

DIE FURCHE, Vienna, Austria (www.furche.at)

Interview with Prof. Yves Gingras

Questions (Proposal)

Today, there is the impression that universities are more and more seen as companies in the free market. Do you share that view and what is your opinion towards it?

(What else can be seen as the social, cultural, economical background of the current ranking hype?)

The idea of a global market for universities must be put in the context of the neoliberal view of the world (weltanschauung) that try to transform all institutions into a market. For universities, this transformation was facilitated by the general decline of government investment into higher education and the recent demographic declines of the usual cohorts of university students (17-21). Taken together, these facts have forced many institutions to try to attract more foreign students to compensate their difficult internal economic situation. All this of course has been sugar coated under the language of “globalization” and “internationalization”, but it is less cultural exchanges that are aimed at than bringing good money from foreign students to compensate the decline of revenues coming from governments. A very important but much less noted effect of this competition is the tendency of non-English speaking university to offer courses in English to attract foreign students, even when this is not their national language. This can transform the mission of national universities, which had been created to train their citizens into a disconnected organization simply seeking to attract foreign students even when they have difficulty speaking the local language. We can even predict a decline in the quality of training in institutions accepting students whose language skills are minimum even in English.

Why do we need university rankings at all?

Strictly speaking, we do not need university rankings. These have emerged recently in relation to the movement I described above that saw universities trying to attract more foreign students. This created a market for tools aimed at helping students to choose the supposedly “best” foreign institutions where they could go. And these new tools have, in turn, been used by universities in their marketing strategies to attract those students. In fact, however, less than 10% of foreign students really used these rankings to make their choice. The reason is that their choice is much more influenced by professors and friends who know that in a given specialty it is better to go to A than to B for many reasons that cannot be summarized by a single number in a ranking!

In your book *Bibliometrics and Research Evaluation. Uses and Abuses* (MIT press, 2016) you have analyzed the problems of the rankings. Since then, do you think is there enough critical debate going on – or still too much „blind belief“ in rankings?

(Has the “International Ranking Group” had any success since then?)

I think there is much more critical thinking now that say ten years ago when the first Shanghai ranking came out and excited many university Presidents, particularly in Europe, which was not well represented in this ranking. Though it seems that many managers still use those rankings in a kind of cynical manner for marketing and promotion, it is more and more difficult not to see that they are very simplistic and artificial and thus cannot really serve as a foundation for making policy decisions in matters of higher education.

I personally do not think that the “International Ranking Group” -- which is in fact composed of the very organizations that create the rankings -- can be considered to have a credible discourse on rankings. Its aims is in fact to promote them and try to circumvent critiques by superficial arguments and the rhetoric of “Excellence”, a buzz word that may signify

different things to different people according to the mission they set themselves. Their “auditing” system is self referential and circular thus creating a self fulfilling prophecy. Looking at some of the members of this Group should raise eyebrows... (I think for example of King Abdulaziz University which has been pointed at in *Science* for playing a dubious game to go up the rankings, see my MIT book p. 84).

Why are many indicators used in academic rankings invalid – what’s the problem with the current system?

This is a bit technical but I explain in chapter 4 of my book that there are very precise criteria that should be met to guarantee the validity of an indicator. Curiously most indicators of the diverse ranking systems are used on the basis of a kind of “common sense” as if that was sufficient to make them valid! In fact most indicators used in existing rankings are bogus and do not measure what they are suppose to measure namely the “quality” of the university.

Which rankings are problematic for example?

The Shanghai ranking contains invalid indicators like Nobel prizes. Who can seriously think that a prize that recognize research done about 20 years ago give us any indication of the quality of the university today! It is thus ridiculous to publish a ranking each year as if universities could change rapidly in the short term. They are in fact very inertial organizations that only change over a long period of time – and by the way this is a good property for it excludes the danger of adapting fast to passing fads and short term demands. So the annual rankings and the changes they make visible are only statistical fluctuations in the data and do not measure real changes. The only reason to publish them each year is to keep the ranking itself into the news. Another problematic ranking is the THES ranking for it uses an opaque system of qualitative assessments by unknown people chosen all over the world on the basis of unknown criteria. This raises a crucial ethical question that is too often forgotten: no evaluation should be done using data that cannot be checked for quality and validity.

Which rankings, on the other side, provide valid approaches? (example)

At this point in time, the best indicators for research impact are those provided by the so-called Leyden Ranking produced by the CWTS for it gives well defined and valid indicators that are not arbitrarily combined to give a single number. U-Multirank also provides an interesting and more diverse set of measures. Here again they avoid the pitfall of adding together heterogeneous measure to create a simplistic single number and ranking.

How do you see evaluation based on the number of publications in specific journals (how useful are the journal impact factors)?

The number of publications is obviously a valid indicator of productivity. Interpreted in the context of the researcher and its specialty and discipline it gives a useful information. Likewise the total number of citations to a paper also provides valid information when taking into account the specificity of the field: citation rates vary much between mathematics and molecular biology or sociology for example. The Journal Impact factor provides a nice case of abuse of bibliometrics for it characterizes a journal not a paper! So it is irrational to use it to measure the paper. After many years of recalling that simple fact, many scientific organizations have finally understood that we should not use impact factors to evaluate researchers. Again impact factors are specific to disciplines and give information on journals not papers or researchers!

In which way do university rankings nowadays affect university life? What is the impact of rankings with regard to various aspects such as international students in Europe, setting of research agenda, individual researchers, etc.? What’s

the dimension of this impact?

Rankings certainly affect institutions, at least those who make the mistake of taking them seriously. It even brings some of them to try to manipulate the data to climb in the ladders of those ranking! This has been proven (and I give examples in my book). As I said, I don't believe most students' choices are really influenced by rankings. It is more the managers of the institutions that play the game of comparing themselves and try to use that as a marketing tool. But it can also push institutions to suggest to their researchers to publish in journals and not others or on trendy topics just to try to get up the ranking, even though that may not be the best way to get research results to the right audiences, particularly in social science and human sciences.

What about diversity in university rankings? Is there a lack of diversity in measuring output, performance, and innovation?

There is clearly a lack of diversity and all rankings go for the easily obtained measures: counting papers or citations is easy. Measuring the quality of teaching is very difficult! And measuring added value is even more difficult. Who can be surprised that an institution that selects only the best students gets the best results without making much effort, contrary to institutions that take more difficult students and make them learn more. Also a university is multidimensional so it does not make sense to summarize it into a single number. It's like collapsing a three-dimensional space into a single point! From this point of view the U-Multirank initiative insists on this multidimensionality.

Why did the Council of Canadian Academies emphasize the fact that indicators cannot replace personal judgement? How do you see the currently strong belief in the automatization process driven by computers and algorithms? Do we give up our capacity to critically evaluate? Will the machines take over?

The conclusion of the Council of Canadian Academies should be considered as good "common sense". In any well-conceived evaluation process, numbers can only provide data to help making decisions but they can never dictate the decision. The reason is in fact simple: for any indicator, one can always make two opposite decisions depending on the context. Suppose a Department performs very well. One can decide to do nothing and let it continue as it is or to give them more resources. On the contrary if it performs badly one can decide to shut it down or to invest new resources to make it better. The decision depends on the larger context and the goal to attain. This should be obvious but, curiously, it is not and the CCA report (of which I was a member) had to remind managers some basic facts about evaluations. Thus any decision based on algorithms would be not only irrational but would have perverse effects, generating behavior that goes against the objective of making things better. Such automatic algorithms also raise major ethical questions.

Today there is a tendency to globalize the evaluation model of the natural sciences, even for the humanities. There is critique that this doesn't really fit for the humanities, which is mostly dependent on the quality of ideas – and probably the most difficult field to evaluate. What would you suggest in order to evaluate research in the humanities properly? (Perhaps also: Why do we need the humanities?)

It is a major mistake to apply the evaluation model of the natural sciences to the social sciences and the humanities. The reason is that the latter disciplines are more interpretive than based on "discoveries" that can be made only once. Moreover social sciences and humanities' most important contributions are often incorporated in books not in small

articles published every month as in biomedical sciences. As I show in the book, the citation practices are also very different: whereas scientists cite only research papers, humanities cite essentially books, even when in their published papers! The most obvious perverse effects of applying the natural sciences is seen in economics. Over the last forty years that discipline tend to cite only papers and denigrate books, as if they were thus more “scientific”. Using the Impact factor as an indicator cannot be applied to books and book chapters. We even observe pressure on scholars to publish papers instead of books so that they can be more easily evaluated with the usual indicators. The situation is clearly absurd when it is the indicator that drives the research practices and not the reverse.

What are the hidden pitfalls in evaluation and why do we need an „ethics of evaluation“, as you have indicated? What about the source of data?

Any decision made on the basis of “back boxes” of data that are proprietary and opaque would not be ethical since one cannot check for accuracy and validity. Discussions of evaluation often focus too much on the indicators but one should not forget that even for a valid indicator, one must also validate the source of data. And these sources should be transparent and accessible to the people or organizations being evaluated so that they can confirm the validity of the data and the reproducibility of the results. It is a basic notion of justice and equity.

Do we have to rethink the evaluation system completely?

I think the “fever of evaluation”, as I called it some years ago, it will pass. There has been so many criticisms that good sense will come back if scientists themselves react and objects to the most simplistic measures like the h-index and the Journal Impact factor. This movement as already begun as the 2013 San Francisco Declaration on Research Assessment (DORA) shows. Many editorials in *Science* and *Nature* have also denounced the abuse of simplistic indicators.

What is your vision for the future?

The future being essentially unpredictable I will pass on this one, particularly in the light of the surprising recent election of Donald Trump in USA...But we can certainly say that the abuses of evaluations will stop if the critical thinking on each and all suggested indicators remains active.