

EVALUATING THE UNEVALUABLE OR TWO KINDS OF PEER*

Ivo De Gennaro

In this paper I would like to show the following: (1) that scientific research as such cannot be evaluated; (2) what it means that scientific research is nonetheless subjected to evaluation, even though evaluation does not capture scientific research as such; finally, (3) how the suggestiveness that research evaluation is actually evaluating scientific research as such is based on the figure of the peer, or, more precisely on a surreptitious inversion of the sense of this figure.

Ad (1). Scientific research cannot be evaluated as such. It is a fact that scientific research is being evaluated wherever it takes place. However, I maintain that, in these cases, it is not evaluated as such. This is the same as saying that research evaluation evaluates scientific research *independently* of what scientific research is; in other words, in the fact that research evaluation evaluates scientific research, the fact that it is scientific research that is being evaluated is accidental. What is the difference between evaluating something and evaluating it as such? Let us consider the case of the use of a table. If I put my lectern on the table in order to be able to give this talk, I am using the table as a table. If, on the other hand, I pick it up and throw it in the direction of some of you in order to hit you, I am using the table not as such, but as a projectile.

In order to show that research evaluation evaluates science, but not science as such, we must say what science consists in. By science we mean a hypothetical or assumption-based knowledge that knows things in such a way as to make them operatively available. The quality of science, or of scientific research, depends on the manner in which a science ascertains and assures these things as its objects, and on how it carries through this ascertainment in its practices. This ascertainment is the moment in which the object of investigation is assumed in a certain form, and thus becomes an object of scientific research in the first place. This moment, which lies at the beginning of a process of investigation, but in fact rules over the entire unfolding of a specific research endeavour, is a very delicate one. In fact, the depth and the scope of sense that any instance of scientific research is capable of, is decided precisely in this critical instant, in which a scientist, to some extent, suffers the struggle for sense and takes the responsibility for a decision

* Paper presented at the workshop on „Research Evaluation“, organized by Fabrizio Durante at the Free University of Bozen-Bolzano on May 10, 2013. The text maintains the oral style of the presentation. Much of what is said here is discussed in more detail and depth in: I. De Gennaro, Gino Zaccaria, *The Dictatorship of Value. Teaching and Research in the Planetary University* (Milan: McGraw-Hill, 2011).

of sense. The quality and fecundity of scientific research — as well as who one is as a researcher — depends crucially on the extent to which this research remains rooted in the problematic sphere from which its assumptions emerge, and on how it deals with a temptation that comes with science itself.

The temptation for scientific research consists in using this instant of ascertainment — and therefore itself as a hypothetical knowledge — *in the first place* in a way that is functional to obtaining an operative result, that is, to assuring a certain performance. What eventually constitutes itself as an object of investigation is, in this case, merely an input, or an occasion for this operation oriented towards performance, while the own sense, which this thing already has in our understanding, is not at all in view. Doing science implies maintaining a delicate balance between the commitment to the sense of what the scientist is assuming as an object, and the task of obtaining an operative knowledge and assuring a performance.

Now, the quality of scientific research *as* this exercise of balance can be *judged*, but it cannot be *evaluated*. To judge scientific research as such implies that he who judges take himself a stance in the same problematic sphere, in the same struggle for sense, and that he expose himself to the same difficulty of decision as he who carried out this research in the first place. Only in this manner can the judge eventually get a sense of the balancing act that we are looking at when we examine a so-called research output. On the other hand, evaluation — the *function*, not the single evaluator — remains outside this problematic sphere, outside this struggle for sense; it remains external to the specific scientific field and assumes scientific research in a purely operative form and for purely operative purposes. In other words, evaluation is blind with regard to the properly scientific content, or blind for science as such. In so far as research evaluation necessarily deals with an operative form (or a format) of science (i.e. with science as a value), and not with science as such, we can say that science is unevaluable.

Ad (2). What does it imply that scientific research is nonetheless subjected to evaluation? In order to answer this question we need to say something more about what evaluation is, namely, where it comes from and what it implies. We said before that research evaluation evaluates science in a purely operative form and for purely operative purposes. What does this say about the scientific character of research evaluation? In the terms of our previous characterization of science, evaluation represents, to say the least, an extreme case of a scientific approach inclined towards the attainment of an operative result, and therefore exposed to the danger of forgetting the problematic and critical nature of the instant of ascertainment, in which basic assumptions are made. In fact, if we look at advanced, increasingly mechanized forms of research evaluation, we

observe that the approach they follow is completely uprooted from a genuine sphere of ascertainment, and alienated from any struggle for sense. In devising their methods and implementing their procedures, these evaluation techniques do not appear to be themselves in need of finding a balance in the terms outlined above. The reason for this is that they are already entirely transposed into an exclusively operational environment, in which *all that counts* is an operative result, an effective performance. In other words, we are talking about a knowledge that, though it *mimics* science and uses tools taken from science (typically, mathematical and statistical tools), is itself not a science, and which however is enthroned as a judge over science, and acts as a controlling body that regulates scientific life.

The question is: Why would we put a knowledge and a procedure that is blind for science as such, a knowledge and a procedure that is itself a-scientific, in charge of science? We all know the arguments that are typically used to uphold the need for research evaluation. These arguments range from needs that are internal to a certain area of research and the community of researchers that belong to this area, to needs that come from the outside, for instance, from funding agencies or scientific policy makers, if not from “society as a whole”. What is common to all these arguments in favour of research evaluation, be they internal or external, is that they have strictly nothing to do with science as such, but are purely operative in kind. They are *purely* operative in that the need that calls for research evaluation has its origin not in a genuine, autonomous problematic sphere, but in a strictly operational demand.

No matter how far we trace back the origin of the demand for research evaluation, we always only find another operatively defined, performance oriented circuit. On the other hand, we never find a sufficient origin or ground of sense, that is, a point of rest, from which we can say: this is what research evaluation is, as such, ultimately good for. In fact, unless we content ourselves with what appears as good only on the face of it, it turns out that the “good” that calls for evaluation is the mere enhancement of operative capacity. This is a rather weird finding, which, however, the diagnostic tools of philosophy allow us to read to some extent. This diagnosis shows that the *need* that research evaluation responds to — which is also the need that somehow we own when we say: well, even though evaluating is problematic, it is nevertheless necessary — is entirely enclosed in the domain of performativity, that is, in a sphere in which performance is enhanced for the sake of enhancing performance, and operations are made for the sake of operating and making things operative.

This need requires that everything be assumed in such a way that it can always and everywhere enter the computation of a performance; in other words, that it be assumed as a value. It is a

need that does not judge, but only evaluate. It wills to evaluate as a pure will to performance, in which this will only wills itself. However, this authorless will has a compelling trait to it. In fact, while on the face of it it promises judgement, this will to evaluation has an exclusive and coercive character: it doesn't really leave us much choice. Moreover, it is a will that sees the rootedness in a problematic field, the struggle for sense, the effort to maintain a balance between this struggle and the operative appropriation of things, as a mere obstacle to its unfolding, as this very balance eludes evaluation. However, if this is true, then applying research evaluation to scientific research bears a danger. Research evaluation puts scientific research under a stress, whose aim is to uproot it from its native soil, that is, to cause it to give in entirely and definitely to the temptation to which scientific research is subjected as such, namely, that of letting the operative trait take over at the expense of the responsibility for sense.

Ad (3). The suggestiveness that research evaluation is actually evaluating scientific research as such is based on the figure of the peer. There is, however, at least one element that makes our diagnosis appear too clear-cut. It is the same element that, no matter how definite a feeling we may have that, when research evaluation is implemented, something that is of no use to science, or is even dangerous for it, is happening, causes us either to justify evaluation altogether, or at least to trust that, ultimately, the genuinely scientific concern is safeguarded. This element slyly instils and preserves the false impression that, in the end, research evaluation *does* have to do with science as such, and that it actually *does* make statements about the quality of scientific research, and not just about a derived concept or a format of science. This element is the fact that research evaluation is based on and done with the help of peers. Peers are — as reviewers or as providers of citations — at this stage the pivotal figure of any system of research evaluation.

Now, there is a good reason for research evaluation to be based on peers, and for the fact that, given that evaluation is based on peers, the appearance is maintained that, in spite of its problematic nature, it does actually have to do with science as such, and deals with it in a manner that ultimately safeguards the scientific character. This reason is that science itself is based on peers. A peer is someone who has the right to be judged by someone of equal rank. This alone makes it clear why scientific research can be carried out, and preserved in its scientific character, only by a community of peers. In fact, it is a need of science itself that any research endeavour be protected, safeguarded, kept clear from extra-scientific assessments, from demands that encroach on it and impinge its commitment to scientific truth. But it is equally a need of science that those who are equals in science — and this means: those who are equally committed to a genuine scientific spirit and responsibility, or simply to scientific truth — sustain and encourage each

other in their endeavours. Scientific research as such is necessarily a domain of *exclusively internal jurisdiction*, and this jurisdictional domain constitutes, for its own sake, a likely community of scientists whose members are peers.

Since this is so, we can see why many researchers, who possibly feel uneasy about the nature of research evaluation and its effects on research, will however say to themselves: whatever this evaluation is, as long as its implementation involves peers and is in the hand of peers, this very fact warrants that, ultimately, it is about (good) science, about truth, about the trust among colleagues. The acceptance of research evaluation in the scientific community; the claim that, if done in a sound manner, it is legitimate; the reason for which the refusal to be evaluated would appear unjustified, and would even be perceived as a gesture that excludes the one who refuses evaluation from science — all this is based on the fact that research evaluation obtains a scientific imprint from being carried out by peers. The fact that peers are used, actively or passively, for procedures of evaluation, upholds the appearance that what is being evaluated is science as such.

However, it can be shown that, in a regime of research evaluation, the very notion of peer undergoes a surreptitious inversion of sense. As a matter of fact, in this regime a peer is not any more someone who has the right to be judged by those who are of equal rank, but rather one who has the right, or rather the mandate, to evaluate others. The peer now does not any more stand for protection and encouragement, but for a somehow aggressive control function. To put it bluntly: being peers is not any more about protecting and sustaining each other, but about going after each other. Equality among peers is not any more given by an equal commitment and service to scientific truth, but by the fact that researchers are equally subjected to control, and coerced into exercising control in their turn. Peers are, in the first place, not any more members of a scientific community that unites them while safeguarding the uniqueness and irreplaceable originality of each researcher; rather, they are assembled in homogeneous groups, in which each one is isolated as a research profile that is transparent to and legible by more or less mechanical procedures of evaluation. Scientific fields are now in the first place defined by the operative needs of evaluation. Their proper name is now “evaluation panel”, or, “scientific sector”. Thus, becoming a member of a group of peers means joining an evaluation panel, that is to say, doing things which need to be scanned and computed by value-generating algorithms overseen by other wheels of the same machinery. Becoming a peer does not imply having proven oneself before the court of scientific truth. There is no selection based on merit: everybody is by default enlisted to become an evaluator, that is, a peer.

The inversion of sense from the right to be judged to the mandate to evaluate bespeaks that the very notion of peer has fundamentally changed. The judging peer, or peer in judgement, has now become the evaluating peer, or peer in evaluation. Among peers in evaluation there is no trust but suspicion, and the need of constant mutual reassurance of operative reliability; there is no solidarity in truth, but changing alliances for value; there is no encouraging each other to serve sense, but chasing each other in the pursuit of ever more effective outputs. Evaluating peers do not sustain each other's freedom of research for the sake of science; they exact from each other the conformation to evaluable research practices. By virtue of the surreptitious mutation of the judging peer into the evaluating peer, the peer acts as a Trojan horse that research evaluation uses in order to infiltrate and take over scientific research for the sake of the will to evaluation. The very figure that still appears as the guardian of the scientific character of research is therefore, in truth, a tool and a functionary of that will to performance, which, because it wills only itself, seeks to uproot scientific research from its adherence to the problematic sphere of sense that it occupies, and to turn it into a practice of self-referred operating.

What is at the basis of the surreptitious replacement of the judging peer with the evaluating peer? This replacement involves a loss of the tension toward the primary scientific responsibility, which, in turn, can only derive from an adaptation and acclimatisation to a changed notion of truth, namely, to truth as performance or effectiveness.

Who is an evaluating peer? Someone who occupies a logistical position, from which he can evaluate without knowing. Someone who is constitutively anonymous, because he does not act in his own name, but merely executes a function. Someone who is essentially irresponsible, for, unlike a judge, he never needs to exhibit the truth of what he does, and only needs to comply with set standards. Finally, someone who is intrinsically inclined towards fanaticism, because the procedures he designs and actions know no measure, and by their own logic aim at eradicating entirely whatever resists the exclusive rule of a merely operative truth.

These unknowing, nameless, irresponsible men, inclined to fanaticism, are not just some of us, but, in a sense, they are all of us who are involved in scientific research — at least as long as we don't find a way to slow down, and linger and critically reflect on our practices. This talk, in the best case, has been a contribution to doing just that.