Honesty and transparency to taxpayers is the long-term fundament for stable university funding

Author(s):
Stark, Wendelin J.

Publication Date:
2016-12-12

Permanent Link:
https://doi.org/10.3929/ethz-a-010745069

Rights / License:
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International
Metrics in Research

Wendelin J. Stark
ETH Zürich

Honesty and transparency to taxpayers is the long-term fundament for stable university funding

The Swiss Society has been trusting the ETH domain for over 150 years: More than 2 billion CHF are annually given to research, engineering and technology transfer. This enormous sum of tax-derived money is spent by about 1000 faculty members leading teams at ETH, EPFL, PSI, EMPA and EAWAG, providing a globally unique teams at ETH, EPFL, PSI, EMPA and EAWAG, providing a globally unique funding situation.

On the more formal side, the Swiss Society has formulated the ETH Law (Art. 2.1.) specifying our job as teaching, research and technology transfer. Art 2.2 is very clear: “we must consider the needs of the country”. Given the amount of money spent, the public is astonishingly patient in letting us do long-term investments of significant costs. This deep trust is based on ETH Zurich’s past achievements and its standing as a “top university” – an unclear term, shaped through the ETH Zurich domain’s view in the public.

We should therefore ask ourselves what determines this image? Most people will not understand the details of a faculty’s work. They will use substitute parameters to evaluate the output of the ETH Zurich: Is this “average” or “outstanding” work? Are these people doing something useful with our investment?

Prominent researchers publicly condemn quantitative measurements of success and in particular university rankings, bibliographic parameters and outreach activity (e.g. altmetric or media coverage). Is this justified?

When evaluating the performance of a faculty, universities often rely on external evaluations. They are commonly performed by a committee delivering a report. A critically thinking layman, a politician or a tax-paying citizen may think about what is more honest or transparent:

1. The application of well-defined, countable, openly-available parameters derived from publicly-available source data (e.g. scientific papers and their citations), or
2. A collection of personal opinions gathered by a small group of invited experts from other universities?

The second procedure appears particularly questionable when the long-term relationship between faculty members at different universities is considered: The same groups of faculty share expert views in different committees, assist review panels, share or assign awards and grants, or organize keynote conferences. There are complex interwoven interests.

However, bibliometrics and other countable data also have severe flaws: Disturbing single-case events (e.g. highly-cited but retracted papers) can be identified in all areas of science. The opponents of objective measurement are quick in citing such flaws. What could we do instead?

Research ethics obliges us to use the best methods available to judge on a problem, i.e. the most objective methods. In addition, we should critically assess our choice of methods. Naturally, comparison depends on the focus of a research team. In all cases, however, a comparison is made with respect to others, considered equal or better. In our case, ETH Zurich should compare itself with world leading universities. At the level of individual researchers, we should also compare ourselves with leading scientists or engineers at leading institutions.

A fair measurement compares with a fair metric

Fundamental science. If a thought of an intelligent person is not published, it is lost. If a thought triggers generations of researchers to develop new thoughts, it has some impact. If two researchers at two good universities work on similar topics, scientific essays (“papers”) become comparable, at least qualitatively.

Applied Sciences and Engineering. Solutions, processes and materials can similarly be irrelevant, if never used, and of no interest to others in the field. If solutions/thoughts are used, and create new research fields or products for companies, they have impact.

Societal implications. If the interactions of a person with society (through any media/means) are inexistent, there is no impact. If these interactions lead to new thoughts, changes in behavior, improvements etc., this scientist has some effect or impact.

Opponents to measurement may cite prominent cases where faculty members of different fields were compared in an unfair way: An artist cannot be compared to a chemist using papers in leading journals as a metric. Such cases, however, are no argument against an adequate effort to use fair parameters. The three following proposals may be used to illustrate such a procedure.

Evaluating a traditional scientific research group. Researcher A at ETH
Zurich works on the metabolites of maritime sponges, using chemical and pharmaceutical methods. She publishes in the leading chemical journals, and occasionally in a multidisciplinary journal. Her natural peers at MIT, Stanford, U. Cambridge, Harvard etc. publish in similar journals. What does this mean? If two groups make 10 or 15 papers in the leading journals, annually, they are at a comparable level. Another team contributes only 2–3 similar papers per year – that is clearly less productive. Second, what happens with that work? One organic chemistry group is cited 1000 times per year – clearly a leading position if compared to other similar groups. Another team is only cited 100 times – it is clearly noted less. Third, research financing is most relevant and clearly measurable (e.g. ERC grant vs. SNF grant).

Evaluating an engineering group. Researcher B works on new chemical processes, publishes papers and patents, and works with the chemical industry. His peers are at a number of leading universities (TU Delft, MIT, Stanford University, University of Minnesota, Harvard University, etc.). The output is at least in three areas, which can be counted separately: Papers in leading journals, patents and industry projects (amount of money; patents; licenses from the university’s technology). Again, we can compare such an output: One group runs projects with large companies (e.g. 500 KCHF contract). That is clearly different from another researcher getting 20 KCHF for a sample analysis. A patent that is licensed and the basis for a 250 Mio CHF cash flow in a Swiss company is more valuable than a non-licensed patent that a university tries to commercialize for 5 years before dropping it. The output of an engineer in terms of papers can be compared at least partially with that of fundamental scientist A (above), since now both compete in the same category, using similar tools (papers, i.e. essays that might be cited and used by others). Ideally, one should look at the corresponding standing of these two persons in their field, e.g. top 1% of the field (excellent) vs. last third (not so good) instead of absolute numbers of citations only.

Evaluating a traditional mechanical engineer. Here, publications are less important and projects are measured against their use in industry (size of a process, cash flow, number of sites running that project etc.), commercial value (patent income, patent citations, licenses), or their impact in the fields (key conferences with formal abstracts, project presentations, contests won, etc.). The commercial fate of a process or product is important (e.g. over a spin-off company). Industry projects, measurable at least in their size, can be compared from one team to the other. The number and future jobs of educated students can also be compared.

Teaching. The success of a university faculty as a teacher, beyond course grade, is ultimately linked to the fate of her/his students: Do they find adequate jobs? Are they working in their field of education? Was their education of any use to their job? What do they earn? How long does it take to find a job? Following up on students a few years later is routinely done in American institutions, and may be significantly intensified at Swiss institutions.

Balancing individual contributions? Most faculties will contribute in several output metrics. How can such different parameters be combined, at least in a given discipline? The ETH tenure committee faces similar challenges, and ETH has developed a number of items where it evaluates an aspiring full faculty. These tenure criteria are a good starting point for faculty evaluation during department internal discussions, and during university-wide evaluations. However, they currently lack the “third part” of our job description (technology transfer), and are too heavily biased towards natural sciences.

The alternative is not to measure. A less objective judgement of performance is prone to personal bias. Bibliometrics and other measurements represent the action of thousands of actors and judges. As such, the outcome is at least less biased compared to the opinion of a small group of experts that partially know one another. If we do not measure, the best personally-connected faculty will win at evaluations done by small groups of evaluators.

Opponents to measurement may add: “In a liberal, free society, everybody is free to think what she/he wants.” Yes, but, most faculty members are employees of a university. Employment comes with a job description (see ETH law), hence, most researchers are not entirely free to do whatever they want. In the case of ETH Zurich, there is a well-defined purpose involved.

The above discussion becomes even more complex if one considers the following example: Researcher C uses 1 Mio CHF per year to hunt for a rare physicochemical effect, his papers are barely read and he does not take part in public outreach. At the same time, dozens of researchers at low-income countries try to improve treatments against diarrhea, using a combined budget of 1 Mio CHF. Here, the question of fair and morally acceptable use of finances becomes an unpleasant topic.

It appears difficult to explain to a layman why scientists should be rigorous with their object of study, but deny the same rigorous approach to their own performance. If we would live in a world of endless resources, this unpleasant discussion would not be needed. Being accountable and transparent is the basis for trust and (hopefully) a continued and generous financing of university research through the Public.

Recommended Reading

Prof. Dr. Wendelin J. Stark
ETH Zürich
Institute of Chemical and Bioengineering
Vladimir-Prelog-Weg 1-5/10, HCI E 107
CH 8093 Zurich
Switzerland
wstark@ethz.ch
http://www.fml.ethz.ch

Citation: Stark W: Honesty and transparency to taxpayers is the long-term fundament for stable university funding. Infozine 2016, Special Issue 1, 13–14, DOI: 10.3729/ethz-a-001054069
Copyright: W.J. Stark, CCBY NC ND 4.0
Published: December 12, 2016