Comparability ≠ Ranking

**Chris Brink**
University of Newcastle upon Tyne

By way of a thought experiment, let us ask how we could compile a gastronomic league table of European cities. In other words, how would we go about ranking cities by the quality of their restaurants?

After some extensive committee work, consultation and workshops, we come up with the following idea. First we invent a rating system for restaurants, as follows:

> 4 stars = World-class
> 3 stars = Of European renown
> 2 stars = Of national renown
> 1 star = Acceptable. Popular in the city
> Zero stars = Poor.

We then invite each city to submit a number of restaurants for rating by a panel of gourmets. After that, the cities are evaluated as follows: of all the restaurants visited, give the percentage that have, respectively, been awarded four stars, three stars, etc.

Imagine, for example, that we are in the city of Someburgh, and that 10% of the restaurants visited were awarded four stars, 25% were awarded three stars, 40% were awarded two stars, 20% were awarded one star, and 5% received no star. That gives what we will call the *quality profile* of restaurants in Someburgh, pictured like this:
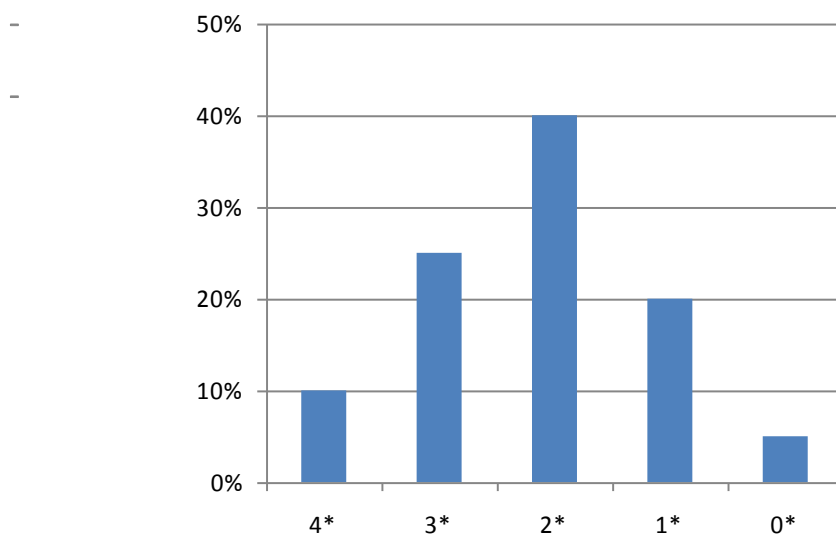
The picture gives us, quite literally, a profile, in the sense that we can see that in Someburgh the quality of restaurants have, like a simple version of a normal curve, a concentration towards the middle of the scale, with smaller representation at the two extremities.

Now imagine the same exercise being conducted in a number of other European cities, such as Otherville, Thistown and Elseport, and that we get a collection of quality profiles such as the following:
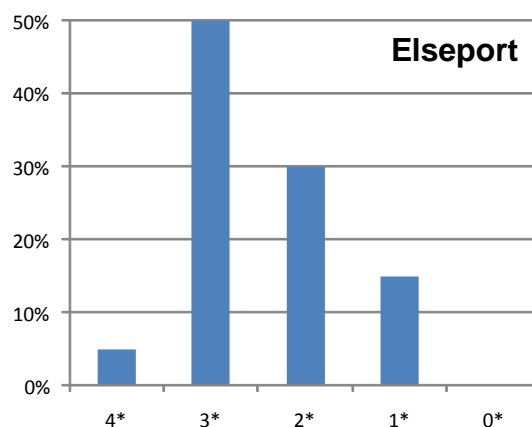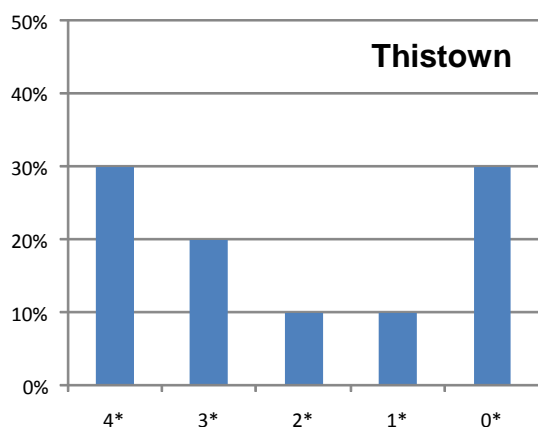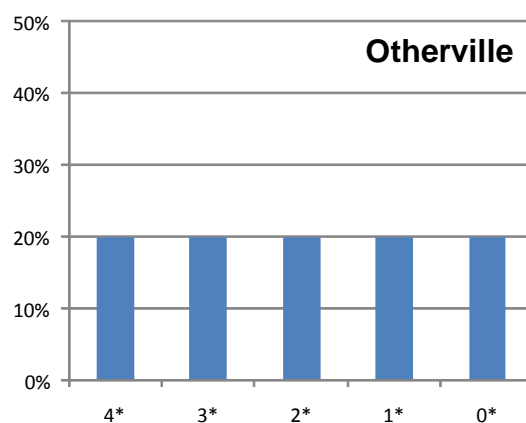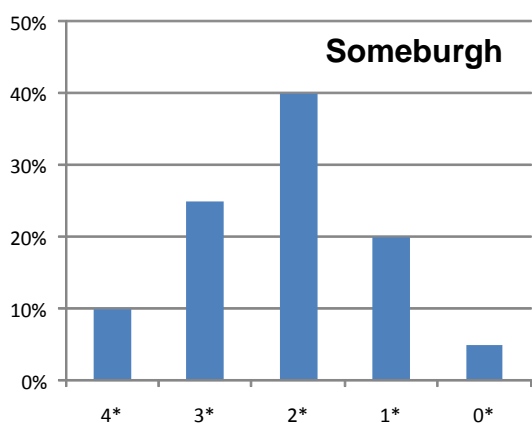
It is easy to see that in Otherville the quality profile is quite level, in Thistown there is a high proportion both of world-class restaurants and of poor ones, and in Elseport there is quite a low proportion of world-class restaurants but a very high proportion of nationally renowned ones. After a while one grows accustomed to this mode of representation, to the extent that the mere listing of the percentages will suffice to evoke the profile. Thus:

    Someburgh: 10-25-40-20-5
    Otherville: 20-20-20-20-20
    Thistown: 30-20-10-10-30
    Elseport: 5-50-30-15-0.

This is good progress, but in order to produce a league table we now have to turn these profiles into a ranking, which will tell us which city is first, second, third and so

on. Doing so is a bit of a conundrum, so we do some more consultation. The weekly magazine *Gastronomy Times* then comes up with the following method, which they call a "weighted grade point average". First, they argue that being world-class is surely more important and prestigious than being nationally renowned, or just popular in the city. Accordingly, they give a weighting to each of the categories (4-star, 3-star, etc.), as follows: the percentage of 4-star restaurants is multiplied by 4, the percentage of 3-star restaurants is multiplied by 3, and so on, giving a numerical score in each category. The sum of these scores is then divided by 100 to give a "GPA" (grade point average), which is a number between 0 and 4. The cities are ranked in order of their GPA. As follows:

$$\text{Someburgh:} \quad [(10 \times 4)+(25 \times 3)+(40 \times 2)+(20 \times 1)+(5 \times 0)] \div 100 = 2.15$$
$$\text{Otherville:} \quad [(20 \times 4)+(20 \times 3)+(20 \times 2)+(20 \times 1)+(20 \times 0)] \div 100 = 2.0$$
$$\text{Thistown:} \quad [(30 \times 4)+(20 \times 3)+(20 \times 2)+(20 \times 1)+(20 \times 0)] \div 100 = 2.1$$
$$\text{Elseport:} \quad [(5 \times 4)+(50 \times 3)+(30 \times 2)+(15 \times 1)+(0 \times 0)] \div 100 = 2.45.$$

In the gastronomic league table published by *Gastronomy Times*, therefore, Elseport is first, Someburgh is second, Thistown is third, and Otherville is fourth.

A rival publication, however, called *Restaurant Fortnight*, thinks that this is not quite right. First, they notice that, mathematically, the "average" part of the GPA (in other words division by 100) is just mumbo-jumbo – it makes no difference to the rankings. Second, and more substantively, they argue that the weightings used by *Gastronomy Times* do not give enough credence to the difference between the categories. Being "world-class", they argue, is not just being of a better quality than "nationally renowned", it is being of a *much* better quality. Accordingly, they decide to do a ranking along the same lines, but with the 4-star category getting a weighting of 16, 3-star with a weighting of 9, 2-star with a weighting of 4, 1-star with a weighting of 1, and 0-star weighted zero. These weights then give rise to the following scores:

$$\text{Someburgh:} \quad [(10 \times 16)+(25 \times 9)+(40 \times 4)+(20 \times 1)+(5 \times 0)] = 565$$
$$\text{Otherville:} \quad [(20 \times 16)+(20 \times 9)+(20 \times 4)+(20 \times 1)+(20 \times 0)] = 600$$
$$\text{Thistown:} \quad [(30 \times 16)+(20 \times 9)+(10 \times 4)+(10 \times 1)+(30 \times 0)] = 710$$
$$\text{Elseport:} \quad [(5 \times 16)+(50 \times 9)+(30 \times 4)+(15 \times 1)+(0 \times 0)] = 665.$$

Accordingly, *Restaurant Fortnight* publishes a league table in which Thistown is first, Elseport is second, Otherville is third and Someburgh is last.

Clearly we could continue playing with the weightings, and arrive at different rankings. At this stage, however, having read the *Gastronomy Times* and *Restaurant Fortnight*, the Association for Haute Cuisine (AHC) decides to publish its own league table. In fact, it does better: it publishes three separate league tables. First, it re-names the categories, as follows: A 4-star award is called a Gold Medal, a 3-star award is called a Silver Medal, and a 2-star award is called a Bronze Medal. The other awards are simply disregarded. The AHC then publishes a Gold Medal Table, which is of the utmost simplicity: it ranks each city by its percentage of gold medals. On the Gold Medal Table, therefore, Thistown is first, Otherville is second, Someburgh is third, and Elseport is last.

Next, the AHC publishes a Gold-and-Silver-Medal Table, in which it counts a gold medal as being worth two silver medals. Thus:

<div style="margin-left:2em">

Someburgh: (10x2)+25 = 45
Otherville: (20x2)+20 = 60
Thistown: (30x2)+20 = 80
Elseport: (5x2)+50 = 60.

</div>

On this table, therefore, Thistown is first, Elseport and Otherville are joint second, and Someburgh is third and last. For completeness, the AHC finally does an All Medals Table, in which a gold medal is worth three bronze medals, and a silver medal is worth two bronzes. Thus:

<div style="margin-left:2em">

Someburgh: (10x3)+(25x2)+40 = 120
Otherville: (20x3)+(20x2)+20 = 120
Thistown: (30x3)+(20x2)+10 = 140
Elseport: (5x3)+(50x2)+30 = 145.

</div>

Which means that in the All Medals Table Elseport is first, Thistown second, and Otherville and Someburgh joint third and last.

We may now pause for breath, and note that the five ranking methods have produced five different rankings. To keep track, we summarise the various rankings as follows, representing each city by the first letter of its name, and putting shared positions in parentheses:

<div style="margin-left:2em">

Gastronomy Times Table:	ESTO
Restaurant Fortnight Table:	TEOS
AHC Gold Medal Table:	TOSE
AHC Gold-and-Silver-Medal Table:	T(EO)S
AHC All Medals Table:	ET(OS).

</div>

In keeping with our initial intention, we may refer to all of these rankings as *quality* rankings. We may, perhaps, be slightly disappointed that there is no unambiguous winner (or loser), and in fact no clear pattern of ranking. At best, if we wish to draw any conclusions at all, we may note that either E or T is always first, that O, S and E have all been last, that T has never been last, and that S and O have never been first.

At this stage (to continue our thought experiment) a new player enters the arena. The Tourist Board points out that in each city it is relevant to ask, not just about the percentage breakdown of ranked restaurants, but also how many restaurants were actually submitted for ranking. When they look into the matter, they find that the numbers vary quite a bit: Someburgh submitted 40 restaurants, Otherville submitted 30, and Thistown and Elseport only submitted 20 each. The Tourist Board then argues that the right way to do a ranking is to factor in the size of the submissions, because that would give the public a good idea of what a city actually offers, which gives more information about the gastronomic quality of that particular city. Accordingly, it reworks the five "quality" rankings as *power rankings*, by the simple expedient of multiplying the number produced by each quality ranking for each city

by the number of restaurants submitted for ranking by that city. This yields the following results:

Gastronomy Times Power Ranking:

> Someburgh: 2.15x40 = 86
> Otherville: 2.0x30 = 60
> Thistown: 2.1x20 = 42
> Elseport: 2.45x20 = 49

SOET

Restaurant Fortnight Power Ranking:

> Someburgh: 565x40 = 22,600
> Otherville: 600x30 = 18,000
> Thistown: 710x20 = 14,200
> Elseport: 665x20 = 13,300

SOTE

AHC Gold Medal Power Ranking:

> Someburgh: 10x40 = 400
> Otherville: 20x30 = 600
> Thistown: 30x20 = 600
> Elseport: 5x20 = 100

(OT)SE

AHC Gold-and-Silver-Medal Power Ranking:

> Someburgh: 45x40 = 1,800
> Otherville: 60x30 = 1,800
> Thistown: 80x20 = 1600
> Elseport: 60x20 = 1,200

(OS)TE

AHC All Medals Power Ranking:

> Someburgh: 120x40 = 4,800
> Otherville: 120x30 = 3,600
> Thistown: 140x20 = 2,800
> Elseport: 145x20 = 2,900

SOET

With these rankings, the previous situation has almost been reversed. With the "quality" rankings, either E or T was always first; with the "power" rankings either E or T is always last. Either S or O is always first under the power rankings, whereas they were never first under the quality rankings.

At this stage, therefore, with two competing methodologies and ten different ranking methods, we already have nine different outcomes. We may, however, still hope that there is some pattern to these outcomes. We note, for example, that so far S and O seem to be often grouped together, as do E and T.

Suppose now (to take the thought experiment further), the matter has gained considerable prominence, and a number of further opinions are voiced. The

Chamber of Commerce, for example, points out that we have only taken account of the number of restaurants *submitted*, instead of the *total number* of restaurants in each city. The gastronomic quality of a city, they argue, is not just about those restaurants carefully selected for submission (probably by the Tourist Board). Gastronomic quality is about what the city offers in the totality of its restaurants. Therefore the power ranking, while correctly putting a size factor into the calculation, errs in using only the number of restaurants submitted. The size factor that should be used instead is the *ratio* of the submitted restaurants to the total number of restaurants in the city. Fortunately the Chamber of Commerce already has a database listing the total number of restaurants per city. From that database they know, for example, that Someburgh has 64 restaurants in total, Otherville has 33, Thistown has 40 and Elseport has 30. That gives the Chamber of Commerce an easy way of calculating what we might call the *Ratio Rankings*, the respective ratios being 40 over 64 for Someburgh, 30 over 33 for Otherville, 20 over 40 for Thistown and 20 over 30 for Elseport.

Gastronomy Times Ratio Ranking:

Someburgh: 2.15x(40/64) = 0.134
Otherville: 2.0x(30/33) = 1.818
Thistown: 2.1x(20/40) = 1.05
Elseport: 2.45x(20/30) = 1.633

OEST

Restaurant Fortnight Ratio Ranking:

Someburgh: 565x(40/64) = 353.125
Otherville: 600x(30/33) = 545.455
Thistown: 710x(20/40) = 355
Elseport: 665x(20/30) = 443.333

OETS

AHC Gold Medal Ratio Ranking:

Someburgh: 10x(40/64) = 6.25
Otherville: 20x(30/33) = 18.182
Thistown: 30x(20/40) = 15
Elseport: 5x(20/30) = 3.333

OTSE

AHC Gold-and-Silver-Medal Ratio Ranking:

Someburgh: 45x(40/64) = 28.125
Otherville: 60x(30/33) = 54.545
Thistown: 80x(20/40) = 40
Elseport: 60x(20/30) = 40

O(ET)S

AHC All Medals Ratio Ranking:

Someburgh: 120x(40/64) = 75
Otherville: 120x(30/33) = 109.091
Thistown: 140x(20/40) = 70
Elseport: 145x(20/30) = 96.667

OEST

With this method of ranking, any supposition that we might have some natural pairings (S with O, or E with T) has disappeared. Moreover, the ratio rankings features something absent from the other rankings: a single clear winner (Otherville).

With the extra data now accumulated we may conveniently enhance our quality profiles. We simply have to use actual numbers instead of percentages, and add another column giving the number of unranked restaurants. The updated quality profiles then look like this:
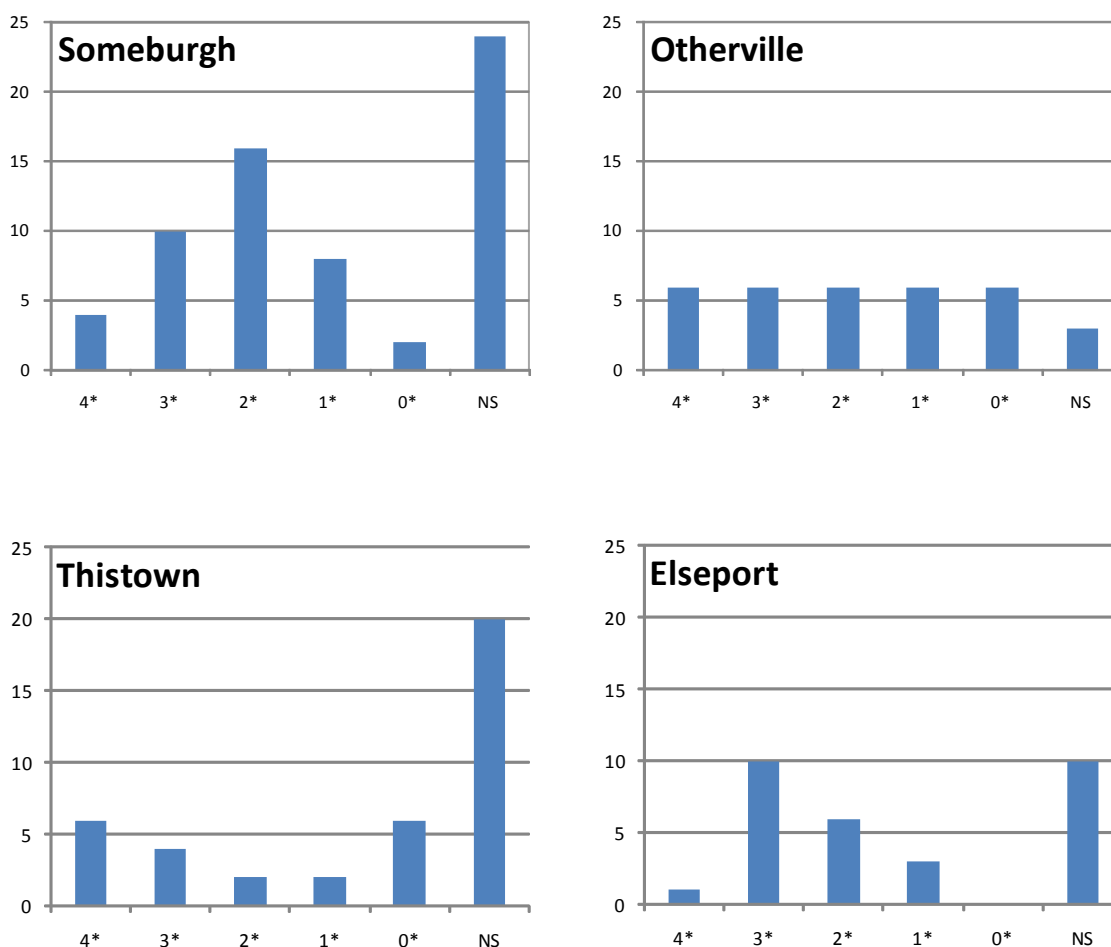
There is much to be said for these enhanced profiles. They still give us a visual impression of quality – indeed, exactly the same profile as before, since the numbers of ranked restaurants are in the same ratio as the original percentages. However, since they are presented in terms of numbers rather than percentages, they also tell us, more or less at a glance, the size of what each city has to offer by way of gastronomic satisfaction – i.e, the total number of restaurants. Moreover, by simply looking at the height of the last column (the unranked remainder), and comparing it with the combined size of the other columns, we have a very clear idea of the tactical approach adopted by each city in the rankings game. Thus we see that Otherville played a very straight game, submitting almost all their restaurants (30 out of 33). Thistown, on the other hand, only submitted half their restaurants (20 out of 40) – possibly because they knew the unsubmitted ones to be very poor. Elseport and

Someburgh are somewhere between these two extremes, having submitted two-thirds (20 out of 30) and five-eighths (40 out of 64), respectively.

We could easily carry on inventing yet further ranking methodologies. We may, for example, start the whole game over, by calculating, for each city and for each category of rating, the number of restaurants in that category as a percentage of the total number of restaurants in the city, not just as the percentage of restaurants submitted. That would lead to another 60 calculations, on top of the 60 we have done already. And, as pointed out earlier, we could obtain yet more variations simply by changing, in any calculation, the relative weightings. Instead, let us put aside our thought experiment and return to reality.

-----/////-----

Readers familiar with higher education in the United Kingdom will have noticed that the various methodologies adopted above for the gastronomic league table of European cities bear a rather close resemblance to newspaper rankings of the research performance of UK universities following the 2008 Research Assessment Exercise (RAE). For the RAE, each university was invited to submit information on its research performance in up to 67 research areas, called units of assessment. These submissions were then assessed, mainly on the best four academic papers of each member of academic staff selected for entry. Each unit of assessment was given a profile based on the percentage of work judged to fall within each of the following ratings:

4* = world-leading
3* = internationally excellent
2* = recognised internationally
1* = recognised nationally
0* = unclassified; below standard.

This exercise was carried out under the auspices of the Higher Education Funding Agency for England (HEFCE), for the purposes of allocating research funding. HEFCE stated the aim of the exercise as follows[1]:

*The primary purpose of RAE2008 is to produce **quality profiles** for each submission of research activity made by higher education institutions. The results will be used by the four UK funding bodies to determine the funding for research to the institutions which they fund from 2009-10.*

The quality profiles in question are exactly as in our thought experiment, explained by HEFCE as follows[2]:

*Sub-panels assessed the research submitted against the published **criteria**. They then made a recommendation to the **main panel** for endorsement. The judgement indicates the proportion of the research that met each of four quality levels or is unclassified. In each case, the panel took account of three*

---

[1] **http://www.rae.ac.uk/news/2008/results.asp**
[2] **http://www.rae.ac.uk/aboutus/quality.asp**

overarching components of the submission - research outputs, research environment and indicators of esteem. The results are published as a graded profile for each UOA for each submission. An example of quality profiles for two hypothetical submissions is below.

Unit of assessment A	Full-time equivalent research staff submitted for assessment	Percentage of research activity in the submission judged to meet the standard for:				
		4*	3*	2*	1*	Unclassified
University X	50	15	25	40	15	5
University Y	20	0	5	40	45	10

The point to note here is that HEFCE felt no need to move from the notion of a quality profile to the calculation of some quality ranking. The profile itself sufficed.

Others were not so restrained. The *Times Higher Education* (*THE*) turned the RAE profiles of the 159 participating higher education institutions into what it called a "Research Excellence Ranking". This was done by calculating a "weighted grade point average", exactly as in our thought experiment, with the same weightings and also without regard to size. Not surprisingly, small specialist institutions tended to do well on this ranking. According to the *THE* ranking, the most excellent research institution in the UK is the Institute of Cancer Research, with a submission of 97 researchers in 2 categories of assessment, beating into second place Cambridge University, which submitted 2,040 researchers in 50 categories. Perhaps recognising this as something of an anomaly, another newspaper, *The Times* of London, adapted *THE*'s ranking by simply leaving out small institutions. A different publication, *Research Fortnight*, adopted the methodology of power rankings, using essentially the algorithm featured in our thought experiment. This reversed some of the earlier rankings. The London School of Economics, for example, which was ranked 4th by *THE* (and 2nd by *The Times*), came 27th in the *Research Fortnight* ranking. The Russell Group of research-intensive universities, generally regarded as the UK's Ivy League, published its own version of power rankings, called the Gold Medal Table and Gold-and-Silver-Medal table, again as illustrated in our thought experiment (adopting the power ranking methodology). On these tables Russell Group universities tended to do well.

Our imaginary gastronomic league table, therefore, is no mere figment of the imagination, but a small illustration of what happens in reality. And that brings us to the core point of this paper. Using quality profiles, we are able to compare universities. Any linear ranking obtained from these profiles is the result of a subsequent calculation. Therefore: **It is perfectly possible to compare universities without ranking them**.

The general point is that difference does not imply ordering. It is a fallacy to conclude from the fact that A ≠ B that there is some substantive sense in which either A is bigger than B or B is bigger than A. In general, if two things are different, it does not

follow from that fact alone that any one of them is better, stronger, more beautiful or more meritorious than the other. The difference between an apple and an orange is not that one is inherently a better fruit than the other, or tastier or healthier. The comparison consists in describing what they are, not in ranking them. Of course it is perfectly possible for any fruit-lover to express a preference, on taste or health or convenience or any other grounds, and thus to impose a ranking, but your ranking will have no more credence than my ranking. Or suppose we ask a child to compare geometric shapes. A very good answer would be to name these shapes as circles, squares, triangles and parallelograms – an answer which gives a comparison, but not a ranking. Again it is perfectly possible to do a ranking, for example in terms of size, or the number of edges. But there is no inherently unique ranking arising from difference in shape.

The quality profiles above, real or imaginary, are a case in point. They are more like shapes, less like numbers, and as such allow comparison without implying an inherent ranking. If we add, as we did, parameters of size, the quality profiles give us an outline, a size-and-shape view, of the entities we are comparing. That is a perfectly good basis of comparison. To list these size-and-shape descriptions in some kind of ranking is perfectly possible, and various algorithms can be devised for this purpose. But the outcome would depend on the algorithm and the weightings assigned to parameters used in the algorithm. With the examples we have used, no ranking gives a better representation of reality than the shape-and-size quality profiles have already given.

It must be acknowledged that we usually do a comparison in order to come to a decision, and that a decision typically involves a ranking, even if that only consists of deciding on a single winner. Thus we may profile applicants for an appointment, or sport stars for an award, but only for the purpose of deciding who to appoint, or to crown. Likewise, HEFCE compiled quality profiles of research performance for a purpose, namely to allocate funding. Since funding comes in an amount, and amounts can be ranked, it may be argued that the HEFCE profiles do amount to a ranking. Which is true, but then only with the caveat that the ranking produced is that of a funding allocation, not a quality ranking per se. HEFCE's decision on funding could have taken any one of a number of directions, such as whether performance should be funded in all categories or only in some, whether there should be weighting assigned to categories and if so what, whether size of submission should be taken into account and if so how. Behind such decisions, as may be expected, lies the prior question of what strategy HEFCE wished to adopt. Should it encourage research concentration, and focus its allocation of resources on those universities which have a high level of performance in a number of different fields? Or should it reward excellence wherever it may be found, no matter whether it is in a small and relatively isolated unit? Evidently these are important questions, and it is the deliberate strategic choice on answering them that would determine the funding methodology, and hence the algorithm for turning quality profiles into funding allocations.

This illustrates a general point. **Ranking reflects the purpose of the ranker at least as much as any inherent properties of the entities being ranked.**

It may be said in response that newspaper league tables do have a purpose, which is to rank universities in order of quality. But that brings us back to the core point: there is no reason to assume that quality inherently presents itself in a linear order.

That, in turn, raises the bigger question of what we take quality to consist of. The curious fact is that university league tables, which purport to rank universities in terms of quality, have made almost no connection with that part of higher education where the meaning of quality has been extensively debated, namely the area of quality assurance. Starting in the late 1980s and early 1990s, when the topic of quality assurance first made its way into academic consciousness, the question of definition dominated. When is a university a good university? What do we mean by quality in higher education? To this question many answers have been advanced, including the answer that no answer is possible.[3]

- The first and most enduring view is that quality is what you have when you answer Yes to the question "is it good?". This is the notion of quality as excellence.

- If we wish to take account of sectoral diversity, and acknowledge that not all universities are the same, we need to relativise the notion of quality to the mission of the university. This is the notion of quality as fitness for purpose, where quality is the answer to the question "Is it good at what it does?".

- A further variation arises when we judge, not the extent to which a university achieves its purpose, but the purpose itself. This is the notion of quality as fitness *of* purpose. It sees institutional quality as the answer to the question "What is it good *for*?".

- Separate but related is the idea of a "good university" being one which brings societal benefit. It delivers a skilled workforce, supports culture and business and industry, helps to create jobs, plays a role in economic development and social mobility, and has a strong corporate social responsibility profile. Here the notion of quality as reflected by esteem has changed into a notion of quality as reflected by impact. The quality question then becomes "Is it beneficial?".

- On a different front there is the notion of quality as value. The obvious version takes the quality question about university education to be "Is it good value?", meaning "Is it a good investment for my future?". There is also a sharper version, namely "Is it value for money?". On this approach, the discourse and methods of the market become part of the quality debate.

- A variation on the theme of value is when we phrase the quality question as being about *added* value. The argument here is that quality of education is not just a matter of taking straight-A students from school and turning them into straight-A graduates. It is also a matter of taking disadvantaged and poorly prepared entrants and turning them into competent and well-rounded graduates.

---

[3] In what follows I draw on my paper "On quality and standards", delivered as a keynote address at the Australian Universities Quality Forum in Alice Springs, July 2009. See *Proceedings of AUQF 2009*, **http://www.auqa.edu.au/files/publications/auqf%20proceedings%202009.pdf** , pp. 12-26.

The quality of what the university provides is then a matter of the educational distance travelled between entry and exit.

Of these various notions of quality, it is predominantly the first which underlies – indeed, is taken for granted in – the construction of league tables. Witness, for example, the "Research Excellence" table of the *Times Higher Education*. And yet the apparent simplicity of the notion of quality as excellence is deceptive.

To illustrate the subtlety of the concept of excellence, here is a simple question. **Is excellence something which can in principle be attained by any university?** Can we envisage a situation where (in a given country, say) all the universities are excellent? If we cannot envisage such a situation, we either have to say that there are some universities for which excellence is unattainable, or we have to say that by definition "excellence" means "being at the top", so that at any given time there is some ranking of universities within which only those in the top layer are the excellent ones. If the former, we have opted for a view that says some universities are inherently second-rate. If the latter, we have opted for a view that may withhold the accolade of excellence from some universities no matter how good they are, just because there are some which are better.

League tables operate on this latter notion of excellence as "being at the top". What that means is that quality (for which "excellence" has been adopted as a proxy) is seen as a relational notion, not a substantive one. The question has morphed from "Is it good?" to "Is it better than the others?". With that, the concept of quality is reduced to nothing more than a position on a list. The shift from "Is it good?" to "Is it better than the others?" represents exactly the fallacy pointed out earlier, to argue from the fact that A ≠ B to the conclusion that one must be better than the other. On this view, we cannot rest content with asking of each of A and B whether they are good, acknowledging that they are different, but feel compelled to create some sense in which one is crowned as better than the other.

There are also other difficulties with the notion of quality as excellence. One is that reputation so easily becomes a proxy for excellence, particularly when there is some kind of polling involved in the construction of the league table. That would give the advantage to the old, the rich and the beautiful. Another difficulty is that the notion of elite universities so easily morphs into an attitude of elitism. We should keep in mind that there is a difference between "elite" and "elitist": the former says that if you are good you can become part of an elite, whereas the latter says that if you are not part of an elite you cannot be any good.

-----/////-----

Thus far, universities have pitched themselves into the ranking game with unbridled enthusiasm. It is quite common now for a university to advertise itself, on its website, in promotional material or recruitment, by reference to some league table. Pre-eminent amongst these would be the *Times Higher Education* "Top 200" table – surely an enviable position for a newspaper to be in. Thus the University of Auckland describes itself as "a top 50 University in a top 5 City", the University of Sydney is "rated as one of the top 40 universities across the globe", and Imperial College London is "ranked 5th best in the world by the THES", while the City University of

Hong Kong's growing international reputation is "evidenced by its surge up the THES rankings"[4]. A university might even characterise itself, not by its current position on the "Top 200" list, but by the position it aspires to. The National University of Singapore "will be a globally-oriented university, in the distinguished league of the world's leading universities"[5]. The Queen's University of Belfast sees itself as a "Destination top 50" university[6]. And the matter goes deeper, even into government policy. Through its "985 Programme", China aims to break into the top ranks of universities world-wide. In Saudi Arabia, when opening the King Abdullah University of Science and Technology (and endowing it with a $10 billion grant), the King stated that "It is my desire that this new University become one of the world's great institutions of research"[7]. In Australia, according to recent reports, the Government is proposing to exempt the country's elite Group of Eight universities from reporting to its new quality and standards agency, on the grounds that it would be an unnecessary imposition on institutions already ranked on the top 100 of the *Times Higher Education*'s world rankings[8].

Such behaviour is common despite the well-known shortcomings of newspaper league tables. These were analysed, for example, in a 2008 report "Counting what is measured, or measuring what counts?", commissioned by HEFCE[9]. Different league tables commonly measure different things, and even where they measure the same thing they would give it different weightings. Within the same league table, the rankings of an individual institution can vary hugely from year to year. The School of Oriental and African Studies in London, for example, in 2004-2007, was successively ranked 44th, 103rd 70th and 243rd in the world by the *Times Higher Education Supplement*. League table methodology is often questionable, particularly when the compilers of the league table make it clear that their ranking cannot be reconstructed from the data they have used. The data itself may be dodgy, for example in rankings based on a so-called "peer survey". Rankings may reflect reputational factors rather than the quality or performance of institutions.

One way of responding to such criticisms is to try and build better ranking systems. The *Times Higher Education* has recently spent a lot of time and effort on improving its methodology, and designing a more authoritative world ranking. At the same time there is a European movement to design a less Anglo-centric ranking system, presumably with the same intention of better capturing some underlying reality. All such efforts, however, are still constrained by the same factors: any ranking system must begin by choosing a methodology, deciding on what to measure, assigning

---

[4] Peter W.A. West, "A Faustian Bargain? Institutional Responses to National and International Rankings", *Higher Education Management and Policy*, 21/1, 2009, pp. 11-18. In the same volume, see also Ellen Hazelkorn, "Rankings and the Battle for World-Class Excellence: Institutional Strategies and Policy Choices", pp. 55-76. Note that the abbreviation "THES" refers to the *Times Higher Education Supplement*, the former name of what is now simply the *Times Higher Education* (*THE*).

[5] **http://www.nus.edu.sg/aboutus/vision.php**

[6] *Times Higher Education,* xxxx

[7] **http://www.kaust.edu.sa/about/kingsmessage.html**

[8] *Times Higher Education* 17-23 June 2010, pp. 16-17.

[9] A report done by the Centre for Higher Education Research and Information of the Open University: "Counting what is measured, or measuring what counts?", HEFCE Issues paper 14, 2008, **http://www.hefce.ac.uk/pubs/hefce/2008/08_14/** .

weights to parameters, and designing an algorithm to produce a list. To build a better ranking system does not address the inherent limitation of ranking.

The point is that ranking systems create a reality as much as they reflect a reality. Such is the promotional value and the public relations potential of being "top-ranked", that universities have become complicit in a system where comparability is seen as nothing more than ranking. It is questionable whether this game-playing is good for higher education. On the other hand, it would seem quixotic to tilt at league tables, as if they could be slain or driven away. Evidently league tables are here to stay.

To address this conundrum, let us return to the question above, with a slight tweak. **Is quality something which can in principle be attained by any university?** It is hard to see how we could, in conscience, answer no to this question. If we answer no, then the same line of argument unfolds as above. Either quality is unattainable for some universities, in which case we would have to question their right to existence, or we would say that at any given time only some universities are of quality, in which case education becomes little more than a lottery for students. It would seem, then, that in any country we should strive for a higher education system where all universities are universities of quality. In that case, however, having answered "yes" to this version of the question, but answered "no" to the previous version, we have to conclude that quality is not the same concept as excellence.

Perhaps we can concede that excellence has some underlying aspect of linearity, particularly since "excellent" means "to excel", which means to stand out above the rest. The caveat, however, would be to make it quite clear that a "table of excellence" always reflects its choice of variables as much as it reflects any reality. As is shown in our thought experiment of the gastronomic league table of European cities, the ranking will change with the methodology, the weightings and the algorithm. Intellectual honesty would seem to demand no less than a health warning with every ranking. With that would come a much more sober and realistic view of what "tables of excellence" actually reflect, and what they create.

The idea that quality is substantive and excellence is relational then gives us a way forward. We may conceive of quality as a notion that allows comparison, but does not demand ranking, by distinguishing it from excellence, which has an underlying aspect of linearity. The advantage of such a distinction is a that it would allow us to consider higher education as a diverse sector, where universities may differ without one being considered better or worse than the other. This would fit the idea of quality as fitness for purpose. If we can get away from the idea that the mission of each university should be to feature in a "top 200" list, universities could differentiate themselves by purpose, and be judged in terms of the extent to which they accomplish that purpose.

All universities have in common the core functions of teaching and research. However, the topics they cover, the manner in which they cover these topics, and the purpose for which they teach and do research, all allow differentiation. Such differentiation makes for a healthy sector, well integrated into civil society. The argument for educational diversity must surely be at least as strong as the argument for biodiversity. The obsession with rankings works against such diversity, since universities all try to play in the same game. Rankings have the effect of turning

diversity into hierarchy. For this reason, if no other, we need a better way of presenting quality.

The idea of quality profiles is a promising way of doing so. As demonstrated by the RAE, it is perfectly possible to compile reputable quality profiles for research. It cannot be beyond our wit to do something similar for teaching and learning – at least it should be possible to agree on the parameters we need to measure. If, then, we take these domains as common, and purpose as university-specific, it does not seem inconceivable that some reasonably stable way can be found in which a university presents its quality profile to the world – particularly if such a profile also contains verifiable data about past performance, not just promises about the future.

Quality profiles do not try to compete with league tables, in the same sense as the RAE profiles did not compete with the newspaper research rankings that followed. On the contrary, quality profiles constructed on roughly the same model, with verifiable data, will provide a rich seam to be mined by those who like league tables. However, quality profiles, if well constructed, could lessen the dependence on rankings. Quality profiles will allow comparison – to the state, to funders, to prospective students and their parents, to academics, and to business and industry. Since it allows comparison, it will allow decision-making, for whatever purpose, by whatever sector of the state or society. Since the profile precedes any ranking, and is richer in information than a mere list, it should lead to better decision-making. If we can, as a sector, put this idea across, and implement it well, we will be doing society and ourselves a favour.

----/////-----