

## EVALUATION? YES, BUT NOT THIS WAY!

## EDITORIAL

Our society, based on the sharing of information and knowledge, is also an evaluation society. The parents of school age children have the right, for example, to know the criteria on the basis of which grades are given. Almost everywhere this generates evaluation grids. A Turin high school requires the consideration of between 7 and 12 items in order to justify a “behavior grade” (that is, a grade assessing general behavior at school) on a scale from 10 to 5. Let us consider the first of such items. In order to deserve a 10, students will have to show, among other things, “total and conscious respect for the others and the educational institutions.” If students showed “total respect for the others and the educational institutions” but without consciousness, all other markers being equal they would deserve a 9. Very refined hermeneutic skills, I would think, are undoubtedly required in order concretely to discriminate between “total respect” and “total and conscious respect.” The grid is thus not immediately applicable.

Someone could also object that it is not always the case that total and conscious respect is superior to total respect. Let us imagine a saint (or a psychotic), full of love for the others and for all forms of life. It could be that such an individual’s degree of consciousness—at least in the sense of rational lucidity—is insufficient. Such a student might perhaps have such a high respect for state institutions to be moved to tears every time the student sees the school gate. On the contrary, some other schoolmate might walk by the gate indifferently and absent-mindedly, already thinking of the Latin translation waiting for the students in class; even though, obviously, this schoolmate would never think of soiling or kicking the gate. Are we sure that we should give a 9 to the fool and a 10 to the one who is aware?

Perhaps one should debate this, complicate the evaluation grid, further differentiate between total and conscious, non-conscious, infra-conscious, super-conscious respect. But defining such concepts is difficult, would require a lot of time, and probably we would not reach a shared solution. Or, given that some hermeneutic exercise is inevitable whether we employ a grid or not, we might be tempted to throw away the grid and use more synthetic concepts such as the traditional notions of *good* and *bad* (and their intermediate). If one were then to ask (with some arrogance?) what behaving *well* means, one could always reply listing between 7 and 12 items (and among them, “total and conscious respect for others and the educational institutions” could not be absent).

Some questions emerge here, though. First, whether there is a limit to interpretation; that is, whether, once we have unfolded the synthesis (behaving *well*) into an analysis of 7 to 12 items, we have really reached some clarity for us and the grid users or whether we have instead produced some new ambiguity (which is in itself a new reason for conflicts), which would in turn require disambiguation. Second, whether in general we may be at a point where we again need more synthesis than simply analysis and lists. Third, whether the epistemic option for a progressive annulment of the interpretative margins, which is a legacy of the analytic tradition, not only risks succeeding only partly

in its promise for clarity and distinction, but may also disclose itself as counterproductive. Do we really want to educate the new generations (and parents, teachers, staff and support personnel, etc.) to conceive the total meaning merely as a sum or mere computation of partial meanings? Fourth, whether educating to a step-by-step process control in the (perhaps vain) hope of eliminating approximation and arbitrariness does not ultimately amount to engineering the educational practice, which ends up neglecting nothing except the essence.

The present issue of “SpazioFilosofico” aims at discussing the impact of techniques and procedures of evaluation on the production of *knowledge* and social *reality* as well as their effect on the very definition of what *knowledge* and *reality* are within the various fields (medicine, economics, management, psychology, sociology, philosophy, public and non-public policies, etc.).

Without doubt such engineered procedures of evaluations have produced clamorous accidents, such as the notorious A1+ rating assigned by Moody’s to Lehman Brothers in March 2008 ([https://www.moody.com/research/Moodys-affirms-Lehmans-A1-rating-outlook-now-stable--PR\\_151071](https://www.moody.com/research/Moodys-affirms-Lehmans-A1-rating-outlook-now-stable--PR_151071)). The systems of research quality evaluation (RQE) have been met with vigorous protests that do not originate from the unwillingness of university professors and researchers to be evaluated (see for example, in Italy, the plea for philosophy by Roberto Esposito, Adriano Fabris, and Giovanni Reale ([http://www.lascuola.it/it/home/editrice\\_detail/un-appello-per-la-filosofia/tutte\\_le\\_news/](http://www.lascuola.it/it/home/editrice_detail/un-appello-per-la-filosofia/tutte_le_news/))).

Evaluations have always occurred. Soul-searching, for example, was a form of examined life that was deemed indispensable for the acquisition of a habit of virtue. In transcendental idealism, the chain of reflections constitutes the dialectical engine of progress by posing as its object of reflection the stage of consciousness that has just been achieved. In both examples, evaluation works not in the sense of photographing reality but rather in the sense of spurring it toward an improvement that has to do with pursuing a closer resemblance with the essence and reality of that which is under examination. Both in soul-searching and in transcendental reflection, one can notice a self-oriented tendency, though; that is, the danger of missing the target of renewal and of instead making oneself happy with the examination, with the picture—an excess of analysis and reflection can be even detrimental to the production of reality and novelty.

The question to be addressed is, first of all, the relation between reflection and novelty. Secondly, there is the issue of how evaluation techniques condition concrete modalities of knowledge production. “To evaluate” does not simply mean to photograph that which exists; in many cases, evaluations force a radical change on the research methods and objects. The evaluator becomes producer, and the evaluation techniques become techniques of knowledge production.

Evaluation and rating agencies are agencies that produce reality; they do not confine themselves to selecting the best products available on the market but rather impose a line of products (a line of economic politics, a line of politics *tout court*, etc.) to the exclusion of others. Thus, evaluation stands in a critical relation with freedom and democracy.

A unilateral evaluation grid may represent a gloomy categorial apparatus that is cast on reality so as to coerce and normalize it and that thus prevents the even minimal

appearance of a *rest* of transcendence, which is conversely that which we appreciate the most—geniality in art, intuition in economics, vision in politics, etc.

In the specific field of the humanities and the social sciences, the issue is that of constructing an epistemological matrix of production and evaluation that we are conversely still missing. The matrix should take into account specific variables such as, for example, the power of an idea. How to evaluate it? We believe that, in the matrix, the reasons of the analysis should not overwhelm those of the synthesis. Although evaluators (as well as many philosophers) think of themselves only as analysts, one should always recall that, as Goethe used to say, “the proof of a successful analysis lies in its synthesis.” This is perhaps the most immediate and natural definition of philosophy and of the sciences related to it: synthetic knowledge.

On the ground of such considerations, at least the following questions may be raised:

1. In which ways, within the specific fields, does evaluation affect production?
2. In what ways does evaluation determine *a priori* that which *must* be seen/recognized and therefore empowered (that is, according to what conception of “must”)?
3. What is the relation between analysis and synthesis in evaluation processes? What is the relation between analytic evaluation and (hyper)specialization? Can we move toward new syntheses of knowledge? How?
4. Is there a connection, and which, between the pervasiveness of the current evaluation methods, the aggressiveness of the new (financial, bureaucratic, epistemic) oligarchies, and the regression of democracy?
5. How can one evaluate *quality*? What do we mean today, and what should we mean, with “impact” of a research program?
6. What anthropology underlies the dominating evaluation processes? In what ways do they allow for considerations of the human being that are not purely mechanistic or reductivistic?
7. Can there be a conflict between evaluation and reality to be evaluated, more or less similarly to the way in which an excess of reflection can be detrimental to life?
8. Is it possible to have alternative evaluation technics in economics, politics, and the humanities? If so, which?

The challenge lays not so much in avoiding as much as in critiquing and transforming evaluation. This cannot be delegated only to the meta-evaluation technicians; rather, it demands a wide community engagement on the side of all those who have some stake in the matter, especially those who are usually kept at the margins of decision-making processes.

*Enrico Guglielminetti*  
*Angelo Miglietta*  
*Paolo Moderato*

(Translated by Silvia Benso)