

ASSOCIAZIONE ITALIANA di SOCIOLOGIA
Per una revisione degli indicatori di Valutazione della Qualità della Ricerca
in ambito non bibliometrico

Premessa.

La finalità emergente dal combinato-disposto del Decreto istitutivo della VQR e della legge 240/2010 assegna alla VQR una funzione esplicita di esercizio finalizzato al riparto delle risorse finanziarie tra le Università, assumendo implicitamente che meccanismi di premi/punizioni in base al “merito” scientifico siano strumenti efficaci di miglioramento del “Sistema universitario”, senza entrare nel merito delle sue funzioni, delle priorità politiche da seguire, del funzionamento effettivo di tali meccanismi.

Questo aspetto ci pare vada sottolineato preliminarmente perché esso presuppone, almeno, che una valutazione orientata al miglioramento di un “Sistema Universitario” sia possibile senza che si chiariscano profili e obiettivi del “Sistema universitario” che si vuole realizzare. Senza esplicitare quale Università vogliamo, non è possibile alcun responsabile esercizio di valutazione di quella che abbiamo, né è possibile distinguere tra risultati voluti e non voluti (perversi?) della valutazione e dei meccanismi che innesca nei comportamenti dei valutati¹.

Pur ritenendo auspicabile e opportuna un’attività di valutazione del Sistema universitario, resa ancor più necessaria dalla adesione al c.d. “processo di Bologna” e agli standard della politica di qualità per l’istruzione superiore adottati a partire dagli accordi di Bergen dai ministri europei, non si può non sottolineare quanto la macchina si sia mossa prestando più attenzione agli strumenti che agli obiettivi in tale quadro normativo².

La parabola della VQR si è così consumata nella ormai nota vicenda che ne ha “curvato” la funzione da esercizio di valutazione costruito sulla strutture e la qualità della loro produzione in “macchina” il cui prodotto finito sono alcuni parametri con cui si “pesano” le persone dei ricercatori/professori. Mentre nella proposta originaria alla VQR è stata assegnata la funzione di valutare le strutture, ad essa si stanno di fatto attribuendo funzioni che finiscono per assumere come oggetto di valutazione i ricercatori/professori individualmente considerati – e non solo per iniziativa di singoli Atenei e Dipartimenti che arrivano a usi che potrebbero essere passibili di sanzione, ma in forma istituzionale, come ad esempio nelle valutazioni dei colleghi docenti dei dottorati a fini dell’accreditamento dei corsi stessi³. Basta solo considerare questi aspetti, per rendere evidente quanto possa cambiare il significato di un indicatore, se usato per valutare la struttura o l’individuo, il processo o il prodotto.

¹ Presuppone anche in modo discutibile che l’accettazione di un processo di valutazione della ricerca comporti l’accettazione del principio “miglior valutazione-maggiore finanziamento o minore contrazione dello stesso”, che non può essere implicito in un sistema di valutazione poiché, ragionevolmente, questo potrebbe essere finalizzato ad un miglioramento perseguibile con altri mezzi e strutture di incentivi/sanzioni.

² La debolezza della direzione “politica” da parte di organi di autogoverno degli Atenei (tra cui non può più annoverarsi a livello centrale il CUN, trasformato in organo di consulenza del Ministero) e dei Ministri che si avvicendavano senza riuscire a incidere sul processo in corso, gli interventi che, attraverso Regolamenti, hanno di fatto esteso e potenziato le prerogative dell’ANVUR hanno concorso ad accentuare opacità e distorsioni nei comportamenti degli attori, individuali e istituzionali, fossero essi valutatori o valutati.

³ Entro l’insieme di indicatori con cui si riteneva di poter descrivere e valutare l’attività delle strutture, ha finito per prevalere la valutazione dei c.d. “prodotti” della ricerca, assolutamente decontestualizzati anche solo da quegli elementi che potrebbero suggerire l’estensione dell’analogia economica (peraltro criticabile), considerandoli come “fattori di produzione”. Tra questi, potrebbero essere considerati, quanto meno, le risorse finanziarie disponibili (stipendi in primis, poi

1. Per valutare la qualità della ricerca accademica

Non essendo questa la sede per discutere quale possa o debba essere il modello di Università e di sistema universitario cui tendere, per metodo non pretendiamo di dare una definizione di ciò che possa intendersi per valutazione della qualità della ricerca, ma riteniamo solo di convenire su alcuni punti⁴ a partire dai quali abbia senso parlare di qualità della ricerca e si possa concordare sulle condizioni alle quali essa diventi valutabile in relazione a certi fini.

✓ *Il primo punto è che la qualità della ricerca non può definirsi se non in relazione ad aspettative sociali nei confronti della ricerca stessa* – provenienti dai diversi destinatari, siano essi gli stessi ricercatori, i diversi “pubblici” per i quali la ricerca pretende rilevanza, (gli *stakeholders* ovvero i portatori di interessi nei confronti dei suoi risultati, delle pubblicazioni che ne derivano, delle sue ricadute in termini di nuove domande di conoscenza, di possibili applicazioni, di artefatti tecnologici di cui induce e/o dei quali richiede l’implementazione per il proprio ulteriore sviluppo. Tra questi ci sono ovviamente anche i finanziatori diretti e indiretti, primi tra tutti i *tax payers* e i politici che si fanno interpreti delle aspettative sociali e le fanno valere nei confronti di una Università pubblica, in quanto finanziata dalla fiscalità generale.

✓ *Il secondo punto* - che deriva quasi come corollario dal primo - *è che la qualità della ricerca non si misura, ma – appunto - si valuta, cioè è oggetto di giudizio in relazione a valori e aspettative sociali* che sono assunti come criteri dai suoi destinatari e che, quindi, variano rendendo *variabile e rivedibile* tale giudizio.

✓ *Il terzo punto* – conseguenza a sua volta del secondo - *è che il significato che assume il concetto di qualità della ricerca, e quindi la valutazione che ne consegue, non è context-free, bensì context-dependent*. Ne consegue che ogni atto di valutazione/giudizio in merito alla qualità di un *item* di ricerca - o di un soggetto cui sia attribuita la responsabilità di quell'*item* come elemento di un curriculum scientifico/accademico - è di per sé un compito assai astratto, qualora non si definiscano una serie di *contesti di riferimento* in relazione ai quali esercitare il giudizio valutativo, tali che lo stesso *item* possa essere considerato valutabile al livello più elevato in un contesto e come assolutamente marginale in un altro.

La conseguenza di ciò è una coerente articolazione delle specificazioni contestuali che sostengono ogni atto di valutazione/giudizio di qualità riferito alla ricerca:

➤ *Oggetto della valutazione*: l’oggetto si diversifica secondo due dimensioni: ricercatore/struttura di ricerca (intesa non come semplice aggregato di ricercatori); processo/prodotto.

➤ *Livello di aggregazione*: micro, meso, macro. Possono essere aggregate sia le strutture (a seconda che, ad esempio, si selezioni come oggetto il Dipartimento (micro) o l’Ateneo (meso) o il sistema Università (macro); possono essere aggregati gli individui⁵.

finanziamenti specifici del lavoro o della struttura), quelle di organizzazione della ricerca (team, reti interne/esterne alla struttura, isolamento o meno degli autori), quelle più propriamente amministrative (unità di personale e competenze disponibili a livello dipartimentale e/o di Ateneo a supporto dei processi di *fund raising*, di internazionalizzazione, di partecipazione ai bandi competitivi e di rendicontazione delle attività svolte, ecc.)

⁴ Si tratta di punti ispirati a risultati consolidati, sul tema della qualità della ricerca e della sua valutazione, nell’ambito dei *Social Studies of Science*. Ai nostri fini, e in relazione alle discipline di area 14, appare particolarmente utile e pertinente la sintesi concettuale in Cannavò (2009), che qui viene parzialmente ripresa e rielaborata.

⁵ E’ la scelta che si è fatta per ri-attribuire i dati 2004/2010, riferibili alla situazione pre-240, alla nuova situazione post-240 in cui i Dipartimenti sono stati disaggregati riducendoli agli individui che li componevano e riaggregati nei nuovi Dipartimenti, secondo le geometrie variabili da cui essi sono di fatto scaturiti e che hanno fatto perdere rilevanza, ai fini valutativi, alle informazioni relative alle strutture di provenienza, che pure erano state rilevate; è, ancora, la scelta che si è fatta per valutare i collegi di Dottorato: ma, ancora una volta, una cosa è l’aggregato “collegio di Dottorato”, altro è il Dottorato di ricerca come struttura per la formazione di terzo livello, in cui diventa rilevante quanto i componenti del collegio riescano a produrre sinergie che valorizzano, attraverso il riferimento reciproco, gli apporti individuali – di cui le incentivate

➤ *Contesto organizzativo-istituzionale entro cui la ricerca si realizza:* ente di ricerca, Università (pubblica/privata e istituzionalmente incardinata entro quale ordinamento/sistema universitario?), impresa. In relazione a quella prodotta in ambito universitario, si tratta di connotare la ricerca scientifica come “accademica”, specificandone i criteri di rilevanza “ulteriori” rispetto a quelli della ricerca che si produce negli Enti di ricerca non accademici e nelle imprese (quest’ultima, in particolare, in Italia sotto-finanziata dal settore privato, rispetto a quanto accade altrove, col rischio di attribuire un ruolo di supplenza a quella universitaria e al finanziamento pubblico).

➤ *Orizzonte temporale:* si tratta di una diversa scansione temporale degli obiettivi cognitivi cui la ricerca è orientata, a seconda che siano diretti al lungo, al medio o al breve termine, e ai quali può essere commisurato il raggiungimento dei risultati e la loro rilevanza (ricerca fondamentale, strategica, applicativa).

Solo chiariti questi punti preliminari, si può proporre quanto segue:

- di ribadire che l’unità di valutazione sia la struttura (plausibilmente il Dipartimento) e non l’individuo facendo attenzione a valutare prodotti e processi, contestualizzando la valutazione dei primi in relazione ai secondi;
- di rivedere la procedura di valutazione utilizzata dalla VQR attraverso un lavoro di ridefinizione di cosa sia valutabile⁶ e una più chiara e condivisa esplicitazione dei criteri di valutazione e del loro significato in relazione agli specifici contesti disciplinari;
- di ridefinire operativamente quei criteri in indicatori e metriche che ne rappresentino adeguatamente il campo di variazione;
- di stabilire con chiarezza a chi possa essere imputata la responsabilità dei giudizi di valutazione, considerando la possibilità - in assenza di alternative praticabili - di attribuirla al GEV nella sua composizione ufficiale, indipendentemente dal fatto che esso scelga di utilizzare un pool di referee di supporto o meno;
- di proporre che l’esito della procedura di valutazione di qualità dei prodotti scientifici sia motivato e articolato attraverso il ricorso a indicatori pertinenti, su ciascuna delle dimensioni richiamate nel D.M., che ne restituiscano lo specifico “profilo” di qualità.

Per le ragioni esposte va chiaramente rigettata l’obiezione che non siano proponibili modifiche a un processo di valutazione già iniziato, senza che se ne infici la comparabilità da una tornata all’altra. Infatti:

- a) non esiste un momento zero dal quale il processo di evoluzione delle strutture possa essere monitorato e valutato; sono inserite entro un processo storico di cui la valutazione non solo deve tener conto, ma di cui è essa stessa un elemento intrinsecamente non irrilevante;
- b) come già ricordato, la prima VQR ha operato su strutture (i Dipartimenti pre-240) che non esistevano più quando i suoi risultati sono stati resi noti e non sono più le stesse strutture su cui il secondo esercizio VQR si realizzerà: qualunque comparazione tra i risultati dell’una e dell’altra non potrà che essere l’esito *costruito* di un qualche algoritmo che li renda comparabili, a certe condizioni che andranno stabilite e che definiranno i limiti – da esplicitare e tenere in conto negli usi che se ne faranno - di quella comparabilità;

pubblicazioni a firma congiunta sono solo un possibile indicatore - e che comporta elementi non rilevabili a livello individuale né concettualizzabili come sintesi di proprietà degli individui che compongono il Collegio.

⁶ Oltre a una migliore specificazione di quali “prodotti” della ricerca siano valutabili, sarebbe importante anche una chiara identificazione dei “processi” attraverso cui le strutture ne sostengono l’attività di produzione e la messa a fuoco delle dimensioni e degli indicatori attraverso cui quei processi possano essere a loro volta resi valutabili.

- c) una volta recuperate le informazioni già rilevate con la prima VQR e riferibili alle strutture e ai processi di produzione della ricerca, sarà dunque possibile costruire condizioni e definire limiti di comparabilità tra la prima VQR e quella che si vorrebbe progettata in forme più consensuali e partecipate dalle comunità scientifiche,
- d) questo processo sarà agevolato se contestualmente si accelererà la costruzione di quella Anagrafe nazionale della ricerca, prevista dalla normativa, come strumento di messa a punto di una base dati trasparente, controllata e accessibile a tutti (Anagrafe Nazionale Nominativa dei Professori e dei Ricercatori e delle Pubblicazioni Scientifiche -AMPRePS, di cui all'art. 3-bis della legge 9 gennaio 2009, n.1, di conversione con modifiche del decreto legge 10 novembre 2008, n. 180);
- e) questo processo contribuirà alla semplificazione delle procedure se si farà in modo che la sperimentazione finalizzata alla messa a punto della SUA per la ricerca dipartimentale ne diventi parte integrante e sia in esso inglobato, riducendo l'onerosità (procedurale, organizzativa ed economica) della stessa VQR.

2. Le proposte operative dell'Associazione italiana di Sociologia

A) Si propone di accogliere la specificazione dei criteri di scientificità formulata dal CUN, ai fini della valutabilità delle pubblicazioni e degli altri prodotti della ricerca scientifica⁷ e si rinvia per una loro contestualizzazione ulteriore al Documento elaborato su questo tema dalla Consulta della ricerca AIS.

B) Si propone, in secondo luogo, di articolare in modo più corretto (e condiviso) i criteri indicati nella VQR, gli indicatori individuati e la metrica con la quale sono “misurati”.

E' appena il caso di premettere che:

- ✓ un indicatore è un elemento osservabile/ rilevabile/ misurabile che viene utilizzato “al posto di” un altro elemento non osservabile/rilevabile/misurabile⁸;
- ✓ la relazione tra il primo e il secondo assume, secondo un'autorevole maggioranza di autori, carattere semantico o stipulativo (a partire da Marradi, 1980), dunque rinvia a un accordo tra una comunità di parlanti o una comunità (nel nostro caso scientifica) di riferimento:
- ✓ la corrispondenza tra indicatori e concetto indicato è di norma parziale e incompleta, sicché l'utilizzo di una pluralità di indicatori copre uno spazio semantico solo in parte sovrapposto a quello del concetto;
- ✓ la metrica con cui sono “misurati” gli indicatori rappresenta un elemento costitutivo degli stessi e non una semplice operazione tecnica e neutrale e va pertanto a sua volta sottoposta ad analisi e riflessione congiunta da parte delle diverse parti interessate;
- ✓ la costruzione di un indice che sintetizzi i diversi indicatori utilizzati è a sua volta operazione che rispecchia la struttura delle relazioni sintattiche e semantiche attraverso cui essi contribuiscono al significato del concetto, così come inteso da chi lo utilizza; pertanto di norma gli indici sono costruiti con procedure stipulative e negoziate, non essendo auto evidenti in alcun modo le regole costitutive degli indici stessi⁹.

Tutto ciò premesso, si propongono le seguenti modifiche/precisazioni e modalità di misurazione in ordine alla VQR:

⁷ Consiglio Nazionale Universitario, “Proposta «Criteri identificanti il carattere scientifico delle pubblicazioni e degli altri prodotti della ricerca» ai sensi art.3-ter, comma 2, l. 9 gennaio 2009, n.1 e successive modificazioni”, Prot. 22011 del 24/10/2013.

⁸ Cfr. Dodd, 1939; Marradi, 1980; 2007; Cannavò, 1999; Parra Sajani, 2009; Palumbo, 2010

⁹ Un recente esempio concernente i ranking degli Atenei basati su una pluralità di indicatori è U-multirank, prodotto dalla Commissione Europea <http://www.umultirank.org/#!/home>

B.1 Le modalità di misurazione: si propone una scala Cantrill (0-10 o 1-10), quindi autoancorata da un minimo a un massimo di presenza sulle tre dimensioni operativizzate sub a), b), c). Questa modalità, se proposta a seguito di un'ampia esplicitazione dei descrittori di ogni indicatore, permette di disporre di un valore che è inteso allo stesso modo sia dal potenziale reviewer che da chi userà la review. Si sottolinea che questa proposta permette di collocare ogni prodotto in una ipotetica distribuzione da 0 a 10 in termini di posizionamento del contributo entro una scala di valore esattamente come richiesto dall'ANVUR nella formulazione del giudizio complessivo; in questo caso, tuttavia, il posizionamento si traduce immediatamente in un punteggio, senza ulteriori mediazioni.

B.2 I criteri di valutazione previsti dal DM possono essere ridefiniti e più analiticamente operativizzati, ricorrendo a indicatori che consentano di superare le contraddittorietà da più parti rilevate nella modalità attualmente seguita e di ricostruire lo specifico “profilo” di qualità di ciascun prodotto da valutare, non in astratto ma in riferimento allo specifico contesto scientifico in cui si colloca. L'elenco di indicatori di seguito suggeriti per operativizzare le tre dimensioni richiamate dal D.M. può essere concepito come un insieme di variabili che traducono la multidimensionalità (e la interpretabilità secondo accezioni diverse) del significato attribuito alla “qualità” dei prodotti della ricerca, senza forzarla entro schemi rigidi. A ciascun prodotto si associa alla fine un punteggio sintetico motivato e specificato dal “profilo” di qualità che risulta dagli indicatori ritenuti pertinenti, su ciascuna delle tre dimensioni, e che danno conto di quel punteggio. Si ritiene infatti – ed è bene ribadirlo - che i reviewers *debbano* dare conto dei punteggi assegnati ai prodotti riferendosi *esplicitamente* agli indicatori pertinenti al prodotto valutato, che giustificano quei punteggi dando loro un significato chiaro e univoco.

Nella motivazione del punteggio va infatti specificato non solo l'indicatore che si ritiene pertinente ai fini del punteggio assegnato, ma anche a quale delle fattispecie richiamate dal descrittore dell'indicatore si intenda fare riferimento.

a) Rilevanza/impatto, da intendersi come valore aggiunto per l'avanzamento della conoscenza nel settore e/o per la scienza in generale (rilevanza scientifico-accademica) o come valore aggiunto in termini di potenziali ricadute e applicazioni nella sfera sociale (rilevanza sociale).

Operativizzazione proposta:

Si assegni un punteggio da zero a dieci al prodotto in base al suo posizionamento rispetto ai seguenti indicatori:

- a.1 rilevanza/qualità sotto il profilo teorico o metodologico del contributo rispetto allo stato dell'arte nella disciplina;
- a.2 centralità, pertinenza e significatività della domanda di ricerca rispetto allo stato dell'arte scientifico o delle potenziali ricadute e/o applicazioni nella società sulla questione affrontata;
- a.3 adeguatezza (rispetto alla letteratura pertinente) dei riferimenti bibliografici rispetto alle esigenze della trattazione e correttezza accademica (assenza di bias particolaristici) nella loro selezione;
- a.4 consistenza/spessore del prodotto in termini di lavoro di ricerca, teorica o empirica, di cui dà conto, rilevanza/qualità della base dati costruita o del materiale bibliografico esaminato;
- a.5 nel caso di ricerca empirica, grado di replicabilità/controllabilità/*accountability* del percorso di ricerca di cui dà conto; grado di accessibilità/ispezionabilità/riutilizzabilità della base empirica; grado di esplicitazione delle elaborazioni compiute sulla base dati e relativa replicabilità /controllabilità/*accountability* dei risultati ottenuti;
- a.6 prestigio/reputazione scientifica visibilità/accessibilità/ampiezza della rete di distribuzione della sede editoriale (casa editrice /collana/ rivista) nella comunità scientifica nazionale.

Esempi:

Punteggio 10: prodotto di elevata rilevanza sul piano teorico, affronta questioni importanti per la disciplina o per la società, è corredato di riferimenti bibliografici completi e aggiornati ed è pubblicato in una sede editoriale che ne assicura la visibilità ed accessibilità.

Punteggio 10: prodotto di notevole spessore metodologico, relativo a un tema rilevante per la società, che si avvale di ampie basi di dati in parte costruite ad hoc, con procedure ben documentate e che sviluppa elaborazioni appropriate per supportare le conclusioni cui giunge, rilevanti per meglio affrontare i problemi sociali trattati.

Punteggio 0: prodotto di modesto spessore teorico o metodologico, che affronta in modo superficiale temi già meglio affrontati dalla letteratura (o dalla ricerca) corrente, con la quasi totale assenza di riferimenti bibliografici ad essa.

b) Originalità/innovazione

Operativizzazione proposta:

Si assegni un punteggio da zero a dieci al prodotto in base al suo posizionamento rispetto ai seguenti indicatori:

- b.1 originalità/innovazione sotto il profilo teorico o metodologico dell'approccio seguito o della base dati costruita e/o delle fonti utilizzate e/o dell'approfondimento analitico su di esse del contributo rispetto allo stato dell'arte nella disciplina;
- b.2 originalità/innovatività della domanda o del tema di ricerca o dei risultati, anche in vista delle potenziali ricadute/applicazioni in ambito sociale;
- b.3 originalità/innovazione nella letteratura cui attinge rispetto allo stato dell'arte sul tema/ai riferimenti più citati nella disciplina.

Esempi:

Punteggio 10: prodotto altamente innovativo sotto il profilo metodologico (affronta temi già trattati con strumenti metodologici innovativi), ovvero sotto il profilo sostantivo (affronta problemi rilevanti per la prima volta o con un approccio mai utilizzato finora) ovvero teorico (affronta sotto una prospettiva originale e in modo efficace un tema già all'attenzione della letteratura o della ricerca internazionale).

Punteggio 0: niente di nuovo sotto il profilo teorico o metodologico o delle ricadute sociali; un contributo di cui la sociologia avrebbe potuto fare a meno tranquillamente.

c) Internazionalizzazione

Operativizzazione proposta:

Si assegni un punteggio da zero a dieci al prodotto in base al suo posizionamento rispetto ai seguenti indicatori:

- c.1. capacità di posizionamento teorico o metodologico rispetto allo stato dell'arte a livello internazionale;
- c.2 profondità/incisività e innovatività del tema rispetto al dibattito internazionale;
- c.3 novità o significatività o interesse dei risultati conoscitivi prodotti rispetto alla letteratura internazionale o alle potenziali ricadute e/o applicazioni nella sfera sociale a livello europeo e/o internazionale;
- c.4 pertinenza dei riferimenti bibliografici alla letteratura internazionale rispetto alle esigenze della trattazione ed eventuale *co-autorship* con autori appartenenti a comunità disciplinari di altri contesti nazionali;
- c. 5 pubblicazione in una lingua diffusa e accessibile nella comunità scientifica di riferimento a livello internazionale e prestigio/reputazione scientifica/ visibilità/accessibilità internazionale della sede di pubblicazione (casa editrice/collana/ rivista).

Esempi:

Punteggio 10: prodotto che si colloca al centro del dibattito internazionale sul tema e che per sede editoriale o livello delle collaborazioni di altri autori, nonché per il taglio originale e innovativo promette di contribuire significativamente al dibattito internazionale.

Punteggio 0: il prodotto non ha alcuna visibilità internazionale, non si confronta con lo stato dell'arte teorico o metodologico internazionale sul tema e tratta di temi considerati poco rilevanti in ambito internazionale.

Ponderazione dei punteggi

Infine, si ritiene che nella valutazione operata dal GEV (che, come da proposta nel paragrafo precedente, è responsabile in solido della valutazione) il peso dei criteri sub a, b, e c non possa essere lo stesso. Le singole comunità scientifiche di riferimento potrebbero anche suggerire pesi diversi a tali criteri, in relazione in particolare all'importanza assegnata al criterio dell'internazionalizzazione nell'ambito del tema trattato.