



WCRI 2019 Hong Kong

The Evaluative Breach

How research staff deal with a challenge of evaluative norms in a Dutch biomedical research institute

Research conducted by: Rinze Benedictus, Guus Dix & Jochem Zijderwijk
(contact: j.b.zijderwijk@cwts.leidenuniv.nl)

Center for Science and Technology Studies (CWTS), Leiden University, The Netherlands

- Project partially funded by ZonMW -



Universiteit
Leiden

Recent discussions on RI link integrity to evaluation

‘How researchers are evaluated reflects what we value most in the research enterprise and powerfully influences researchers’ behavior, including research integrity. [...]...[E]vidence implies that modifying current incentives and rewards is an important next step to optimize societal value and strengthen research integrity’ (Hong Kong Manifesto)

BUT HOW?

High-level statements and recommendations can mobilize people, but are not so useful for guiding realities of implementation

A 'natural' breaching experiment

Case: Dutch University Medical Center

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue

Fewer numbers, better science

Scientific quality is hard to define, and numbers are easy to look at. But bibliometrics are warping science — encouraging quantity over quality.

Leaders at two research institutions describe how they do things differently.

REDEFINE EXCELLENCE

Fix incentives to fix science

*Rinze Benedictus and
Frank Miedema*

An obsession with metrics pervades science. Our institution, the University Medical Center Utrecht in the Netherlands, is not exempt. On our website, we proudly declare that we

publish about 2,500 peer-reviewed scientific publications per year, with higher than average citation rates.

A few years ago, an evaluation committee spent hours discussing which of several faculty members to promote, only to settle on the two who had already been awarded particularly prestigious grants. Meanwhile, faculty members who spent time crafting policy advice had a hard time explaining how this added to their scientific output, even when it affected clinical decisions across the country.

Publications that directly influenced patient care were weighted no higher in evaluations than any other paper, and ▶

A 'natural' breaching experiment

Case: Dutch University Medical Center



Fewer numbers, better science

Scientific quality is hard to define, and numbers are easy to look at. But bibliometrics are warping science – encouraging quantity over quality. Leaders at two research institutions describe how they do things differently.

REDEFINE EXCELLENCE Fix incentives to fix science

Rinze Benedictus and Frank Miedema

An obsession with metrics pervades science. Our institution, the University Medical Center Utrecht in the Netherlands, is not exempt. On our website, we proudly declare that we

publish about 2,500 peer-reviewed scientific publications per year, with higher than average citation rates.

A few years ago, an evaluation committee spent hours discussing which of several faculty members to promote, only to settle on the two who had already been awarded particularly prestigious grants. Meanwhile, faculty members who spent time crafting policy advice had a hard time explaining how this added to their scientific output, even when it affected clinical decisions across the country.

Publications that directly influenced patient care were weighted no higher in evaluations than any other paper, and ▶



Harold Garfinkel

HOW DO ACTORS WORK TO (RE-)ESTABLISH ORDER?

Research design:

A detailed institutional analysis

- 33 Interviews
 - Dean, support staff, early-career and senior researchers
- Two focus groups
 - Early career researchers
 - Senior researchers
- Document analysis
 - Press statements, internal communication, minutes of internal meetings, evaluation guidelines
- Fieldwork
 - Research group meetings, support staff meetings

We find three types of responses to the evaluative breach

- (1) Accounts of potential**
- (2) Re-affirmative accounts**
- (3) Accounts of uncertainty**

Results: (1) *Accounts of potential*

Researchers that provide **accounts of potential** highlight what could or has become possible in a new order:

- More room for discussion about what is valuable
- More room for societal relevance
- Increasing possibility that different activities and forms of work will be recognized & rewarded

'I do think that **it has fired up the discussion** whether we need to score everyone on the basis of: how many papers do you have? What is your H-factor?'

'**...there is more and more attention for feeding our results back to patients...**'

'..."what have you done for society?" [...] Because you now ask people about this, they also realize: "oh, but **apparently that is appreciated too!**"'

Results: (2) *Re-affirmative accounts*

Researchers try to re-establish order in response to the breach through **accounts that re-affirm established norms**:

- Indicator legitimacy:
Specific indicators (JIF, H-index), based on output
- Evaluative standards of objectivity, measurement, quantifiability, transparency
- Compatibility with external evaluation systems
- The importance of old norms for specific positions: valuing fundamental research, goals for PhDs

'...But you have to be able to **quantify it**. Because I've got the idea that because it's still a little vague and **not measurable**...[...] ...there is **a lot of room for interpretation**. And that actually makes it **less transparent**. Look, you can say what you want about an H-index, but it is **a hard number**'.

'If you want to be eligible for certain grants (...), **your impact factor & H-index certainly still count**. It's definitely not off the table'.

'...that's all well and good, but I simply do have to have something to tell people to aim for in terms of their thesis. [...]. **You need these things to give people guidelines**'.

Results: (3) *Accounts of uncertainty*

Researchers that encounter problems re-establishing a sense of order provide **accounts of uncertainty**:

- **Organisational**: Uncertainty about criteria, workload, opaque and subjective evaluation processes
- **Field**: Uncertainty about evaluation outside UMC
- **Epistemic**: Uncertainty about valuing basic research
- **Status**: Uncertainty about the value of positions with new criteria (professors)

'...it is very unclear what you're being judged on. There are people here that are doing fantastic work and they aren't even nominated for associate professorship'.

'.....if I have to hand in a funding proposal in the Netherlands or Europe tomorrow **and I'm confronted with the old evaluation norms, than this evaluation norm has only been a blip on the radar**'.

'Lets be honest. **It is incredibly important to do fundamental research**, because **without it we can never make any progress**'.

'....it devaluates professorship'.

Conclusions:

From a call to change to a call to experiment

Accounts reveal the crucial role of researchers in maintaining old and establishing new evaluative orders

There is an urgent need to:

- move beyond high-level guidelines
- follow reform initiatives up close and learn about best practices; avoid overambitious calls for change and interventions that neglect to address context
- provide richer understandings of evaluation, organizational change & human action in RI discussion