

STORIA PRESENTE

VALUTARE LA VALUTAZIONE. QUALITÀ DELLA RICERCA SCIENTIFICA E «SCIENTOMETRIA»

*Pone seram, cobibe, sed quis custodiet ipsos custodes?
Cauta est et ab illis incipit uxor*
Giovenale

L'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (Anvur), nel quadro dell'elaborazione dei criteri e parametri per la valutazione dei candidati ai fini delle procedure di abilitazione scientifica nazionale (Legge Gelmini), come è noto, ha adottato l'utilizzazione delle metodologie «scientometriche» (1).

Tale decisione risponde a una tendenza, recentemente diffusasi in Italia, a presentare l'utilizzo di tali metodologie, da un lato, come un doveroso adeguamento del nostro paese a principi ormai invalsi nei paesi più avanzati (le cosiddette *best practices*, alle quali si fa spesso riferimento), e, dall'altro, come il metodo in grado di combattere i clientelismi e le pratiche «baronali» che – secondo tale tesi – avrebbero riempito le Università italiane di ricercatori non meritevoli con i concorsi tradizionali, affidati ai giudizi soggettivi dei membri delle commissioni concorsuali.

Chiunque tenti di avanzare obiezioni all'introduzione di tali metodologie viene – più o meno esplicitamente – accusato di oscurantismo scientifico e di voler difendere metodi baronali. Ovviamente, invece, i sostenitori delle metodologie «scientometriche» vengono implicitamente presentati come gli unici difensori del principio del merito e del rigore. È appena il caso di osservare

(1) Nel prosieguo sono stati presi come punti di riferimento i criteri contenuti nel documento approvato dal Consiglio Direttivo dell'Anvur il 22 giugno 2011, unicamente quale esempio di impostazione della valutazione della ricerca scientifica basato su criteri scientometrici; le considerazioni svolte tuttavia si rivolgono al principio e non alle specifiche proposte contenute nel documento Anvur.

che tali argomenti vengono fatti valere da docenti universitari che hanno ottenuto la cattedra proprio grazie a quei concorsi da essi oggi criticati.

Non è affatto vero che nei paesi più avanzati le metodologie scientometriche siano utilizzate per la selezione dei ricercatori e dei docenti universitari né, tanto meno, che tali metodologie siano accettate dalla comunità scientifica internazionale. Coloro che insistono sulla presunta utilizzazione della scientometria all'estero dovrebbero indicare con precisione quali sarebbero i Paesi in cui essa sia utilizzata e le norme o i regolamenti che la adottano.

Per quanto riguarda, in particolare, le discipline umanistiche l'uso, anche parziale, di metodi «scientometrici» è stato criticato con forza proprio dalle istituzioni scientifiche ritenute migliori sul piano internazionale. Ad esempio, in Francia l'autorevolissima rivista «Annales» ha pubblicato, nel 2008, un intervento redazionale nel quale si spiega per quali ragioni ritenga del tutto inaccettabile la proposta di classificazione delle riviste e in cui, fra l'altro, si afferma:

On peut gager toutefois que la résistance au classement des revues aurait été moins vive si ce dernier n'était pas lié à la mise en place de nouvelles formes d'évaluation des chercheurs. Le premier risque d'une telle évaluation est de mettre l'accent sur des critères strictement quantitatifs au moment même où la communauté scientifique prend conscience des limites des outils bibliométriques et de la vanité des mesures comme le «facteur d'impact», y compris dans les sciences physiques ou naturelles, sur lesquelles on prétend aligner coûte que coûte les sciences humaines et sociales. Même si la quantité de publications n'est souvent pas sans rapport avec la qualité de l'activité scientifique d'un chercheur et qu'elle peut avoir une place dans la mesure de son activité scientifique, en fournissant des indicateurs certes faillibles mais objectivables, l'absence d'une corrélation directe et le risque de céder à la facilité des méthodes quantitatives incitent à la prudence, d'autant qu'à la différence d'autres disciplines, les travaux de référence en sciences sociales ne passent pas toujours par les revues, les livres jouant un rôle fondamental dans la structuration du débat intellectuel. D'autre part, une telle conception de l'évaluation témoigne d'une méprise. Un comité de rédaction n'a pas pour fonction de distribuer des notes à la place des évaluateurs institutionnels et il n'y a pas lieu de se délester sur lui de la part la plus importante de l'évaluation, la part qualitative, à un moment où il faudrait au contraire défendre les instances collectives d'évaluation. Un comité de rédaction travaille à garantir la plus grande qualité scientifique possible des articles publiés, mais aussi à défendre une conception de la recherche qui est propre à chaque revue. Les membres de ces comités font des choix intellectuels qui ne sont pas neutres et qui s'inscrivent dans l'histoire et l'iden-

tité de chaque revue, ce qui invalide l'utilisation mécanique d'un classement des revues comme principal outil d'une évaluation des chercheurs ou des unités de recherche (2).

Anche in Italia la metodologia «scientometrica» è stata oggetto di critica da parte di molti studiosi. Riteniamo utile esporre brevemente le ragioni di fondo per le quali la cosiddetta «scientometria» non ha alcuna base scientifica come metodo privilegiato e oggettivo per la selezione e la valutazione dei ricercatori e dei docenti universitari.

* * *

È bene anzitutto chiarire le origini e le finalità dei metodi proposti.

La «scientometria» si propone di applicare il metodo delle scienze nomotetiche (o delle «scienze esatte») alla valutazione dei risultati di ricerca delle scienze stesse, facendo ricorso alle metodiche «bibliometriche». Queste ultime sono state elaborate negli anni Cinquanta allo scopo di misurare la distribuzione o «disseminazione» dei risultati della ricerca scientifica, utilizzando vari strumenti di rilevazione di tale distribuzione, fra cui in particolare la frequenza delle citazioni nelle riviste scientifiche. La finalità delle metodiche «bibliometriche» non è generalmente quella di esprimere delle valutazioni qualitative dell'attività degli studiosi, ma appunto quella di misurare la «disseminazione» dei risultati delle ricerche attraverso l'analisi delle pubblicazioni scientifiche, essenzialmente aventi forma di pubblicazione periodica (riviste scientifiche).

Negli anni recenti si è invece diffusa la tendenza a ricorrere alle metodiche «bibliometriche» per sostituire alle soggettive (e certamente imperfette) valutazioni qualitative operate dalle commissioni di docenti, sistemi quantitativi e presuntivamente oggettivi di valutazione dei risultati della ricerca. In tal modo, secondo le intenzioni di coloro che sostengono l'uso di tali metodiche, non vi dovrebbero essere più valutazioni soggettive, corrotte da opinioni personali, ma solo comparazioni fra dati numerici asettici e incorrotti.

A tale scopo è stato introdotto, inizialmente nella prassi di alcune discipline, l'uso di metodiche «bibliometriche» (*Impact Factor*, *H-Index* e simili)

(2) *Classer, évaluer*, in «Annales. Histoire. Sciences Sociales», 63, 2008, 6, pp. 1-4.

che dovrebbero sostituirsi alle valutazioni eminentemente soggettive, tipiche della prassi accademica. Più recentemente, come si è detto, l'utilizzazione delle metodiche «scientometriche» è stata prescelta come prassi istituzionalizzata sia per le valutazioni dell'attività di ricerca di Dipartimenti universitari, sia per le valutazioni individuali della produzione scientifica di docenti o aspiranti docenti a fini concorsuali e non solo in relazione alle scienze nomotetiche, ma anche – sia pur con particolari adattamenti – per discipline cosiddette umanistiche o sociali. In particolare l'Anvur ha proposto l'adozione di metodiche «bibliometriche» per le valutazioni dei docenti nelle procedure di abilitazione scientifica nazionale (3). L'Anvur ha differenziato le metodiche «scientometriche» da utilizzare rispettivamente per i settori scientifici inclusi nelle Aree CUN da 1 a 9 (inclusi alcuni settori delle aree da 10 a 14), ai quali si applicano i metodi basati sul numero delle citazioni e dell'«indice H», e i settori inclusi nelle Aree CUN da 10 a 14, ai quali si applicano invece misurazioni del numero delle pubblicazioni ponderate.

In via generale, le metodologie «scientometriche» identificano i risultati della ricerca scientifica in una serie di categorie omologate di «prodotti» discreti della ricerca scientifica, che si assumono quantitativamente paragonabili fra loro, all'interno di un dato settore di ricerca scientifica. Un ricercatore è valutato in base al numero di «prodotti» resi pubblici (e censiti da banche dati) in un determinato periodo di tempo e raffrontando tale numero con quello di altri ricercatori dello stesso campo di scienza. Ogni prodotto viene ponderato sulla base di alcuni parametri, che variano a seconda del settore scientifico, desunti dal cosiddetto «impatto» che il prodotto della ricerca ha registrato nella comunità scientifica (di qui il termine *Impact Factor*), impatto che viene «misurato» mediante la rilevazione di elementi indiretti presuntivi. Tale ponderazione consiste in un fattore elaborato attraverso l'uso di banche dati basate sul numero delle citazioni registrate da una pubblicazione (ad es. H-Index) ovvero, più semplicemente, quando tale metodo non è riconosciuto come valido per un determinato settore scientifico, mediante un semplice parametro numerico predeterminato per una determinata categoria di prodotti (ad es. articoli di rivista, monografie, ecc.).

Dunque i principi concettuali su cui si muovono i metodi in questione – a prescindere dall'esame dei metodi particolari proposti – sono due: (a) che si debbano comparare le quantità di prodotti discreti per unità di tempo di

(3) Documento Anvur approvato il 22 giugno 2011, in http://www.anvur.org/sites/anvur-miur/files/documento01_11.pdf

ciascun ricercatore; e (b) che la quantità dei prodotti debba essere ponderata sulla base dell'«impatto» del singolo prodotto sulla comunità scientifica.

Esaminiamo distintamente questi principi.

a) *Misurazione della quantità dei prodotti di ricerca in una unità di tempo*

Il punto di partenza delle valutazioni basate sul numero dei prodotti di ricerca è l'assunto che un ricercatore, se è attivo, deve produrre un dato numero minimo di risultati di ricerca in un dato tempo, al di sotto del quale viene qualificato come inattivo, e che l'attività scientifica di un ricercatore debba essere valutata comparativamente a quella degli altri ricercatori sulla base della quantità di prodotti. Ancorché i metodi proposti affermino di essere intesi a valutare la «qualità» della ricerca scientifica, tale qualità viene, *di fatto*, tradotta in numero di prodotti aventi determinate caratteristiche ovvero ponderati mediante ricorso a fattori compensativi: ad esempio, secondo i parametri proposti dall' Anvur, si deve prendere in considerazione «il numero degli articoli su riviste e di monografie censite su ISI o *Scopus* negli ultimi dieci anni» o «il numero di pubblicazioni negli ultimi dieci anni ponderato ...».

Il presupposto implicito di tale impostazione è che si possano classificare i prodotti di ricerca in una tipologia di prodotti «discreti» (un articolo o un libro, ecc.) e di considerare ogni prodotto, all'interno di una medesima categoria di prodotti, direttamente comparabile, sotto il profilo quantitativo, agli altri prodotti considerati analoghi solo per il fatto di rientrare in quella particolare categoria di prodotti.

Tale presupposto è evidentemente infondato. Un articolo non è quantitativamente confrontabile con un altro articolo (o un libro con un altro libro), se non si prendono in esame altre variabili, interne alla categoria «articolo» (o «libro»). Ad esempio la diversa complessità di un articolo o di un libro, rilevabile semplicemente dalle sue dimensioni (numero di pagine, di parole o di battute, ecc., elementi che spesso corrispondono a una maggiore o minore attività di ricerca svolta), il fatto che un articolo sia frutto di una ricerca originale ovvero sia una sintesi di ricerche altrui, la quantità di documenti, di dati o di altri elementi costitutivi della ricerca utilizzati e analizzati dal ricercatore, e così via.

Tali variabili fanno sì che la semplice sommatoria del numero di articoli di un ricercatore non sia paragonabile con la sommatoria degli articoli prodotti da un altro ricercatore e che tale paragone non sia sufficiente a concludere che uno dei due ricercatori sia più meritevole dell'altro.

Vi sono discipline per le quali il numero delle pagine o delle parole di

cui è composto un articolo su rivista è irrilevante: ad esempio in matematica, è ovvio che conta il risultato, non il numero delle parole con cui il risultato viene raggiunto. Appare tuttavia evidente che in tali casi sono quantitativamente confrontabili solo pubblicazioni che hanno dato luogo a risultati omogenei: ad esempio articoli contenenti risultati corretti non sono quantitativamente confrontabili con articoli che contengono risultati errati. La situazione è diversa per pubblicazioni aventi a oggetto la ricerca storica ovvero la critica letteraria o artistica, per le quali invece una pubblicazione di poche pagine difficilmente può avere un contenuto paragonabile – sempre sotto il profilo quantitativo – direttamente con un'altra pubblicazione, senza tener conto della sua ampiezza. Dato che si sta tentando di effettuare un paragone quantitativo fra il numero di articoli prodotti, non ha senso effettuare un confronto quantitativo fra prodotti quantitativamente disomogenei fra loro. Omologare i prodotti affermando che un articolo «equivale» per definizione quantitativamente a un altro articolo è un'operazione arbitraria che confligge con l'obiettivo di partenza, vale a dire eliminare gli arbitrii.

Ma anche a voler immaginare che gli articoli (o i libri) prodotti siano quantitativamente confrontabili fra loro, il criterio della quantità di prodotti si rivela in ogni caso inadeguato a permettere una valutazione comparativa dei ricercatori. È sufficiente considerare alcuni esempi di ricercatori che hanno dato un contributo solitamente ritenuto molto rilevante al proprio campo di ricerca per rendersene conto. *La produzione di merci a mezzo di merci* di Piero Sraffa, ad esempio, una delle opere più importanti nel campo della ricerca economica, fu concepito nel 1925 ma pubblicato a Cambridge nel 1960. Ci sono voluti ben 35 anni per pubblicare i risultati di una ricerca che ha avuto un impatto senza dubbio significativo. Ma parliamo di un'altra opera di grande importanza: il *Tractatus logico-philosophicus*, l'unica opera pubblicata in vita da Ludwig Wittgenstein. Sulla base dei criteri oggettivi proposti dall'Anvur si può supporre che Wittgenstein sarebbe stato considerato un ricercatore non meritevole, avendo pubblicato troppo pochi «prodotti» di ricerca. Molti ricercatori che hanno innovato profondamente hanno scritto poco e pubblicato ancora meno e quindi, applicando i criteri «bibliometrici» sarebbero battuti da ricercatori che hanno pubblicato prodotti molto più numerosi, ma del tutto irrilevanti per il progresso scientifico. Sul fronte opposto possiamo considerare ricercatori che hanno pubblicato «prodotti» sterminati. *Settecento riformatore* di Franco Venturi o le biografie di Cavour e Mussolini di Rosario Romeo e Renzo De Felice, ad esempio, sono opere che si articolano in numerosi tomi, ma che sarebbero verosimilmente valutate numericamente come un solo prodotto

e confrontate con la produzione numericamente più elevata ma molto meno importante (anche a voler considerare solo la quantità delle pagine stampate) di un altro ricercatore.

Tali esempi paradossali rendono evidente che il confronto fra quantità di prodotti appartenenti a una determinata categoria di prodotti omologata per definizione non è un metodo che possa garantire alcuna oggettività.

Il metodo di valutare la produttività sulla base del numero dei prodotti di ricerca pubblicati porta a un solo risultato, che è sotto gli occhi di tutti i ricercatori: la pubblicazione di un numero molto elevato di prodotti di ricerca privi di qualsiasi utilità, sia perché il loro contenuto normalmente è privo di interesse scientifico, sia perché, in ogni caso, nessuno li legge. Inserendo le parole *cold war* su *Googlescholar* si ottengono 1.080.000 risultati e su *Jstor* se ne ottengono oltre 60.000. A cosa potranno servire un milione di «prodotti» sulla guerra fredda? Chi li potrà mai leggere? Senza dire, poi, che molti tra i prodotti segnalati da *Googlescholar* appartengono alla pseudo-storiografia e che l'inserimento di una rivista nel catalogo di quella grande e sicuramente meritoria *impresa commerciale* che è *Jstor*, è, in ogni caso, subordinato *anche* al pagamento di un “dazio d'ingresso” da parte dell'editore della pubblicazione periodica.

Appare, comunque, evidente che è errato spingere i ricercatori a pubblicare quanti più prodotti di ricerca possibile, ma, al contrario, bisogna spingerli a pubblicare solo ciò che ha – o può avere – interesse per la comunità scientifica. L'adozione di criteri «bibliometrici» e numerici quale unico o prevalente metodo di valutazione dei ricercatori spinge la comunità scientifica a privilegiare le quantità anziché il valore e l'originalità dei risultati, che invece dovrebbero essere gli unici obiettivi.

b) *Metodi «scientometrici» di ponderazione dei prodotti della ricerca scientifica*

I metodi «scientometrici» dovrebbero servire proprio allo scopo di evitare la distorsione sopra indicata, poiché intenderebbero valutare *la qualità* della ricerca, ma in modo oggettivo. Tali metodi di valutazione sono fondamentalmente basati su due gruppi di criteri, entrambi indiretti e presuntivi, per desumere la «qualità» del singolo prodotto di ricerca: (i) i criteri basati sull'«impatto» del prodotto sulla comunità scientifica, attraverso il numero di citazioni che il prodotto ha ricevuto oppure attraverso il numero di citazioni che in media gli articoli pubblicati su quella rivista hanno ricevuto in passato, attribuendo al ricercatore l'*Impact Factor* della rivista sulla quale il suo prodotto viene pub-

blicato; e (ii) quelli basati sulla selezione effettuata dalle case editrici o dai comitati delle riviste scientifiche o altri similari selezioni preventive delle pubblicazioni.

Molte critiche sono state rivolte alle esistenti «banche dati» (o motori di ricerca) solitamente utilizzate per applicare tali criteri: ISI, *Scopus*, *Googlescholar*, *Publish or perish*, ecc. In realtà vi è una generale concordia sul fatto che le banche dati utilizzate ai fini «bibliometrici» presentino difetti e che quindi non siano del tutto affidabili. In sostanza, non sembra contestato né che il numero di citazioni (comunque sia rielaborato dalle banche dati stesse in forma di parametri più o meno complessi) riportato da un articolo di rivista sia inattendibile e incerto, né che tale numero possa essere influenzato da fattori controllabili o influenzabili da parte dell'autore: è ben noto che molti ricercatori adottano il metodo delle citazioni reciproche, proprio allo scopo di aumentare il rispettivo «Impact Factor».

Ma quello che qui si vuole evidenziare è che, se anche venisse creata una banca dati perfetta, in grado di fornire, in modo «scientificamente» affidabile, la quantificazione delle citazioni ricevute da un determinato prodotto di ricerca, ciò non risolverebbe il problema principale, che è insito proprio nell'assunto di poter valutare la qualità scientifica di un prodotto della ricerca sulla base del numero delle citazioni che il prodotto ha ricevuto.

Una ricerca scientifica può essere citata molte volte per le più varie ragioni, fra le quali c'è anche quella di aver rappresentato un risultato giudicato importante dalla comunità scientifica, ma possono esservi anche altre ragioni, che nulla hanno a che vedere con la qualità o l'importanza della ricerca. È evidente che una ricerca su un campo poco «battuto» da altri ricercatori riceverà, soprattutto in un primo periodo, poche citazioni, mentre una ricerca su un campo di studio sul quale molti altri ricercatori stanno lavorando riceverà molte citazioni. L'effetto probabile di un criterio come quello degli indici di citazioni su una disciplina scientifica è di concentrare un gran numero di ricercatori sullo stesso settore di studio, ciò che garantisce un elevato numero di citazioni, e di evitare settori innovativi in quanto inizialmente non verrebbero citati.

La classificazione delle riviste scientifiche sulla base delle citazioni ricevute dai prodotti che hanno pubblicato in passato non fa altro che creare un circolo vizioso: le riviste con alto indice di citazioni tendono a pubblicare articoli di ricercatori che hanno già un alto numero di citazioni e i ricercatori che hanno un alto numero di citazioni tendono a pubblicare solo su riviste che hanno un alto indice di citazioni: in tal modo è sempre più difficile per un

ricercatore nuovo entrare nel novero dei ricercatori molto citati: in base a quale criteri, infatti, la rivista con alto indice di citazioni dovrebbe accettare il lavoro di un neofita? Per quale ragione dovrebbe rischiare di abbassare il proprio indice? E come dovrebbe fare una rivista che ha un basso indice a aumentarlo, se i ricercatori che hanno un alto indice scelgono solo le riviste con un alto indice? Inoltre gli articoli pubblicati su riviste con alto indice tendono a citare altri articoli pubblicati su riviste analoghe e non quelli pubblicati su riviste con un basso indice.

Si tratta di un metodo che spinge inevitabilmente l'ambiente scientifico a creare un *establishment*, nel quale è sempre più difficile entrare, e nel quale sono ipotizzabili accettazioni per cooptazione che introducono (e l'ipotesi è tutt'altro che peregrina) principi di «ortodossia» scientifica, di appartenenza a «scuole» o, peggio, di favore personale e, in certi settori, di alleanza politica. In definitiva, come è stato da più parti rilevato (4), si tratta di un criterio che tendenzialmente lede il principio della libertà della ricerca scientifica.

Ancora più evidenti sono i motivi per i quali appaiono criticabili i metodi alternativi, solitamente proposti per le scienze umanistiche e sociali, di ponderazione dei prodotti di ricerca per valutarne induttivamente la qualità: ci riferiamo ai criteri basati sulla classificazione dei *media* sui quali il prodotto viene pubblicato, le case editrici, le riviste scientifiche, e la lingua in cui il prodotto è pubblicato. Tali metodi sono basati sul principio di valutare non il contenuto del prodotto di ricerca, bensì l'editore o la rivista, partendo dal presupposto che sia il *medium* a validare la qualità scientifica.

I parametri proposti dall'Anvur, ad esempio, prevedono che se un articolo è stato pubblicato su una rivista internazionale, «censita» dalla banca dati ISI, riceve un punteggio di 1,5, mentre se è pubblicato su una rivista nazionale riceve un punteggio di 0,5. Pertanto tre articoli pubblicati da una rivista nazionale «equivalgono», per definizione, sotto un profilo quantitativo, a un solo articolo pubblicato su una rivista «internazionale». Se una monografia è pubblicata da un editore «internazionale» ha un peso 3, mentre se è pubblicata da un editore «nazionale» ha un peso 1, con un risultato proporzionalmente identico alla valutazione degli articoli.

(4) G. GALASSO, *I piccoli editori nel ghetto*, in «Corriere della Sera», 24 febbraio 2012, su cui la replica di A. Graziosi e la contro-replica di G. Galasso, *ivi*, 2 marzo 2012; L. MASCILLI MIGLIORINI, *Il vizio dell'indice editoriale*, in «Il Sole 24Ore», 26 febbraio 2012; A. BURGIO e M. R. MARELLA, *Università, la Valutazione sbagliata*, in «Il Manifesto», 21 marzo 2012.

Anzitutto si deve rilevare che anche questo metodo, in definitiva, porta comunque a un raffronto puramente quantitativo: infatti se un ricercatore ha un numero di pubblicazioni con un basso punteggio (valutato con i metodi di cui si è detto sopra) superiore al triplo delle pubblicazioni di un altro ricercatore con un alto punteggio, il primo conseguirebbe un punteggio complessivo più alto rispetto al secondo. Pertanto il criterio proposto finisce per essere un criterio quantitativo, benché parzialmente corretto dalla ponderazione, e, in definitiva, favorisce pur sempre il ricercatore che produce molta quantità ma poca qualità (come definita sulla base del criterio anzidetto).

Passando ai criteri specifici proposti per attribuire la ponderazione, che dovrebbe riflettere la qualità relativa del singolo prodotto, le proposte elaborate dall'Anvur attribuiscono parametri differenziati essenzialmente sulla base del fatto che il prodotto valutato sia una monografia pubblicata da un editore internazionale ovvero nazionale o un articolo pubblicato su rivista internazionale (essendo definita come tale una rivista censita da ISI o *Scopus*) ovvero nazionale. Altri criteri proposti in altre sedi attribuiscono parametri differenziati (più elevati) a monografie pubblicate rispettivamente in collane di case editrici commerciali, sottoposte a selezione da parte di un comitato di redazione, ovvero articoli pubblicati in riviste che adottano il metodo della *peer review*.

Per quanto concerne i parametri basati sul carattere «internazionale» delle pubblicazioni, il primo problema che si pone è quello di comprendere per quale motivo una monografia pubblicata da un editore «internazionale» dovrebbe avere un valore qualitativo (e quindi un parametro) maggiore di una pubblicata da un editore nazionale. Le ragioni addotte dall'Anvur per giustificare tale criterio sono le seguenti: (1) che «un volume pubblicato all'estero, in lingua diversa dalla lingua madre (italiano) comporta di norma uno sforzo maggiore per l'autore rispetto a un libro pubblicato in italiano»; (2) che «esso è stato sottoposto a una selezione che inevitabilmente si è basata su una competizione più ampia e più severa in quanto per definizione più numerosi sono i concorrenti»; e (3) che «esso raggiunge grazie alla distribuzione una platea più vasta di lettori e utilizzatori, realizzando in questo modo un ampliamento della comunicazione scientifica (che è un valore in sé), una maggiore visibilità della ricerca italiana nel mondo, ma anche un impatto più incisivo della spesa pubblica in ricerca» (5).

(5) Vedi http://www.anvur.org/sites/anvur-miur/files/Documento%2002_11.pdf.

Non è difficile constatare che queste tre affermazioni non sono vere. Va premesso che non è chiaro se tali affermazioni si riferiscano a qualunque lingua diversa dall'italiano ovvero a una lingua o ad alcune lingue in particolare. Nel primo caso tutte le affermazioni dell'Anvur sopra citate sono palesemente non vere. Nel secondo caso, se ci si riferisce alle lingue maggiormente diffuse nel mondo (l'inglese, lo spagnolo, il francese), è parzialmente vera la terza affermazione, nel senso che l'uso di tali lingue rende possibile (in linea teorica) una maggiore diffusione, ma non è detto che tale maggiore diffusione si verifichi, dipendendo da numerosi fattori diversi dalla lingua (ad es.: numero di copie stampate e distribuzione effettuata dalla casa editrice).

Le altre affermazioni non sono attendibili nemmeno per le lingue più diffuse nel mondo. Non è vero che pubblicare una monografia in inglese comporti un maggiore sforzo, poiché si tratta solo di tradurla (o più spesso di farla tradurre da altri, il che comporta solo un maggiore costo economico, non uno sforzo scientificamente apprezzabile). Inoltre occorre osservare che un ricercatore candidato a una selezione in Italia potrebbe anche non essere di lingua madre italiana, e in tal caso si vedrebbe favorito da una circostanza che non indica un merito (o una qualità scientifica) maggiore. Non è neanche vero che sia «inevitabile» che la competizione per pubblicare una monografia in lingua inglese sia più severa, né che ci sia «per definizione» una selezione maggiore, in quanto più numerosi sono i concorrenti, asserzione nella quale l'uso dell'espressione «per definizione» appare quanto meno poco rigoroso. Tali affermazioni ovviamente non sono valide per una qualsiasi casa editrice che risponda al solo requisito di non essere italiana; ma non lo sono nemmeno per quelle che vengono solitamente considerate di «maggiore prestigio» (ad. es. la Princeton University Press, l'Oxford University Press, ecc.), le quali è ben vero che pubblicano monografie selezionate in base alla loro qualità scientifica, ma non solo. Basta sfogliare il catalogo di queste case editrici per avvedersi che contengono anche pubblicazioni di ricercatori sconosciuti e sovente opere di scarsissimo rilievo scientifico, accanto a quelle di alto valore scientifico.

La migliore riprova del fatto che la pubblicazione di una monografia in lingua non italiana o da parte di una casa editrice «internazionale» non è di per sé garanzia di alto livello scientifico si trova nella circostanza che, quanto meno in alcuni settori scientifici, *le migliori monografie dei migliori ricercatori italiani sono state pubblicate in italiano e non in altre lingue*. Ad esempio, *per quanto riguarda la storia moderna e contemporanea, tutte le più importanti monografie dei più noti storici italiani sono state pubblicate in lingua italiana*

e da editori italiani, nessuna in lingua diversa dall'italiano (6). La stessa considerazione può essere ripetuta anche per molti altri settori scientifici, quali ad esempio le lingue e letterature antiche. Per non parlare di settori, quali il diritto positivo, nei quali le pubblicazioni «internazionali» sono praticamente inesistenti in quanto si tratta di settori di interesse puramente nazionale.

Considerazioni analoghe valgono per le classificazioni, spesso proposte dai sostenitori di questi criteri, delle case editrici italiane in editori nazionali o locali e in editori commerciali o editori non commerciali. L'assunto su cui si basano simili distinzioni è che un editore nazionale o un editore commerciale effettua una selezione più rigorosa di uno locale e di uno non commerciale. Tale presunzione è infondata e spinta alle sue estreme conseguenze, come pare sia nelle intenzioni dell'Anvur, si configura come una vera e propria "turbativa del mercato editoriale". Gli editori commerciali perseguono, ovviamente, il lucro e pertanto pubblicano i libri che presumono si vendano e non quelli che avrebbero un limitato commercio. Alcuni editori commerciali, inoltre, domandano un contributo per la pubblicazione di un'opera. Questa pratica, tuttavia, contraddistingue anche note case editrici commerciali nazionali che, secondo alcune indiscrezioni, i responsabili dell'Anvur sarebbero intenzionati a collocare in posizione privilegiata in una futura classifica nazionale per la loro presunta fisionomia di "editori puri".

La qualità scientifica di una monografia prescinde totalmente, comunque, dalla sua vendibilità e viceversa. Ci sono molti esempi di monografie scientifiche che hanno anche avuto un successo di pubblico. Ce ne sono anche numerosi altri di monografie di alto livello scientifico che non hanno raggiunto un alto numero di vendite, ma che non per questo cessano di essere importanti contributi scientifici. Soprattutto vi sono innumerevoli esempi di libri, pubblicati da case editrici commerciali, che hanno avuto un larghissimo successo di pubblico e che hanno uno scarsissimo o addirittura nullo contenuto scientifico.

(6) Aggiungiamo che storici e maestri come Romeo, Venturi, De Felice, le cui opere principali non sono state tradotte nelle lingue maggiormente diffuse, sarebbero stati penalizzati da questo criterio. Quello che inoltre va sottolineato è, ad esempio, che la mancata diffusione degli studi di De Felice nel mondo anglosassone, caparbiamente e unicamente anglofono, ha fatto sì che la ricerca storica sul fascismo, in Inghilterra e negli Stati Uniti, si rifaccia ancora, per molta parte, alle opere di Salvemini (pubblicate in inglese, durante il periodo dell'esilio americano di questo studioso). Inutile aggiungere che, in questo caso e in altri casi, il monolinguisma dei ricercatori anglosassoni ha prodotto un grave depotenziamento della cultura storica dei loro Paesi. Varrebbe la pena, quindi, di liberarsi, una volta tanto, della così tanto proclamata volontà di apertura internazionale, che spesso nasconde il vecchio vizio italiano del cosmopolitismo provinciale, e considerare anche quest'ordine di problemi.

Neppure può essere ritenuto accettabile come criterio automatico di validazione della qualità scientifica di un «prodotto», sia esso una monografia o un articolo di rivista, la circostanza che la rivista o la collana di casa editrice sottoponga le pubblicazioni proposte a una selezione di un comitato scientifico fatta da un valutatore esterno (la cosiddetta *peer review*), invece che a quella di un comitato scientifico interno, affiancato dal comitato di direzione. Non è stato spiegato per quale motivo un nucleo di valutatori esterni (il quale può essere composto, secondo i suggerimenti dell'Anvur, anche da giovani ricercatori che non hanno ancora espresso tutte le loro potenzialità) dovrebbe avere un valore più oggettivo rispetto alla selezione da parte di un comitato editoriale interno e pubblico (dove figurano nomi d'illustri studiosi italiani o di altre Nazioni) o da parte di una commissione di concorso.

I comitati di redazione e i comitati di revisori esterni (*peer reviewers*), anonimi, delle riviste e delle collane editoriali sono scelti a insindacabile giudizio dell'editore, del direttore responsabile, o del proprietario della casa libraria. Quale garanzia può essere fornita sul rigore scientifico di tali comitati dal fatto che li abbia scelti l'editore o l'azionista della casa editrice rispetto alle commissioni di concorso? Queste ultime, quanto meno, devono pubblicare i propri giudizi motivati, esponendosi essi stessi a un giudizio negativo quando non sono in grado di addurre motivazioni adeguate per le proprie scelte. Nessuna collana libraria e nessun comitato editoriale (o *peer review*) di una rivista scientifica rende *universalmente* pubblici i criteri e le motivazioni in base ai quali una monografia o un articolo scientifico è stato accettato o respinto. Il giudizio favorevole o sfavorevole deve essere comunicato solo all'autore del prodotto e conservato per essere messo a disposizione di un'eventuale ispezione da parte di un organismo di controllo che, almeno per il momento, sembra che l'Anvur abbia né individuato né designato. Un autore, i cui lavori siano respinti, non può, tuttavia, opporre ricorso contro questa valutazione, come almeno *fino ad ora* è accaduto, invece, nelle valutazioni comparative per l'accesso all'insegnamento universitario o per il passaggio da una fascia inferiore a una superiore della carriera accademica. Infine l'anonimato di un valutatore lungi dal costituire un fattore di oggettività può produrre proprio l'effetto contrario. Il "valutatore senza volto", che può essere anche portatore di logiche dettate da interessi economici, dall'invidia, dalla gelosia, dal pregiudizio politico, ideologico, religioso, razziale, sfugge, infatti, al controllo della comunità scientifica nazionale e internazionale che sono il "giudice naturale" chiamato a esprimersi sulla validità di un prodotto scientifico e a vigilare sulla cor-

rettezza deontologica dell'attore della valutazione (7). Insomma, anche e soprattutto, in questo caso, vale l'antico adagio di Giovenale:

*Pone seram, cobibe, sed quis custodiet ipsos custodes?
Cauta est et ab illis incipit uxor*

** *

Le osservazioni qui sopra svolte, pur nella loro sinteticità, dovrebbero consentire di evidenziare gli ostacoli che si oppongono all'accettazione dei criteri sinora proposti per l'adozione di metodologie pseudo-scientifiche nella valutazione della ricerca scientifica. Non si tratta di confutare la validità di singole metodiche «scientometriche» sinora proposte (in particolare di quelle proposte dall'Anvur), ma, più in generale, di *confutare l'impianto concettuale ed epistemologico alla base di simili metodologie*.

I sostenitori dei metodi «scientometrici», di fronte a obiezioni come quelle sopra esposte, sono soliti chiedere che sia formulata una proposta alternativa di un metodo oggettivo di valutazione della ricerca scientifica, e in mancanza di tale formulazione alternativa reclamano la necessità di adottare quella da essi proposta. Tale impostazione è scientificamente inaccettabile. Se uno scienziato ha proposto una teoria che si è dimostrata errata, il fatto che contro di essa non sia stata proposta una teoria alternativa non conferisce alcun diritto alla prima teoria di essere considerata non erronea.

Non esiste un metodo oggettivo per valutare la qualità di una ricerca scientifica. Questa non è misurabile e può essere giudicata solo attraverso la sopravvivenza di una teoria scientifica alle confutazioni e la sua affermazione nella comunità scientifica, elementi non desumibili in modo automatico dal numero

(7) J. V. BRADLEY, *Pernicious Publication Practices*, in «Bulletin of the Psychonomic Society», 1981, 18, pp. 31-34; M. DAVID, *Politicizing Peer Review: Scientific Perspective*, in W. WENDY-R. STEINZOR, (eds.), *Rescuing Science from Politics: Regulation and the Distortion of Scientific Research*, Cambridge, Cambridge University Press, 2006, pp. 224-230. Gli inconvenienti del *blind referee* o del *double blind referee* (referaggio cieco; doppio referaggio cieco) possono (e dovrebbero) essere ridotti, ma non possono essere eliminati, dalla pubblicazione (a scadenza annuale o biennale) della lista dei valutatori utilizzati dal periodico o dalla collana editoriale. Questo rimedio, parziale, però, riduce il carattere anonimo dell'attore del referaggio a un vero e proprio "segreto di Pulcinella", perché la biografia scientifica del valutatore permette d'identificare, con larghi margini di certezza, quale prodotto egli ha valutato.

delle citazioni o dalla notorietà dei ricercatori. Basti qui ricordare il recente episodio della presunta scoperta che la velocità del neutrino sarebbe superiore a quella della luce, successivamente e clamorosamente smentita. La diffusione di quella «scoperta», citata su innumerevoli riviste scientifiche e su tutti i maggiori organi di stampa, non ha nulla a che vedere con la qualità e l'importanza della ricerca o con il merito dei ricercatori.

È quindi auspicabile che le istituzioni governative prendano atto dell'abbaglio costituito dall'attuale tendenza e attuino una decisa inversione. In mancanza di questo ripensamento toccherà al mondo della cultura secernere gli antidoti necessari, per neutralizzare i vecchi e nuovi agenti patogeni, in una franca e aperta discussione, dove, però nessuno deve pretendere di detenere in regime di monopolio esclusivo la verità rivelata. Come ha, infatti, scritto Giuseppe Galasso:

È in questo dibattito che sono stati sempre battuti le baronie e le fazioni culturali, i vuoti campanilismi o i dialettalismi insensati, le meschinità intellettuali, la falsa o sciocca scienza, il tradizionalismo come incongruo passatismo, e altro ancora. Senza contare le baronie e fazioni e altre negatività *che anche il nuovo sistema potrà produrre* (se, come dice qualcuno, non sta già producendo) e che, per l'appunto, solo per la «vecchia» via di confronto e dibattito potranno essere indicate e battute (8).

EUGENIO DI RIENZO

Università degli Studi di Roma – La Sapienza

FRANCESCO LEFEBVRE D'OVIDIO

Università degli Studi di Roma – La Sapienza

(8) G. GALASSO, *I piccoli editori nel ghetto*, cit.

The Authors analyse the conceptual premises of the use of “bibliometrics” for the evaluation of scientific publications, both in general, for any scientific field, and, more particularly, for social studies and humanities. The Authors do not disagree that it would be desirable, in an ideal world, to “measure” objectively and mathematically the quality of scientific (and particularly historical) research, but they contend that there exists no such method to attain this result. Bibliometrics have been introduced to “measure” with numerical and mathematical methods the “dissemination” of scientific literature, not to evaluate the quality of such literature. The principle underlying the use of bibliometric methods, such as “impact factor” or similar methods, for the evaluation of researchers is both conceptually wrong and misleading and causes distortion in the development of scientific research and publications, placing emphasis on the quantity of publications and on the purpose of obtaining citations – which may be increased, as is well known, through the recourse to reciprocal citations, or similar practices – rather than on the intrinsic quality of the research and the importance of its results. Moreover, the Authors contend that the use of valuations of research publications based on the ranking of the journal on which they are published, rather than the valuation of the publications on their own merits, is also the cause of distortions and that it damages the diffusion of less known publications. In the end, the Authors contend that the use of bibliometrics to evaluate scholars tends to limit freedom of thought and of scientific research, by entrusting implicitly the power to judge the authors of the papers to the members of panels of reviewers who tend to apply their own ideas and concepts in the selection process. In principle there is no reason to consider peer reviews by a single anonymous reviewer, selected by the editors of journals, as preferable to public and collective peer reviews by panels of scholars appointed by universities or public bodies. Still less acceptable, in the opinion of the Authors, is the principle – which has recently been adopted in Italy – to assign higher evaluations to research products that are published abroad and in foreign languages, on the wrong assumptions that it is deemed more difficult to publish in a foreign language and that such publications automatically reach a wider audience. Both these assumptions are considered by the Authors to lack any evidence and to be unfounded.