

Epistemic Consequences of Bibliometric Evaluation: A Reply to Rip and Stöckelová
Tommaso Castellani, Institute for Research on Population and Social Policies;
Emanuele Pontecorvo, Sapienza University of Rome; Adriana Valente, Institute for
Research on Population and Social Policies, National Research Council of Italy

In writing a contribution on the consequences of bibliometric evaluation on the practices of doing science, we were aware of dealing with a very delicate matter, which may be easily subject to objections and misunderstandings.

We are grateful to Arie Rip and Tereza Stöckelová for having carefully read and commented our paper, helping us to identify in which directions our arguments can be reinforced and, not least, giving us the opportunity to further reflect on our work. Starting from their comments, we are going to further clarify and expand our reasoning. Namely, we are going to address three main issues:

1. The methodology of the work, in relation to its objectives;
2. The logical sequence of our reasoning;
3. The conclusions of the study.

We will develop these points in the following three sections.

Preparing a Debate: Raising Questions from Within the Community

A paper based on qualitative interviews is particularly subject of methodological objections (Berg and Lune 2004). Both Rip (2014) and Stöckelová (2015) raised issues on the methodology we followed, and we are going to address them in more details in this section.

A first clarification is needed on the aim of our paper. The article is not about identifying trends—for which we would have chosen a completely different and basically quantitative methodology—but it aims to investigate the impact of the bibliometrics-based evaluation system at the epistemic level, which to our knowledge is a topic almost not dealt in the literature. By means of our qualitative interviews, we aim to raise the relevant problems from within the scientific community, opening a research line that could eventually set the ground for a follow-up based on a quantitative methodology. Opening such an investigation we set a bunch of questions that asked for a first feedback within the scientific community.

Clarifying the aim of our paper also helps us to reject the objection, raised by Stöckelová, that the diversity of responses depending on panellists' gender and discipline is not enough discussed in our article. As we just stated, our paper is not about detecting differences of positions among gender and disciplines. For such a task, qualitative interviews with a so narrow group are not at all the right instrument. Even if we had noted specific differences among the points of view of the six women and the six men of our panel, of course we would not have dare to present them as 'gender differences'. This aspect may of course be interesting for a future quantitative follow-up.

Regarding the delicate question of how the points of view expressed in the paper reflect the points of view of the interviewees, it is obviously a central issue of social science research. The power and limits of the interviews as a social science research method have been broadly discussed in literature, recognizing the risks contained in applying this technique but detecting at the same time that, according to estimates, about the 90% of all social science investigations is interviews-based (Briggs 1986).

From our side, we dealt with this problem in the most direct way: we sent the paper to all the interviewees to have a feedback from them. None of them raised objections on the argument we developed basing on their observations; on the contrary, we received several appreciations to the way in which we represented their positions.

Knowing the Italian research system, we are also able to reject the suspicion, raised more or less directly in Rip's and Stöckelová's comments, that interviewees' declaration may be instrumental, for they are interested in avoiding evaluation and protect their careers. Actually one should carefully analyse the peculiarity of the Italian system to take care of this argument. Once achieved a permanent position in Italy, careers are still strongly protected and, differently to other countries, in no way scientists with permanent positions can feel threatened of losing their job because of the results of bibliometrics-based evaluation. The interviewees of our panel are all permanently hired in the research system. As authors, we consider important to remark that only one of us has a permanent position, while two of us belong to the class on young(er) researchers with non-permanent positions, who are generally supporting the introduction of more rigid bibliometrics-based evaluation system for individual careers, on the way of Stöckelová's report about her scholar experience.

Are the Rules of the Game Part of the Definition of the Game Itself? Our Flowchart

We are grateful to Arie Rip who made the exercise to draw up the flowchart of our argumentation as we have surely missed to make it clear: indeed we have to say that his outline does not match at all the logic sequence we followed.

In his commentary, Rip wonders if the epistemic effects we address in our paper are consequence of the proliferation of scientific articles which is in turn caused by many aspects where the bibliometrics-based evaluation plays a role but is not the only neither the main factor: on the other hand he addresses in a general notion of "competition" the main one. We have to be very clear on this point: most of the epistemic effects considered in our paper are not in a causal relationship with the proliferation of scientific articles which cannot thus be considered the central box of an hypothetical logic flowchart.

In the introduction of our paper we reviewed some literature which, among the causes of this proliferation, mentions also the bibliometric indicators; but we actually examined this subject only in the last paragraph of our discussion. In that paragraph we pointed out that the growing in the number of publications actually happened, many decades before the introduction of the bibliometrics-based evaluation. The growing of the scientific communities may be a simpler cause of the proliferation of articles; in any case the

growth rate of the scientific production is a highly debated issue and subject of an intense research since many decades (Merton 1973; Ziman 1980). In our paper we clearly explain how bibliometric indicators have been created as a *response* to this proliferation of articles: we discuss the contributions of Bernal, Bradford and others, who—more than half a century ago—were wondering how to cope with the rapid growth of scientific publications (Bernal 1960; Bradford 1948). Only subsequently—much more recently—bibliometric indicators have been applied to the evaluation of scientists and research centres, but according to our interviewees bibliometrics still plays a fundamental role in the information retrieval from both aspects of selection and accessibility. The simple impact of bibliometrics on the proliferation of the scientific literature might have been a merely quantitative ‘non epistemic’ effect if we just look at the scientific journals as static knowledge reservoirs: nevertheless it is in the dynamic aspects, as the attitudes in retrieving and selecting the information that becomes ‘science’, that an epistemic issue may arise.

In any case all the scientists we interviewed entered the world of scientific research when the ‘explosion’ of the scientific publications had already started since long time. This event is part of their background and we are not questioning how the world would be without it. But *all* of the interviewees declared that they detect a change of attitudes after the introduction of the bibliometrics-based evaluation. We aimed to investigate whether this change of attitudes in their own way of doing science could unveil some common traits and suggest possible epistemic implications.

We are aware that, among the consequences of bibliometrics-based evaluation mentioned by the interviewees, there are some of them which may move in the direction of further aggravating the problem of excessive proliferation of papers, as the tendency of ‘publishing everything’ or the salami publishing.

But at the same time we registered many other consequences of the bibliometrics-based evaluation which have very little to do with the proliferation of scientific articles. Mentioning only some of them, bibliometrics-based evaluation:

- Impacts on the choice of the research topic;
- Discourages long and theoretical essays, towards short and empirical papers;
- Pushes to chose a clear school of thought, to increase the citations from a well-defined epistemic community;
- Discourages the interdisciplinary approach and the change of research topic during the career;
- Hinders the repetition of scientific experiments.

As Rip states, among the causes of these features there is the competition in science, eventually declined in a projects-oriented funding ruled by agencies or within the more general frame of the new public management. But the only competition is not sufficient to explain these elements: the explanation rather involves the normative aspects that are ruling the competition. Whereas scientists compete for the prestige, it is very relevant how this ‘prestige’ is defined. According to what our interviewees declared, the competition for increasing own bibliometric indicators does not coincide with the

competition for quality: this researchers' perception of a strong mismatch between visibility dynamics and quality dynamics has been very well reported by Aknes and Rip (2009) in their study based on the Norwegian community. One of the very aims of our contribution is to go deep into researchers' 'folk theories', and give a look into their daily habits, in order to improve our awareness and understanding of this crucial point.

Another issue that needs to be clarified is the role of scientists in the origin this scenario. We thank Rip and Stöckelová for both having alerted us on the risk to be misunderstood: we do not at all consider the scientists as 'hapless victims' of the bibliometrics-based evaluation. On the contrary, we are convinced that this practice, as we wrote in our essay, 'relies on rules and practices which scientists have deeply accepted' and that its validation 'relies on the wide acceptance and diffusion within the scientific community so that bibliometrics is substantially self-sustained by its broad application'. But we of course need to stress more this point: none of the interviewees perceived the problem in terms of a dialectic clash between 'scientists' and a duty top-down imposed by scientific policies.

Scientists themselves are ruling and defining their evaluation practices, creating those 'intersecting games' they like to play. But this is hearth of the matter: we would like to go beyond the simple statement of Sindermann that a scientific game is just a 'series of transactions and strategies which legitimately enhances progress in the many interpersonal relations which surround the act of doing good science' (Sindermann 2001), turning this statement into a question regarding one of such games, namely bibliometrics-based evaluation, which is assuming an overcoming importance. In other words, our interviews try to investigate how the rules of the scientific games are impacting 'on the act of doing good science' suggesting the risk that what should be one of the instruments to achieve information inside and about the research systems, becomes a quality assessment exercise up to the individual level, so adapting not only the strategies but the same idea scientists have about 'good science'.

Which Evaluation for a Socially Accountable Science?

From Rip and Stöckelová's commentaries, we got the impression that our conclusions may have been misunderstood. For this reason, we try to rephrase them in this section. As we explained, bibliometric indicators have been originally introduced as an aid to organising and accessing scientific knowledge. The practice of counting citations was intended to help to get oriented in the scientific publications. Our interviewees helped us to understand that, if the measure becomes the aim, the instrument is distorted and is no longer useful for the aim for which it was intended. The lesson we learnt from our study is that it is urgent to encourage a discussion on how bibliometric indicators can continue to be useful for the aim they were originally conceived, and if and how they can be used for evaluation purposes without generating monsters.

Therefore, the issue is how to frame bibliometrics-based evaluation in the general problem of accountability. In this respect this work has nothing to do with accepting or refusing any kind of evaluation but to refer the issue to the real stakeholders and tackle the even more general problem of the social accountability of science. We strongly

believe that science must be accountable towards society, and for this reason we believe that the rules that the scientific community gives to itself have to be continuously questioned and improved in order to guarantee a better science for the society. About this issue it is illuminating the work of Paula Stephan and we must acknowledge Arie Rip for suggesting her book: just for instance the risk aversion of many scientists is shown to be driven by the game of costs and incentives rising from the funding systems and job market. This strongly impacts on the possibility of a really transformative research to be performed and thus to have a significant return to society from the investments in research and development. A similar investigation should be rephrased on the epistemic level wondering whether the strong normative role of bibliometrics-based evaluation brings to a significant cultural and practical return to society in terms of quantity and quality of knowledge production (Stephan 2012).

Contact details: t.castellani@irpps.cnr.it, adriana.valente@cnr.it

References

- Aksnes, Dag W and Arie Rip. "Researchers' Perceptions of Citations." *Research Policy* 38, no. 6 (2009): 895–905.
- Berg, Bruce L and Howard Lune. *Qualitative Research Methods for the Social Sciences*, Vol. 5. Boston, MA: Pearson, 2004.
- Bernal, John D. "Scientific Information and Its Users." *Aslib Proceedings* 12, no. 12 (1960): 432–438.
- Bradford, Samuel C. *Documentation*. London: Crosby Lockwood, 1948.
- Briggs, Charles L. *Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge University Press, 1986.
- Castellani, Tommaso, Emanuele Pontecorvo and Adriana Valente. "Epistemological Consequences of Bibliometrics: Insights from the Scientific Community." *Social Epistemology Review and Reply Collective* 3, no. 11 (2014): 1-20.
- Merton, Robert K. *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press, 1973.
- Rip, Arie. "On Epistemic Effects: A Reply to Castellani, Pontecorvo and Valente." *Social Epistemology Review and Reply Collective* 4, no.1 (2014): 47-51.
- Sindermann, Carl J. *Winning the Games Scientists Play: Strategies for Enhancing Your Career in Science*. Basic Books, 2001
- Stephan, Paula. *How Economics Shapes Science*. Cambridge, MA: Harvard University Press, 2012
- Stöckelová, Tereza. "Unspoken Complicity: Further Comments on Castellani, Pontecorvo and Valente and Rip." *Social Epistemology Review and Reply Collective* 4, no. 2 (2015): 17-20.
- Ziman, John M. "The Proliferation of Scientific Literature: A Natural Process." *Science* 208, no. 4442 (1980): 369–371.