Performance-Based Funding and the Research Mission

When the Indicator Becomes the Institution

Marek Novák

Marek Novák, Institute of Sociological Studies, Charles University, Prague, Czech Republic

The research mission of a European university is now funded through indicators, and the indicators have become the description that the institution gives of itself. This article examines the mechanism: the four families of indicator in use, the sample that replaces the count in the systems that have reformed, the department that reorganises around a deadline, the young researcher whose contract follows the cycle, the teaching that pays for the research, the monograph in a field where the article is the metric, and the evidence that a system can publish to know whether the mission survived. It argues that the indicator is not a measurement of the mission but a constitution of it, and that a system which funds what it can count has decided, without stating it, what its universities are for.

Contents **Page**

1. The Indicator as a Constitution **33**
2. Four Families of Indicator **33**
3. The Sample and Its Cost **33**
4. The Department and the Deadline **34**
5. The Young Researcher **34**
6. The Teaching That Pays **35**
7. The Monograph and the Metric **35**
8. The Data That a System Should Publish **36**
9. The Appeal **36**
10. The Small Discipline **37**
11. The Ranking That Follows **38**
12. The Contract Research **38**
13. The Infrastructure of the Assessment **38**
14. The International Comparison **39**
15. The Student in the Assessment **39**
16. The Public Money and the Public Good **40**
17. The Department That Reads Itself **40**
18. The Sequence of a Reform **40**
19. The Mission That Is Not Measured **41**
20. Conclusions **41**

1. The Indicator as a Constitution

A research assessment does not describe an institution and it constitutes one, because the institution organises itself around the description it expects to be judged by.

The mechanism is not cynical and it is structural. A department that knows the criteria of the next evaluation allocates its time, its appointments and its internal rewards to satisfy them, and the allocation is a decision about what the department is. A criterion that counts articles produces a department of articles; a criterion that reads a sample of the work produces a department that prepares work worth reading.

The consequence is that the choice of indicator is the most consequential decision that a ministry takes about its universities, and it is taken in the technical annex of a funding regulation. The instrument is to state the choice as a decision about the mission, debate it as such, and review the mission when the indicator changes.

2. Four Families of Indicator

The indicators in use across Europe fall into four families, and each produces a distinct institution.

The first is the count: publications, citations, doctoral degrees, grants. It is cheap, it is comparable and it rewards volume. The second is the sample: a number of outputs submitted by the institution and read by peers of the field, with a narrative of the contribution. It is expensive, it is defensible and it rewards quality as the field defines it.

The third is the mission: the indicators that an institution proposes for itself, negotiated with the funder and assessed against its own statement. It permits diversity and it rewards the drafting of a document. The fourth is the use: the patents, the contracts, the policy reports, the public engagement, which measures the effect beyond the academy and which is the hardest to attribute.

3. The Sample and Its Cost

The systems that replaced the count with a sample have produced a better assessment and a heavier burden, and the burden falls on the panel rather than on the institution.

A panel that reads five outputs from each of two hundred researchers reads a thousand texts, which is a year of work for a group of twenty. The cost is real and it is smaller than the cost of the distortions that a count produces, which is the argument that the reforming systems have made and that the others have not accepted.

The consequence is a division of Europe into systems that read and systems that count, and a mobility of researchers from the second to the first. The instrument is a regional panel: a group of systems that share a pool of readers and recognise each other's assessments, which reduces the cost per system and increases the comparability of the result.

4. The Department and the Deadline

The cycle of a research assessment is between four and six years, and the cycle organises the life of a department in a way that is visible to anyone who works in one.

In the two years before the submission, the department hires for the outputs it lacks, postpones the projects that will not produce in time and asks its staff to submit the work that is nearly ready. In the two years after, it repairs what the assessment cost: the teaching that was carried by the staff who remained, the students who were supervised by the people who were writing.

The consequence is a department that runs in a cycle and a research culture that is periodic. The instrument is a submission that is continuous: the outputs are recorded when they appear, and the panel reads the record of five years at the moment of the assessment. A continuous record removes the deadline from the calendar and leaves the assessment as an assessment.

5. The Young Researcher

The person whose career is most determined by the indicator is the researcher on a fixed-term contract, and the contract is frequently shorter than the cycle.

A postdoctoral researcher is appointed for two or three years on the strength of the outputs that the department needs for the next submission, and her contract ends before the assessment in which her work will be counted. The consequence is a workforce that produces the

outputs and does not share the benefit, and a career structure in which the assessment of a department is built on the labour of people who are not in it.

The instrument is a requirement that the outputs submitted by an institution include the work of every researcher employed in the period, with the contract stated, and a rule that a postdoctoral post is funded for the duration of the project rather than for the cycle. The two provisions make the contribution visible and the employment compatible with the work.

6. The Teaching That Pays

The research of a European university is funded, in a share that no system publishes, by the teaching income of the same institution.

A professor who buys out a part of her teaching with a grant leaves the course to a colleague or to a doctoral candidate, and the cost of the course is met from the teaching income while the grant pays for the research. The transaction is legitimate and it is invisible in the accounts, and the consequence is a teaching capacity that is systematically smaller than the student numbers require.

The instrument is a cost accounting of the buy-out: the hours released, the cost of the replacement, and the source from which the cost is met, reported in the annual accounts of the institution. A report of that kind permits a ministry to see whether its research funding is buying research or buying a subsidy from its teaching funding.

7. The Monograph and the Metric

The disciplines in which the book is the form of the contribution are the disciplines that the count serves worst.

A monograph takes five years, it is read by a few hundred people and it is cited rarely, and in several fields it remains the work through which a career is judged. A count of articles or of citations places the author of a monograph below a colleague who has divided the same research into six articles, and the placement is not a difference of quality.

The consequence is a decline of the form in the systems that count, and the decline is a loss to the disciplines rather than to the individu-

als. The instrument is a submission in which the form is declared and assessed by the peers of the field, with the reading of a chapter and the evidence of its use, which is what the sample systems have done and what the count cannot do.

8. The Data That a System Should Publish

A system that funds research through indicators can publish four figures about its own instrument, and the four are almost never published.

The first is the distribution of the funding by institution, before and after the assessment, which shows whether the concentration is a policy or an effect. The second is the share of the submitted outputs that were produced by researchers on fixed-term contracts. The third is the cost of the assessment itself, in the time of the panel and of the institutions. The fourth is the number of the appeals and the share of them that succeeded, which measures the reliability of the judgement.

The four are administrative facts and their publication is the instrument through which the assessment can be criticised with evidence rather than with principle.

Indicator	Cost	Institution produced
Count	Low	A department of volume
Sample	High	A department of arguments
Mission	Medium	An institution with a stated purpose
Use	High	An institution turned outward

Table 1 Four families of indicator and the institution each produces

9. The Appeal

An assessment that changes a budget requires a route of appeal, and the route is the point at which the reliability of the judgement is tested.

A panel of peers makes mistakes: a text read by a reader who does not know the field, a declaration misunderstood, a category misapplied. An appeal that is heard by the same panel is a formality, and an appeal that is heard by a body of the funder without the field is a lottery.

The instrument is an appeal to a panel of the same field with a different membership, a deadline of sixty days, and a written decision that states the ground. The three are cheap and they are the condition under which the assessment can be described as a judgement rather than as an administrative act.

10. The Small Discipline

A discipline with twelve researchers in a country is assessed by the same instrument as one with two thousand, and the instrument does not fit.

The sample of five outputs per researcher is a sample of the field itself when the field is small, and a reader from abroad may be the only person competent to judge it. The consequence is a discipline that is either over-rewarded by the excellence of its few or under-rewarded by the absence of a community to cite it.

The instrument is a provision for the small disciplines: a panel with a member from abroad, a statement of the size of the field, and an assessment that is reported separately rather than averaged into a total. A provision of that kind costs a paragraph and it prevents a decision that the field cannot contest.

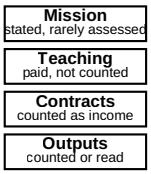

Figure 1 What a research assessment reaches and what it misses

11. The Ranking That Follows

The commercial rankings use a subset of the indicators that the systems use, and their effect on an institution is larger than their methodological quality.

A ranking that weighs citations and international staff produces a strategy in the institutions that are ranked, because the position in the ranking is read by applicants, partners and ministries. The consequence is a second indicator system, less accountable than the first, which sets the priorities of the institution.

The instrument is a published statement by the institution of the indicators it uses to judge itself, with the reasons, and a refusal to manage the position in a ranking that it did not choose. A statement of that kind is a defence that requires a rector with a mandate and a council that supports her.

12. The Contract Research

The part of the research income that comes from contracts with industry and with public bodies is counted in most systems and it changes the relation between the university and the funder.

A contract produces a report for a client rather than a contribution to a field, and the institution that counts the income rewards the report. The consequence is a research capacity that is oriented to the needs of the payer, which is a legitimate mission and a different one from the production of public knowledge.

The instrument is a separate reporting of the two: the income from the competitive public funding and the income from contracts, each with its own indicator, so that a system can see what it is buying. A separate reporting does not judge the contract and it prevents the contract from being counted as the other.

13. The Infrastructure of the Assessment

An assessment that reads a sample requires a database of the outputs of the system, and the database is the instrument that makes the reading possible.

A record of the publications, the grants, the doctoral theses and the contracts, with the identifiers, the authors and the affiliations, is the foundation of any assessment and it is expensive to maintain. The consequence is that the systems with a good record can reform and those without one continue to count what they can.

The instrument is a national record, funded and maintained, with the obligation of the institutions to deposit, and an open access for the researchers to their own record. A record of that kind costs less than one assessment cycle and it serves the institution for a decade.

14. The International Comparison

The comparison of research systems across Europe is made on the share of the publications in the top decile of citations, and the figure is used to rank systems that are not comparable.

The differences in size, in language and in the structure of the fields produce differences in the figure that have nothing to do with quality. A system that publishes in a language of limited diffusion has a lower share and not a weaker research base, and the two are conflated in every comparison that is published.

The instrument is a comparison with the field normalised and the language reported, and a narrative from the panel of the field. A comparison of that kind is slower and it is the only one on which a minister can act without damaging the system she is trying to improve.

15. The Student in the Assessment

The students appear in the research assessment as the doctoral candidates who produce the degrees, and they appear in no other way.

A department that is assessed on its outputs has an interest in the doctoral candidates who will produce them, and a student who needs five years for a question that will not yield three articles is a cost. The consequence is a selection of the candidates by the metric and a supervision that is oriented to the output.

The instrument is a submission that states the completion and the destination of the doctoral candidates as part of the evidence of the research environment, which is where the effect on the students belongs.

A statement of that kind makes the formation of researchers a part of the assessment and not a by-product of it.

16. The Public Money and the Public Good

The research that a system funds is public knowledge, and the arrangements of publication determine whether the public can read what it paid for.

A system that funds research and permits the results to be locked behind a subscription has produced knowledge that its citizens cannot consult. The consequence is a public expenditure whose product is a private asset, and the anomaly is now recognised in the policies of open access and rarely in the funding formula.

The instrument is a requirement that the outputs submitted to an assessment are deposited in an open repository at the moment of acceptance, with the metadata and the version, and a statement of the licence. A requirement of that kind is a condition of the funding and it is the cheapest reform in this article.

17. The Department That Reads Itself

The instrument with the largest effect on the research culture of a department is not the assessment and it is the internal seminar.

A department in which the work in progress is presented and criticised by colleagues produces a standard that its members internalise, and the standard travels to the assessment without being aimed at it. A department without the seminar produces work for the metric and no common standard.

The instrument is a seminar in the workload, with a convenor who receives the hours and a rule that every member presents once a year. The cost is a dozen hours per person and the effect is the only durable form of research development that a system can fund.

18. The Sequence of a Reform

A system that decides to reform its research assessment can do so in a sequence, and the sequence begins with the record.

The first step is the national record of the outputs, funded and mandatory. The second is a panel of the field that reads a sample, applied to one discipline as a pilot. The third is the replacement of the count in that discipline, with the funding held constant in the first cycle. The fourth is the appeal and the publication of the four figures.

A sequence of that kind takes eight years and it can be started by a ministry that has a record. A reform that begins with a new indicator produces a new distortion and a committee to explain it.

19. The Mission That Is Not Measured

A university has a mission that no indicator captures: the education of the people who will run the institutions of a region, the preservation of a language, the provision of a clinic in a district.

The four are the work that a system funds through its teaching component and that its research component ignores. The consequence is a hierarchy of institutions in which the ones that perform the mission are described as the ones that do not research.

The instrument is a mission statement that is negotiated and assessed, with the indicators that the institution proposes for the work that the national indicators do not reach. A statement of that kind is the third family of indicator and it is the one that permits a system to have more than one kind of university.

20. Conclusions

A research assessment is not a measurement of the mission and it is a constitution of it, and a system that funds what it can count has decided, without stating it, what its universities are for.

Three conclusions follow. The judgement must read the work: a sample submitted by the institution, read by peers of the field, with the narrative declared and the form respected, and a panel that includes a member from abroad for the small disciplines. The contribution must be visible: the outputs of every researcher employed in the period with the contract stated, the doctoral completion and destination as evidence of the environment, and the deposit of every output in an open repository at the moment of acceptance. And the instrument must be accountable: the distribution of the funding before and after, the share

produced on fixed-term contracts, the cost of the assessment, and the appeals with their outcomes, published every cycle.

A system that has done the three has funded a mission it can describe. A system that has counted what was easy has funded a mission it did not choose, and the choice was made in an annex.

References

[1] Novák, M. (2022). Research assessment and the Czech reform. Prague: Charles University Press.

[2] Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. Nature, 520.

[3] Wilsdon, J. (2016). The metric tide: independent review of the role of metrics in research assessment. London: HEFCE.

[4] European University Association (2022). Research assessment in Europe: mapping national practices. Brussels.

[5] OECD (2023). Assessing research performance: policy lessons. Paris.

Biography:

Marek Novák is a researcher at Charles University working on research policy, the assessment of universities and the sociology of academic work.