutional mandates for self-archiving. Another important initiative is the San Francisco Declaration on Research Assessment (DORA; <http://am.ascb.org/dora/>), produced by the American Society for Cell Biology (ASCB) in collaboration with a group of editors and publishers of scholarly journals:

There is a pressing need to improve the ways in which the output of scientific research is evaluated by funding agencies, academic institutions, and other parties... [there exists]

- the need to eliminate the use of journal-based metrics, such as Journal Impact Factors, in funding, appointment, and promotion considerations;
- the need to assess research on its own merits rather than on the basis of the journal in which the research is published.

The jury is still out on the construction of multi-parametric indices. When it comes back in, be ready for a surprise.

The solutions

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 is supreme. There could be a large range of possible global models (Marginson 2006a). In a very recent report for UNESCO, Hazelkorn (2013, p.1) asked the question: ‘Should higher education policies aim to develop world-class universities or should the emphasis be on raising standards worldwide, i.e. making the entire system world-class?’

Her report ends by explaining very clearly that the development of an elite group of world-class universities is ‘the neoliberal model’ while the development of a world-class system of higher education is ‘the

social democratic model’. We have seen that the current state of affairs and the host of harmful effects produced by the rampant rankings of apprentice ranking agencies is due to the former. As with the US academic mutiny of 2007, the problems are giving birth to new solutions that have the potential to put social democracy back on the map.

The Leiden Ranking is the only ranking system in *Rankmart* that does not aggregate weighted indicators. Instead, it uses a simple counting method for calculating ranks based on the rank of each university according to each indicator treated separately (Waltman et al. 2012). Its indicators are also correctly normalised to account for differences existing between subject areas/faculties, and it does not rely on data supplied by the universities themselves and avoids potential bias associated with self-inflationary policies. However, like the HEEACT Ranking, the Leiden Ranking focuses exclusively on research performance, and it limits these data to only journals listed in the Web of Science. This is problematic due to the existence of a power law associated with journal impact factors (Taylor et al. 2008). The Leiden Ranking also assesses universities as a whole and therefore cannot be used to draw conclusions regarding the performance of individual research groups, departments or institutes within a university. It also does not attempt to capture the teaching performance of universities. In a very welcome step, the designers of the ranking themselves admit that the ranking focuses exclusively on universities’ academic performance, and we give them credit for noting that scientific performance need not be a good predictor of its teaching performance (Waltman et al. 2012). Furthermore, other ‘input’ variables such as the number of research staff of a university or the amount of money a university has available for research, are not taken into account and again, the authors of the ranking explain that this is due to a lack of accurate internationally standardized data on these parameters (Waltman et al. 2012).

A second interesting development is from the Centre for Higher Education Development (CHE) in Gütersloh, Germany (www.che.de), which assesses German university departments and assigns them to more ‘fuzzy’ top, middle and lower tiers. It also allows the user to choose which indicators they want to use to sort universities. CHE collects survey data from faculty and students. These data are used to generate specific comparisons between institutions of research, teaching and services within each separate discipline and function. The CHE data are provided via an interactive web-enabled database that permits

each student to examine and rank identified programmes and/or institutional services based on their own chosen criteria (CHE 2006). The idea of basing the classification on fuzzy sets and what we call 'interactive ranking' is something we strongly support. While the CHE data themselves are subject to the same problems described in the section on *Rankmart*, it does offer a new and less problematic methodology. Those interrogating the data decide themselves which indicators meet their desired objectives; this is healthy acknowledgement that 'quality is in the eye of the beholder'. The CHE rankings dispense with the holistic rank ordering of institutions, noting that there is no 'one best university' across all areas (Marginson 2006a).

The real problems start to arise when rankings try to be both heterogeneous and comprehensive at the same time. The attraction of a single score judging between entities that would otherwise be impossible to compare is logical, but statistically-flawed. 'A ranking can be heterogeneous (universally applicable) as long as it doesn't try to be too comprehensive (including all possible variables); it can be comprehensive as long as it doesn't try to measure things that are heterogeneous' (Gladwell 2011, p. 3).

For several years, Jeffrey Stake, a professor at the Indiana University law school, has run a Web site called '*the Ranking Game*' (<http://monoborg.law.indiana.edu/LawRank/play.shtml>). It contains a spreadsheet loaded with statistics on every law school in the US and allows users to pick their own criteria, assign their own weights and construct any ranking system they want. Stake's intention is to demonstrate just how subjective rankings are, to show how determinations of 'quality' rely on arbitrary judgments about how different variables should be weighted (Gladwell 2011). Rankings, it turns out, are full of implicit ideological choices, like not including the price of education or value for money, or efficacy (how likely you are to graduate with a degree): 'Who comes out on top, in any ranking system, is really about who is doing the ranking' (Gladwell 2011, p. 13).

It is vital that rankings systems are crafted so as to serve the purposes of higher education, rather than purposes being reshaped as an unintended consequence of rankings (Marginson & Van der Wende 2007). Gladwell (2011) eloquently expressed the difficulty of creating a comprehensive ranking for something as complex and multi-faceted as higher education. The reality is that the science of how to construct multi-parametric indices is still embryonic and disputable. Multi-parametric indices reflect the

diachronic, ontological and epistemological tension between the need of simplification and quantification on the one hand, and the apparent integrative and qualitative character of phenomena on the other. The limited socio-economic interests commonly represented through the ranking systems and the fact that they tell us little in relation to social exclusion and stratification should also be taken into account. In any case, ranking systems should be subject to theoretically informed analysis and discussed within a wider framework that encompasses society, economy, restructuring and the underlying uneven forces that determine such procedures. If analysed from a critical realist perspective, it is possible perhaps that rankings may be able to offer estimates that are helpful for comparative analyses between different academic systems and environments. Having said that, intricate and socially and historically determined procedures such as knowledge production, academic performance and teaching quality, are not easily reducible to simple quantitative groupings and categorisation. This was also verified by Bollen et al. (2009), who, in the context of research impact, performed a principal component analysis of 39 scientific and scholarly impact measures and found that impact is a multi-dimensional construct that cannot be adequately measured by any single indicator – in particular citation counts. Their analysis included social network statistics and usage log data, 2 important alt-metrics parameters.

The semantic web, armed with metadata harvesting tools, provides a rich source of new information that should also be harnessed to support the assessment of higher education. Our belief is that alt-metrics have a central role to play. Perhaps a system like that proposed by the CHE in Gütersloh or the new U-multirank (www.u-multirank.eu/) can be expanded to include indicators related to the real educational experiences of students, i.e. how they perceive their educational training, the quality of contact time, tuition fees, the cost of living, future job prospects and social development. Such things can even be crowdsourced to students who can rate universities and the departments (and even staff) with which they have experience. This source of collective wisdom is extremely valuable and should be given more prominence. In the context of research assessment, alt-metrics are gaining in popularity as early indicators of article impact and usefulness. Why is this so? The fundamental reason is that traditional citations need time to accrue and hence they are not the best indicator of important recently published work. Alt-metrics appear more rapidly. A typical article is likely to be most

tweeted on its publication day and most blogged within a month or so of publication (Thelwall et al. 2013). The fact that citations take typically over a year to start accumulating casts into doubt the annual validity of rankings that rely heavily on citation counts. Social media mentions, being available immediately after publication (and even before publication in the case of preprints), offer a more rapid assessment of impact. Evidence of this trend was published in a paper last year by Shuai et al. (2012), who analysed the online response to 4606 articles submitted to the preprint server arXiv.org between October 2010 and May 2011. The authors found that Twitter mentions have shorter delays and narrower time spans and their volume is statistically correlated with the number of arXiv downloads and also early citations just months after the publication of a preprint. Brody et al. (2006) described how early web usage statistics can be used as predictors of later citation impact, and Taraborelli (2008) outlined how social software and distributed (bottom-up) methods of academic evaluation can be applied to a form of peer-review he called 'soft' peer-review. In the context of crowdsourcing student opinion about the quality of their higher education, alt-metrics appear to have the potential to be fast, transparent and more objective. More evidence supporting crowdsourcing comes from a recent paper published in the journal PLOS One that developed a method to derive impact scores for publications based on a survey of their quantified importance to end-users of research (Sutherland et al. 2011). The process was able to identify publications containing high-quality evidence in relation to issues of strong public concern. We see alt-metrics as a welcome addition to the toolbox for assessing higher education quality.

Universities can be very complex institutions with dozens of schools and departments, thousands of faculty members and tens of thousands of students. In short, as things stand: 'There's no direct way to measure the quality of an institution—how well a college manages to inform, inspire, and challenge its students' (Gladwell 2011, p. 6).

CONCLUSIONS

Picking a university to attend can be a harrowing experience for students and their parents, and can resemble an exercise in e-shopping. Rankings appear to offer a quick-fix solution. However, as we have shown, the rankings have been produced by apprentices. They are a quick-fix precisely because, like the sorcerer's apprentice tired of doing chores,

the ranking agencies are not skilled in the handling of multi-dimensional data and have come up with a simpleton's 'solution'. They have let out a beast that has been reproducing and exaggerating global higher education inequality. The inequalities produced by global economics and production, and academic hierarchy are then reproduced by the ranking system, which serves as a commercial and stratifying tool at the service of global academic competition. This system is now threatening to destroy not just national higher education systems, but also the sustainability of local and cultural wisdom. Metrics have made their way into many spheres of academia, and the same old problems keep resurfacing, viz. the inaccuracy of single-parameter metrics to capture holistic features as well as inconsistencies between various multi-parametric indices. Taylor et al. (2008) argued that the citation-based impact factor was responsible for creating and exaggerating a journal monopoly that laid siege to science. Lessons seem not to have been learned. Global university rankings are doing to universities and the higher education sector what the journal impact factor did to tenure evaluation and the research sector.

Rankings are not representative of the average quality of education at a university or faculty, nor are they representative of the average quality of researchers at these institutions. In some of the global university ranking systems, the academic establishments of entire continents (like Africa or South America) are not even present in the top 200. We have seen also how the brain *gain* of the top 200 universities and their host countries is a brain *drain* for the rest of the world. The Global Financial Crisis of 2008 and the on-going Eurozone crisis is testament to this, as we saw how the case of Greece brings to light many of the contradictions. The Greek drama unfolding will sooner or later have a negative impact upon universities' performance, a fact that would be expected to correlate with them descending the ranks. Yet, rankings themselves are misleading as they do not measure many of the aspects highlighted here (such as free tuition), and will continue to mislead as long as they are constructed in the current manner. Essentially, as the cumulative-causation effect explains, as Greece is devaluated, the country's universities are also expected to devalue, with neoliberal rankings re-enforcing and accelerating this devaluation process. This is a huge warning bell of possible future trends in other nations like Cyprus, Spain, Portugal, Italy and Ireland in Europe also being subjected to severe austerity. With contagion spreading, other countries will no doubt suffer similar fates.

The loss of highly skilled labour via brain drain and the under-funding of higher education is a threat to knowledge acquisition and therefore also wealth creation in these countries and indeed, whole continents. This loss of local knowledge and the local culture it embodies is not in the interests of humanity. Action needs to be taken immediately to conserve knowledge creation and re-production and to tame the rankings and the policy-makers that exploit them. What is needed is a return to the global village—a network of universities worldwide where each is valued and accessible, and where quality is clearly defined and can be monitored. At the close of the last century, Stiglitz (1999) described how knowledge is one of the global public goods and how it is a requirement for international cooperation in the 21st century. To conserve our collective educational heritage, we are going to have to embrace this idea and flatten the global higher education hierarchy. Isaac Newton, in a letter to his rival Robert Hooke in 1676, said, 'we stand on the shoulders of giants'. A good higher education, wherever we receive it, and in whichever language, is one that gives us the helping hand we need to stand on these shoulders and look out onto new knowledge horizons. Pedagogical gems exist the world over; our job is not to count them, but to find them, embrace them and conserve them.

The digital age, while giving birth to new liberating modes of scholarly communication, has also given birth to many sorcerers' apprentices—financial agencies producing credit ratings, proprietary citation providers producing journal impact factors and now media agencies producing global university rankings. Every time, the apprentices have wielded their power irresponsibly. In the case of credit ratings, the destruction of national economies has sparked rebellions spreading like wildfire across continents. Citation impact has led to tenure insecurity and citation-based malpractices, but also gave birth to a global free open-access movement. Global university rankings are increasing brain drain and leading to knowledge extinction in large areas of the planet, but here too there is strong resistance; national policy makers and university administrators are starting to refuse to play this rigged game. We know that the sorcerer is out there and has hidden the truth about the quality of higher education in the numbers. Our job is to unravel the mystery hidden in the 1s and 0s so that the sorcerer, freed from the apprentices' 'services', can regain control of the workshop. Until that time, we propose the following protocol be followed to tame wayward apprentices and their rankings:

- Reject all rankings based on irreproducible reputational survey data
- Reject all rankings that involve arbitrary weightings in the aggregation of their indicators
- Reject all rankings that focus *exclusively* on research performance *or* on teaching
- Reject all rankings that do not include cost of education indicators
- Reject all rankings that do not normalise for discipline and language
- Reject all higher education policies based on rankings alone
- Foster interactive, user-specified combinations of indicators that are open and transparent
- Foster the issuing of statements of purpose to accompany all rankings
- Foster the use of independent agents to collect, process and analyse data
- Foster research into new proxies and assess their applicability and co-linearity.

Acknowledgements. M.T. is supported by a FP7-PEOPLE-2011-IEF grant. P.P. was supported by grants JCI-2010-06790, ECO-2011-23634 and P1-1B2012-27.

LITERATURE CITED

- Aguillo IF, Ortega JL, Fernández M (2008) Webometric ranking of world universities: introduction, methodology, and future developments. *High Educ Eur* 33:233–244
- Aguillo IF, Bar-Ilan J, Levene M, Ortega JL (2010) Comparing university rankings. *Scientometrics* 85:243–256
- Altbach PG (2007) Empire of knowledge and development. In: Altbach PG, Balán J (eds) *World class worldwide: transforming research universities in Asia and Latin America*. Johns Hopkins University Press, Baltimore, MD, p 1–28
- Amsler SS, Bolsmann C (2012) University ranking as social exclusion. *Br J Sociol Educ* 33:283–301
- Baty P (2012) Rankings don't tell the whole story – Handle them with care. *University World News* 227. Available at www.universityworldnews.com/article.php?story=20120626171938451
- Billaut JC, Bouyssou D, Vincke P (2009) Should you believe in the Shanghai ranking? An MCDM view. *Scientometrics* 84:237–263
- Bloom D (2005) Raising the pressure: globalization and the need for higher education reform. In: Jones GA, McCahey PL, Skolnik ML (Eds) *Creating knowledge: strengthening nations: the changing role of higher education*. University of Toronto Press, Toronto, p 21–41
- Bollen J, Van de Sompel H, Smith JA, Luce R (2005) Toward alternative metrics of journal impact: a comparison of download and citation data. *Inf Process Manag* 41: 1419–1440
- Bollen J, Van de Sompel H, Hagberg A, Chute R (2009) A principal component analysis of 39 scientific impact measures. *PLoS ONE* 4:e6022

- Bookstein FL, Seidler H, Fieder M, Winckler G (2010) Too much noise in the Times Higher Education rankings. *Scientometrics* 85:295–299
- Bremner J, Haub C, Lee M, Mather M, Zuehlke E (2009) World population prospects: key findings from PRB's 2009 World Population Data Sheet. Population Reference Bureau, United Nations, Washington, DC
- Brody T, Harnad S, Carr L (2006) Earlier Web usage statistics as predictors of later citation impact. *J Am Soc Inf Sci Technol* 57:1060–1072
- Buela-Casal G, Perakakis P, Taylor M, Checa P (2006) Measuring internationality: reflections and perspectives on academic journals. *Scientometrics* 67:45–65
- Butler D (2007) Academics strike back at spurious rankings. *Nature* 447:514–515
- Cameron WB (1963) *Informal sociology, a casual introduction to sociological thinking*. Random House, New York, NY
- Center for Higher Education Development (CHE) (2006) Study and research in Germany. University rankings, published in association with *Die Zeit*. <http://www.daad.de/deutschland/studium/hochschulranking/04708.en.html>
- Enserink M (2007) Who ranks the university rankers? *Science* 317:1026–1028
- Florian RV (2007) Irreproducibility of the results of the Shanghai academic ranking of world universities. *Scientometrics* 72:25–32
- Frank R, Cook P (1995) *The winner-take-all society*. The Free Press, New York, NY
- Gewirtz P, Ball SJ, Bowe R (1995) *Markets, choice and equity in education*. Open University Press, Buckingham
- Gladwell M (2007) *Blink: the power of thinking without thinking*. Back Bay Books, New York, NY
- Gladwell M (2011) The order of things. *New Yorker* 87: 68–75
- Grewal R, Dearden JA, Lllien GL (2008) The university rankings game. *Am Stat* 62: 232–237
- Hazelkorn E (2013) World-class universities or world class systems? Rankings and higher education policy choices. In: Marope PTM, Wells, PJ, Hazelkorn E (eds) *Rankings and accountability in higher education: uses and misuses*, UNESCO, Paris, Forthcoming
- Held D, McGrew A, Goldblatt D, Perraton J (1999) *Global transformations: politics, economics and culture*. Stanford University Press, Stanford, CA
- Herman EYC, Chomsky NN (1988) *Manufacturing consent: the political economy of the mass media*. Pantheon, New York, NY
- Holmes R (2012) Power and responsibility: the growing influence of global rankings. *University World News* 228. Available at www.universityworldnews.com/article.php?story=20120630083447805
- International Rankings Expert Group (IREG) (2006) The Berlin Principles on ranking of higher education institutions, 20th of May, 2006. Available at www.ireg-observatory.org/index.php?option=com_content&task=view&id=41&Itemid=48
- Interuniversity Higher Education Policy Network (2013) The Greek higher education institutions in the world rankings: universities, technical institutions, private structures. Study 8, 2013. Available at <http://hepnet.upatras.gr/xfiles/pubs/HEP%20translation%20final.pdf>
- Ioannidis JPA, Patsopoulos NA, Kavvoura FK, Tatsioni A and others (2007) International ranking systems for universities and institutions: a critical appraisal. *BMC Med* 5:30
- Ishikawa M (2009) University rankings, global models, and emerging hegemony: critical analysis from Japan. *J Stud Int Educ* 13:159–173
- Jeong H, Néda Z, Barabási AL (2003) Measuring preferential attachment in evolving networks. *Europhys Lett* 61: 567
- Jöns H, Hoyler M (2013) Global geographies of higher education: the perspective of world university rankings. *Geoforum* 46:45–59
- Karabel J (2005) *The chosen: the hidden history of admission and exclusion at Harvard, Yale, and Princeton*. Houghton Mifflin, Boston, MA
- King C (2007) Of nations and top citations: Whose papers rate as elite? *ScienceWatch* 18(3)
- López-Illescas C, De Moya-Anegón F, Moed HF (2011) A ranking of universities should account for differences in their disciplinary specialization. *Scientometrics* 88: 563–574
- Marginson S (2006a) Dynamics of national and global competition in higher education. *High Educ* 52:1–39
- Marginson S (2006b) Global university rankings at the end of 2006: Is this the hierarchy we have to have? OECD/IMHE & Hochschulrektorenkonferenz Workshop on Institutional Diversity: rankings and typologies in higher education, Bonn, 4–5 Dec
- Marginson S (2006c) The Anglo-American university at its global high tide. *Minerva* 44:65–87
- Marginson S, Van der Wende M (2007) To rank or to be ranked: the impact of global rankings in higher education. *J Stud Int Educ* 11:306–329
- Marope M (2013) *Rankings and accountability in higher education: uses and misuses*. Education on the move. UNESCO, Paris
- Morley L (2003) *Quality and power in higher education*. Open University Press, Buckingham
- Myrdal G (1957) *Economic theory and underdeveloped regions*. Gerald Duckworth Press, London
- Newman ME (2001) Clustering and preferential attachment in growing networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 64:025102
- Organization for Economic Cooperation and Development (OECD) (2006) *Education at a glance*. OECD Publishing, Paris
- Ortega JL, Aguillo IF (2009) Mapping world-class universities on the Web. *Inf Process Manag* 45:272–279
- Rauhvargers A (2013) *Global university rankings and their impact - Report II. Report on Rankings 2013*. European University Association, Brussels
- Saisana M, d'Hombres B, Saltelli A (2011) Rickety numbers: volatility of university rankings and policy implications. *J Reg Policy* 40:165–177
- Schevitz T (2008) Stanford, UC Berkeley to set up Saudi college. *San Francisco Chronicle*, March 5. <http://www.sfgate.com/bayarea/article/Stanford-UC-Berkeley-to-set-up-Saudi-college-3292715.php>
- Seglen PO (1997) Why the impact factor of journals should not be used for evaluating research. *BMJ* 314:498–502
- Shuai X, Pepe A, Bollen J (2012) How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations. *PLoS ONE* 7: e47523
- Social Mobility and Child Poverty Commission (ed) (2013) *Higher education: the fair access challenge*. Social Mobility Report 2013, UK Government, London
- Stergiou KI, Lessenich S (2013) On impact factors and uni-

- versity rankings: from birth to boycott. *Ethics Sci Environ Polit* 13:101–111
- Stiglitz J (1999) Knowledge as a global public good. In: Kaul I, Grunberg I, Stern M, (eds) *Global public goods: international cooperation in the 21st century*. Oxford University Press, New York, NY, p. 308–325
- Sutherland WJ, Goulson D, Potts SG, Dicks LV (2011) Quantifying the impact and relevance of scientific research. *PLoS ONE* 6:e27537
- Taraborelli D (2008) Soft peer review: social software and distributed scientific evaluation. In: Hassanaly P, Ramrasingh A, Randall D, Salembier P, Tixier M (eds) *Proc 8th Int Conf Design of Cooperative Systems (COOP '08)*, Carry-Le-Rouet, May 20–23, Institut d'Etudes Politiques d'Aix-en-Provence, Aix-en-Provence
- Taylor M, Perakakis P, Trachana V (2008) The siege of science. *Ethics Sci Environ Polit* 8:17–40
- Thelwall M, Haustein S, Lariviere V, Sugimoto CR (2013) Do alt-metrics work? Twitter and ten other social web services. *PLoS ONE* 8:e64841
- Trachana V (2013) Austerity-led brain drain is killing Greek science. *Nature* 496:271
- UNESCO (2003) World conference on higher education partners, final report: foreword by J. Daniel. Paris 23–25 June 2003. <http://unesdoc.unesco.org/images/0013/001352/135213e.pdf>. June 2003.
- Van Leeuwen TN, Moed HF, Tijssen RJ, Visser MS, Van Raan AF (2001) Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics* 51:335–346
- Van Noorden R (2012) 366 days: 2012 in review. *Nature* 492:324–327
- Van Raan AF (2005) Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics* 62:133–143
- Van Raan AF, Van Leeuwen TN, Visser MS (2011) Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics* 88:495–498
- Waltman L, Calero Medina C, Kosten J, Noyons E and others (2012) The Leiden Ranking 2011/2012: data collection, indicators, and interpretation. *J Am Soc Inf Sci Technol* 63(12):2419–2432
- Wilson KA, McBride MF, Bode M, Possingham HP (2006) Prioritizing global conservation efforts. *Nature* 440:337–340
- Wintour P, Adams R (2013) Universities failing on equality of access, research suggests. *The Guardian*, Sunday 16 June. <http://www.theguardian.com/education/2013/jun/16/universities-failing-equality-state-school-admissions>
- Woods P, Jeffrey B (1998) Choosing positions: living the contradictions of Ofsted. *Br J Sociol Educ* 19:547–570
- Zitt M, Filliatreau G (2007) Big is (made) beautiful: some comments about the Shanghai ranking of world-class universities. In: Sadlak J, Liu NC (eds) *The world-class university and ranking: aiming beyond status*. UNESCO-CEPES, Institute of Higher Education, Bucharest, Shanghai Jiao Tong University, Shanghai, and Cluj University Press, Cluj-Napoca, p 141–160

Editorial responsibility: Konstantinos Stergiou, Thessaloniki, Greece

*Submitted: July 19, 2013; Accepted: November 23, 2013
Proofs received from author(s): February 20, 2014*