

Science evaluation – As it is, as it could be

Social Science Information

1–5

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissionsDOI: [10.1177/0539018211021205](https://doi.org/10.1177/0539018211021205)journals.sagepub.com/home/ssi**Andrea Mubi Brighenti** 

University of Trento, Italy

Abstract

In this short commentary, I discuss Olof Hallonsten's argument against science evaluation. My invitation is to place the discussion in a wider frame that takes into account the many challenges scientists and scientific institutions face today. I argue in favor of a 'start from the middle' approach that gives emphasis to the inner principle of valorization of the scientific endeavor.

Keywords

measures, pragmatism, science evaluation, valorization, valuation studies

Résumé

Dans ce court commentaire, je reviens sur la prise de position d'Olof Hallonsten contre l'évaluation de la science. J'y invite à placer la discussion dans un cadre plus large, qui prenne en compte les nombreux défis qui se posent aujourd'hui aux scientifiques et à leurs institutions. Je suis pour ma part favorable à une approche consistant à « commencer par le milieu » et qui met en avant le principe de valorisation intrinsèque à la recherche scientifique.

Mots-clés

études d'évaluation, évaluation de la recherche scientifique, mesures, pragmatisme, valorisation

In his piece, Olof Hallonsten (2021) advances the argument that we should stop the folly of science evaluation before it's too late: evaluation, he points out, inflates bureaucracy in unnecessary and counterproductive ways, wasting and misdirecting precious resources. While I am broadly sympathetic with such view, I also believe that it holds for a

Corresponding author:

Andrea Mubi Brighenti, Università di Trento – Sociology and Social Research, Via Verdi, 26 Trento 38122, Italy.

Email: andrea.brighenti@unitn.it

historically specific version of evaluation. My aim here is thus to specify and nuance Hallonsten's central claim. Historically, it is not surprising that evaluation in the guise of performance assessment has reached higher education and research institutions, as this same trend has been seen in many other sectors since the 1980s, in conjunction with the neoliberal social-Darwinist philosophy – 'resources are scarce, you have to struggle to survive'. The current lexicon of research excellence is, in this sense, clearly related to the 'survival of the fittest' motif.

But one should not overlook that, if scientists are now trapped inside evaluation, the trap is in good part of their own making. Suffice to recall a figure such as Eugene Garfield, the renown linguist and businessman, inventor of the Science Citation Index (SCI), and a manager at Thomson Reuters corporation. Scholars both crafted the tools and set the tune of what is now a burgeoning industry, especially because the quest for visibility has contradistinguished scientists well before neoliberalism. One could not fully explain the current evaluation epidemics without taking into account the seduction element it contains: scholars are collaborating with evaluation because pleasure can be derived from performing, being recognized, and standing out. That is why it is not enough to 'leave science to scientists', as called for by Hallonsten: evaluation capitalizes on ambition. The situation in the natural sciences, for instance, is one of increasing gigantism of research projects, with large infrastructure investments: in this context, leading scientists turn into de facto managers, overseeing labs organized as mid-size or even large-size firms.

Once evaluation becomes compulsory, however, it also turns into 'just another drill' of academic life: in its institutionalized form, it has grown into a specialized field, with its expertise and methods. Side by side with the neoliberal seductive aspect of evaluation, whereby the latter is sold to scientists in pursuit of visibility-as-recognition and its accompanying rewards (in terms of career, salary, budget, bonuses, incentives etc.), another facet has been unveiled and denounced, which has been called 'disciplinary'. More precisely, using Michel Foucault's categories, science evaluation should be called not so much a form of surveillance-discipline (oriented towards the possibility of an unverifiable inspection) as much as a 'confessional' practice (oriented towards a certain, ethical self-examination) (Foucault, 1975, 2018). Once the 'assessment exercise' is on, each one – not only individuals but also institutions, such as educational programs, departments, research centers, etc. – is required to produce a detailed documentation of one's achievements *as well as* of one's shortcomings, a *redder rationem* meant to reduce the gap between stated objectives and obtained results, so as to *fine-tune* one's performance, in a self-improvement process that is not without religious resonances.

Such 'self-regulation' could not be underestimated, as it is perfectly complementary to institutional transformation. Institutions think in ways that can make an *a priori* neutral tool become amply dysfunctional. This way, the rush to evaluate has resulted in what, elsewhere, I have proposed to call a 'precession of measures', whereby measures replace the object they initially sought to measure (Brighenti, 2018). One may also speak, technically, of *perversion*. One striking illustration from Italy is, for instance, the faulty national pandemic-preparedness plan (not updated for 15 years, and never tested), which in 2020 Italy self-rated 'excellent' just few weeks before the country was knocked down by Covid-19. While the scientists who filed that document are surely to blame, it is also

clear that their main practical problem at the time of writing the report was more to pass the evaluation benchmark than work on any actual preparedness. This is what happens when concern for evaluation becomes more important than concern for the real phenomena evaluation is supposed to assess.

Another case in point is peer review. Peer review is undoubtedly an important mechanism of control, but as we all know by first-hand experience, it is far from perfect. In particular, peer review often ends up cutting off both the worst *and* the best submissions, whether they are articles or grant applications. Not only does it reward conformism over real innovation, but it promotes a conformist scientific production cloaked in an empty rhetoric of originality (the sort of ‘path-breaking-cutting-edge’ perfunctory declarations we are all acquainted with). This way, ‘originality’ becomes a ritual object and an illocutionary speech act. That there is some idiocy to this procedure is attested by the fact that peer review is increasingly conducted with the aid of artificial-intelligence software. The full circle will materialize before our eyes when a bot will write an academic paper which will be assessed by another bot, then cited by further bots, boasting a citation bonanza for an nonexistent academic profile. Some of this stuff is already happening.

The situation would sound comic, were it not for its human costs. As currently performed, science evaluation is responsible for a load of human suffering. Stress and poor mental health are reported to be a mass-scale issue among researchers as well as students in higher education. Younger, female and minority scholars and students often bear the brunt of these dynamics. Particularly pernicious effects follow when evaluation is embedded within a network of other institutional regulations – including, for instance, student debt, career promotion, etc. The ways in which evaluation is conducted and scientific production measured impact deeply on psychological well-being, and many promising scholars have been left disenfranchised and traumatized by the straightjacket of ruthless evaluation, usually camouflaged as the pursuit of ‘excellence’.

While there are good reasons to leave this model behind, one needs to remain alert to the wider political context in which we live. At a time when ideological and economic pressures weight heavily upon academia, it is all the more important that we restate out loudly our professional, scientific and human values: science as *Beruf* (work in the sense of profession and vocation). Politically-propelled conspiracy theories, authoritarian state ideologies, populist smearing campaigns and fundamentalist obscurantism are just some of the powerful forces that science confronts, and needs to keep restraining. From this perspective, evaluation, widely understood as rational scrutiny, is one of the very few tools we have at our disposal not to give in to the arrogance of power. Scrutiny is necessary to avoid both that dogmas, forgeries and malevolent falsifications pass as science and that, vice versa, science is discredited on baseless grounds.

But we also should not forget that the current evaluation system has itself put upon the shoulders of individual scientists such a pressure that this has pushed some of them towards resorting to unethical behavior (a classic case is the production of ‘science’ based on nonexistent or falsified datasets). In times of big uncertainties, then, my proposal is that we return to pose some naïve questions, which are also foundational ones. One such question sounds as: ‘Can we work for a while under the assumption that we do not know the true value of a piece of scientific work?’ And I think we can. This is what, in a sense, the

philosopher Imre Lakatos was recommending, when he pointed out that scientific programs (regardless of whether they are large-scale or individual) need time to bear fruit (Lakatos, 1978; Motterlini, 1999). If science needs evaluation, then it needs non-arrogant forms of evaluation that respect science's fundamental dynamic: the latter consists in advancing towards the new, into the unknown.

For his part, Friedrich Nietzsche saw well that, in all compartments of life, humans are evaluators through and through. It is on the terrain of the modes, styles and techniques of evaluation that there are the broadest differences as there are between freedom and slavery. A radically re-appropriative transformation of current evaluation might then come from an old rally call: 'Start from the middle!' This means caring for what really matters in our work, its *inner principle of valorization*, with respect to which other considerations are subsequent and ancillary. Here again the temporal horizon is of the essence. If one wants, I am describing a pragmatically-oriented stance, which invites letting scientific research proceed in ways that are relatively unencumbered and shielded from interferences and pressures: 'Do not block the way of inquiry', as Charles Sanders Peirce (1955: 54) effectively put it. Anyone with minimal experience in gardening, for instance, knows how much care goes into the production of a single aubergine. Culture and science are, from this point of view, a continuation of agriculture by other means. Such an attitude does not necessarily lead to complacency, cynical or fanatical manipulation. Nor even lack of production. To return to the towering giant evoked at the outset, the original Charles Darwin had an overall more benign message than all subsequent social-Darwinists in the row. To Darwin, nature appeared as interconnections, multiplicity, and experimentation: through diversity and variation, life carries out constants experiments with itself. I do not see any real obstacle to conceive of science in a similar way.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Andrea Mubi Brightenti  <https://orcid.org/0000-0002-7498-9857>

References

- Brightenti AM (2018) The social life of measures: Conceptualizing measure-value environments. *Theory, Culture & Society* 35(1): 23–44.
- Foucault M (1975) *Surveiller et punir. Naissance de la prison*. Paris: Gallimard.
- Foucault M (2018) *Les Aveux de la chair. Histoire de la sexualité IV*. Paris: Gallimard.
- Hallonen O (2021) Stop evaluating science: A historical-sociological argument. *Social Science Information* 60(1): 7–26.
- Lakatos I (1978) *The Methodology of Scientific Research Programmes*. Cambridge: Cambridge University Press.
- Motterlini M (ed.) (1999) *For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence*. Chicago: University of Chicago Press.
- Peirce CS (1955) *Philosophical Writings*. New York: Dover Publications.

Author biography

Andrea Mubi Brighenti is a social theorist based at the Department of Sociology of the University of Trento, Italy. His research topics broadly cover space, power and urban culture. His most recent monograph, co-authored with Mattias Kärrholm, is *Animated Lands. Studies in Territoriology* (University of Nebraska Press, 2020).