

SPAZIOFILOSOFICO

1/2015

Numero 13
Valutazione



Fondatori

Enrico Guglielminetti
Luciana Regina

Comitato scientifico

Enrico Guglielminetti (Direttore)
Silvia Benso
Gianfranco Dalmaso
Ugo Perone
Luciana Regina
Brian Schroeder

© 2015 SpazioFilosofico
Tutti i diritti riservati

ISSN: 2038-6788

Gli articoli filosofici della rivista sono sottoposti a blind review. La pubblicazione è subordinata per ogni articolo all'approvazione dei valutatori anonimi esterni alla direzione e all'accoglimento di eventuali richieste di revisione.

SPAZIOFILOSOFICO

1/2015

VALUTAZIONE

a cura di Enrico Guglielminetti

INDICE

E. GUGLIELMINETTI-A. MIGLIETTA-P. MODERATO, <i>Valutazione? Ma non così! Editoriale</i>	11
E. GUGLIELMINETTI-A. MIGLIETTA-P. MODERATO, <i>Evaluation? Yes, but not This Way! Editorial</i>	15

TEORIA

P. DAHLER-LARSEN, <i>The Evaluation Society: Critique, Contestability and Skepticism</i>	21
E. GUGLIELMINETTI, <i>Valutazione e novità</i>	37
R. MANZOTTI-P. MODERATO, <i>Neuroscienze e psicologia: valutazione e paradigmi</i>	55

POLITICHE

D.M. MERTENS, <i>Philosophical Assumptions and Program Evaluation</i>	75
J.C. GREENE, <i>Evaluation as a Socio-Political Intervention</i>	87
G. SAPELLI, <i>Breve pensiero sulla valutazione di un economista roussoiano</i>	97

PRATICHE

P. LEMMA, <i>Interventi di promozione della salute: una valutazione appropriata</i>	103
P. RICCHIARDI-E.M. TORRE, <i>Valutazione della scuola e del sistema scolastico: qualità formale e qualità effettiva</i>	113
M. CASSELLA-D. MUTTI, <i>Su alcuni aspetti della valutazione della ricerca nelle scienze umane. Bibliometria e dintorni</i>	123
A. MIGLIETTA, <i>La valutazione della qualità della ricerca nelle discipline manageriali: buoni propositi, cattive pratiche e urgenti cambiamenti</i>	139
E. FRANZINI, <i>Interrogare la valutazione nelle discipline umanistiche</i>	155

STUDI

S. JACOB-W.S. AFFODEGON, <i>Conducting Quality Evaluations: Four Generations of Meta-Evaluation</i>	165
---	-----

VALUTAZIONE

VALUTAZIONE? MA NON COSÌ!

EDITORIALE

La società dell'informazione e della conoscenza è anche una società della valutazione. I genitori dei ragazzi in età scolare hanno, per esempio, diritto a sapere in base a quali parametri i professori assegnano un voto. Nascono così un po' ovunque griglie di valutazione. Un liceo torinese prevede da 7 a 12 *items* per giustificare ciascun voto di condotta da 10 a 5. Consideriamo il primo *item*: per meritarsi un 10, lo studente dovrà mostrare, tra le altre cose, «pieno e consapevole rispetto degli altri e delle istituzioni scolastiche»; se invece dovesse mostrare solo «pieno rispetto degli altri e delle istituzioni scolastiche», senza consapevolezza, e se gli altri indicatori andassero di pari passo, meriterebbe il 9. Ci vuole, immagino, una raffinata ermeneutica per distinguere in concreto la fattispecie del pieno rispetto da quella del pieno e consapevole rispetto. La griglia quindi non è immediatamente applicabile. Qualcuno potrebbe anche obiettare che non sempre il rispetto pieno e consapevole è superiore al pieno rispetto. Immaginiamo un santo (o uno psicotico), ebbro di amore per gli altri e per tutte le forme di vita. Forse il suo grado di consapevolezza – perlomeno nel senso della lucidità razionale – potrebbe essere insufficiente. Magari ha tanto rispetto per le istituzioni dello Stato che piange di commozione ogni volta che vede il cancello della scuola, mentre il suo più consapevole compagno lo oltrepassa con distrazione, già pensando alla versione di latino che lo aspetta, anche se – ovviamente – non si sognerebbe mai di imbrattarlo o di prenderlo a calci. Siamo sicuri di voler dare il 9 al folle e il 10 al consapevole? Forse bisognerebbe allora discutere, complicare la griglia di valutazione, fare differenze ulteriori tra il pieno rispetto consapevole, non consapevole, infra-consapevole, sopra-consapevole (è difficile definire questi concetti, occorrerebbe molto tempo, ed è probabile che non si giungerebbe comunque a una soluzione condivisa). Oppure – visto che comunque l'esercizio di ermeneutica è inevitabile, griglia o non griglia – si potrebbe essere tentati di buttare a mare la griglia, e di affidarsi a concetti maggiormente sintetici, come quelli più tradizionali di comportamento *buono* o *cattivo* (e intermedi). Che se poi qualcuno chiedesse (con protervia?) che cosa vuol dire comportarsi *bene*, si potrebbe sempre rispondere elencando da 7 a 12 *items* (tra i quali un «pieno e consapevole rispetto degli altri e delle istituzioni scolastiche» non potrebbe mancare).

E tuttavia, alcune questioni si pongono qui: 1) se vi sia appunto un limite all'interpretazione, se cioè svolgendo la sintesi (comportarsi *bene*) in un'analisi di 7 o 12 *items*, abbiamo davvero fatto chiarezza, per noi e per l'utenza, o se non abbiamo piuttosto prodotto nuova ambiguità (che costituisce a sua volta un motivo di conflittualità), che richiederebbe di essere disambiguata; 2) se in generale non torniamo ad avere piuttosto bisogno di sintesi, che non (solo) di analisi ed elencazioni; 3) se l'opzione epistemica per un tendenziale azzeramento dei margini interpretativi, figlia della cultura analitica, non solo rischi di non mantenere che in parte le proprie promesse

di chiarezza e distinzione, ma possa addirittura rivelarsi controproducente: davvero vogliamo educare le nuove generazioni (e i genitori, i docenti, il personale tecnico-amministrativo...) a non sapere concepire il senso complessivo se non come una somma o una mera computazione di significati parziali?; 4) se educare al controllo passo per passo dei processi, nella (forse vana) speranza di eliminare approssimazione ed arbitrio, non costituisca un'ingegnerizzazione dell'atto educativo, che rischia di non trascurarne alcunché, eccetto l'essenza.

Questo numero intende discutere l'impatto delle tecniche e procedure di valutazione sulla produzione del *sapere* e della *realtà* sociale, nonché sulla definizione stessa di ciò che sono il *sapere* e la *realtà* in ambiti differenti (medicina, economia, management, psicologia, sociologia, filosofia, politiche pubbliche e non...).

È indubbio che tali procedure ingegnerizzate abbiano dato luogo a clamorosi infortuni, come la famosa A 1 con *outlook* positivo elargita da Moody's a Lehman Brothers nel marzo del 2008 (<https://www.moody.com/research/Moodys-affirms-Lehmans-A1-rating-outlook-now-stable--PR-151071>). Gli stessi sistemi di VQR (valutazione della qualità della ricerca) sono andati incontro a vibrante proteste, che non sembrano derivare dall'indisponibilità di professori e ricercatori universitari a essere valutati (si veda ad esempio, in Italia, l'appello per la filosofia di R. Esposito, A. Fabris, G. Reale: http://www.lascuola.it/it/home/editrice_detail/un-appello-per-la-filosofia/tutte-le-news/).

La valutazione c'è sempre stata. L'esame di coscienza, per esempio, era una forma di vita esaminata, che si riteneva fosse indispensabile per acquisire l'abito della virtù. La catena delle riflessioni, ponendo ogni volta a oggetto lo stadio di consapevolezza appena raggiunto, costituisce, nell'idealismo trascendentale, il motore dialettico del progresso. In entrambi questi esempi la valutazione non serve a fotografare l'esistente ma a sollecitarlo verso un miglioramento che ha che fare con il perseguimento di una maggiore somiglianza con l'essenza e con la verità di ciò che viene esaminato. Sia nell'esame di coscienza sia nella riflessione trascendentale, però, possiamo riscontrare una tendenza egologica della valutazione, il rischio di mancare l'obiettivo del rinnovamento e di accontentarsi dell'esame, della fotografia: un eccesso di analisi e di riflessione può addirittura nuocere alla produzione di realtà e di novità.

La questione che si pone è quindi in primo luogo quale sia il rapporto tra la riflessione e la novità. In secondo luogo, si pone il problema di come le tecniche di valutazione condizionino le concrete modalità di produzione del sapere. "Valutare" non significa semplicemente fotografare quello che c'è, ma – in molti casi – costringe a una mutazione radicale dei metodi e degli oggetti stessi della ricerca. Il valutatore diventa quindi produttore, le tecniche di valutazione tecniche di produzione del sapere.

Se le agenzie di valutazione e *rating* sono agenzie di produzione della realtà, che non si limitano a selezionare i prodotti migliori disponibili sul mercato, ma impongono una linea di prodotti (una linea di politica economica, una linea politica *tout court*...) a esclusione di altre, la valutazione intrattiene un rapporto critico con la libertà e con la democrazia.

Una griglia di valutazione unilaterale può costituire un greve apparato categoriale, gettato sulla realtà per coartarla e per normalizzarla, impedendo anche solo che si affacci

quel *resto* di trascendenza, che pure nei differenti ambiti è quanto apprezziamo di più: il genio nell'arte, l'intuizione in economia, la visione in politica...

Nell'ambito specifico delle *humanities* e delle discipline sociali, si tratta di costruire una matrice gnoseologica della produzione & valutazione, di cui ancora si avverte la mancanza. Essa dovrebbe tenere conto di variabili specifiche, come – ad esempio – la potenza di un'idea. Come valutarla? In tale matrice, le ragioni dell'analisi non dovrebbero – crediamo – sopravanzare quelle della sintesi. Sebbene i valutatori (come molti filosofi) si concepiscano solo come analisti, andrebbe sempre ricordato che – come diceva Goethe – “la prova di una analisi ben riuscita sta nella sintesi”. E forse, proprio questa è la più immediata e naturale definizione della filosofia e delle scienze che vi si connettono: sapere sintetico.

Sulla base di tali considerazioni, si pongono (almeno) le seguenti domande:

- 1) in che modo, nei differenti ambiti, la valutazione condiziona la produzione?
- 2) in che modo la valutazione determina a priori ciò che deve (e in base a quale concezione di “dovere”) essere visto/riconosciuto, e per conseguenza potenziato?
- 3) qual è il rapporto tra analisi e sintesi nei processi di valutazione? Che rapporto c'è tra valutazione analitica e (iper-)specialismo? Possiamo andare verso nuove sintesi del sapere? Come?
- 4) c'è un nesso, e se sì quale, tra la pervasività degli attuali metodi di valutazione, l'aggressività delle nuove oligarchie (finanziarie, burocratiche, del sapere), l'arretramento della democrazia?
- 5) come è possibile valutare la *qualità*? Che cosa s'intende oggi, e che cosa si dovrebbe intendere per “impatto” di un programma di ricerca?
- 6) quale antropologia è sottesa ai processi dominanti di valutazione? In che modo questi lasciano spazio a considerazioni non meramente meccanicistiche o riduzionistiche dell'essere umano?
- 7) può darsi un conflitto tra la valutazione e la realtà da valutare, più o meno a quel modo in cui un eccesso di riflessione può risultare nocivo alla vita?
- 8) sono possibili *tecniche* alternative di valutazione in economia, in politica e nelle *humanities*? Se sì, quali?

La sfida non consiste tanto nel sottrarsi alla valutazione, ma nel criticarla e trasformarla. La qual cosa non può essere demandata solo ai tecnici della meta-valutazione, ma richiede un ampio coinvolgimento comunitario di tutti coloro che avrebbero titolo a prendere la parola, specialmente di coloro che di solito sono tenuti ai margini dei processi decisionali.

Enrico Guglielminetti
Angelo Miglietta
Paolo Moderato

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). 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/)).

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)