

La valutazione della ricerca nell'ambito delle Scienze dell'educazione: un problema di metodo

ROBERTO SANI

Abstract: *The evaluation culture in Italy is quite fragile and, above all, lacking in organization system. Worldwide, however, research evaluation systems adopted are subject to verification in progress and constantly called into question. The debate on this issue is based on one point: a clear awareness that there is no evaluation system "perfect" and that the model "best" will never be free from defects and limitations. This awareness is an integral part of modern culture and a serious evaluation. In addition it is stated that there is no single method for the evaluation of research, but it is now possible to have different approaches and qualitative and quantitative models, which potentially can be used independently or in synergy with each other. Referring to bibliometric indices and quantitative evaluation models which were recently proposed to the attention of scholars, this paper intends to offer a number of ideas, which enable the reader to grasp the more general trend that characterizes the recent international debate on this matter. It also wants to focus on certain critical issues and problem areas that impact – sometimes decisively – on the evolution and destiny of the university research conducted in our country, especially in the humanities.*

Riassunto: *La cultura della valutazione in Italia è assai fragile e, soprattutto, carente sotto il profilo dell'organizzazione di sistema. In tutto il mondo, tuttavia, i sistemi di valutazione della ricerca adottati sono oggetto di verifica in progress e di costante rimessa in discussione. Il dibattito su tale materia si fonda su un unico punto fermo: la chiara consapevolezza che non esiste nessun sistema di valutazione "perfetto" e che anche il modello "migliore" non sarà mai esente da difetti e limiti. Tale consapevolezza è parte integrante di una seria e moderna cultura della valutazione. Occorre altresì aggiungere che non esiste un solo metodo per la valutazione della ricerca, ma è possibile oggi disporre di diversi approcci e modelli di tipo qualitativo e quantitativo, i quali, potenzialmente, possono essere utilizzati ciascuno autonomamente o in sinergia tra di loro. Il presente contributo intende offrire, attraverso il riferimento a taluni tra gli indici bibliometrici e tra i modelli di valutazione quantitativa che recentemente sono stati proposti all'attenzione degli studiosi, una serie di spunti di riflessione che consentano di cogliere il più generale trend che caratterizza il recente dibattito internazionale su questa materia. Si vuole altresì focalizzare l'attenzione su talune criticità e su taluni nodi problematici de-*

stinati a riflettersi – talora in modo determinante – sull'evoluzione e sugli stessi destini della ricerca universitaria condotta nel nostro Paese in ambito umanistico.

Parole chiave: *valutazione della ricerca, metodo, approcci e modelli, indici bibliometrici.*

1. Vecchi e nuovi sistemi di valutazione della ricerca: un confronto tra approcci “perfettibili”

La cultura della valutazione in Italia è assai fragile e, soprattutto, carente sotto il profilo dell'organizzazione di sistema (Reale, Barbara, Costantini, 2006; Università di Macerata, 2009). Uno sguardo a quello che succede, su questo versante, in altre parti d'Europa e del mondo rivela tuttavia, che, un po' ovunque, i sistemi di valutazione della ricerca adottati sono oggetto di verifica *in progress* e di costante rimessa in discussione (si veda il caso recente e particolarmente significativo del Regno Unito) (Brinn, Jones, Pendlebury, 2001; Harnad, Carr, Brody, Oppenheim, 2003).

Il dibattito su tale materia si fonda su un unico punto fermo: la chiara consapevolezza che non esiste nessun sistema di valutazione “perfetto” e che anche il modello “migliore” non sarà mai esente da difetti e limiti (De Vries, 2001; Di Donato, 2005; De Robbio, 2008). Questa consapevolezza è parte integrante di una seria e moderna cultura della valutazione. Occorre altresì aggiungere che – specie *nell'era digitale e nella società della conoscenza dominata da internet* – non esiste un solo metodo per la valutazione della ricerca, ma, su questo versante, è possibile oggi disporre di diversi approcci e modelli di tipo qualitativo e quantitativo, i quali, potenzialmente, possono essere utilizzati ciascuno autonomamente o in sinergia tra di loro (Guédon, 2004).

Non solo. Proprio *internet* e le nuove tecnologie digitali consentono di ampliare la portata e l'efficacia degli stessi sistemi di valutazione tradizionali e accrescono le possibilità di affinare le metodologie e di accedere a modelli sempre più sofisticati e rigorosi (Guédon, 2001). È chiaro, allora, che la scelta della/e metodologia/e influisce notevolmente sull'efficacia del processo di valutazione e sulla rappresentatività dei suoi risultati (Etzkowitz, Webster, Gebhardt, Cantisano, 2000).

Com'è noto, i metodi qualitativi, come il cosiddetto *peer review* (giudizio dei pari, affidato a due o più valutatori competenti), si basano essenzialmente su una valutazione di merito realizzata da esperti esterni che operano nello stesso settore scientifico-disciplinare o in settori affini a quello oggetto della valutazione (De Vries, 2001; Pistotti, 2005; Di Donato, 2007).

I metodi quantitativi, invece – e il riferimento qui è agli “indici bibliometrici” –, si fondano sulla misurazione dei risultati prodotti dalla ricerca, allo scopo di fare emergere eventuali correlazioni statistiche (Harnad, 2008; Joint, 2008). Entrambi gli approcci – quello di tipo qualitativo e quello di carattere quantitativo – presentano limiti oggettivi ben noti (Lavorato, Marchetto, 2003; Campbell, 2008).

Nel primo caso si corre inevitabilmente il rischio di una valutazione inficiata da limiti soggettivi e caratterizzata da discrezionalità, nonché dal possibile prevalere degli orientamenti e indirizzi di ricerca prevalenti e delle scuole di pensiero dominanti; nel secondo, ci si trova di fronte ad un approccio sostanzialmente deterministico, ma solo in linea teorica autenticamente neutrale, che talora, specie in ambito umanistico, si rivela poco adatto a rappresentare la complessità e poliedricità della ricerca e dei suoi risultati. In entrambi i casi, lo sforzo costante di coloro che si avvalgono di tali metodi è quello di ridurre – e tendenzialmente di annullare – gli effetti distorti prodotti dalle incongruenze e anomalie insite nel modello, in modo da garantire la normalizzazione dei risultati attraverso la ponderazione dei dati, l'introduzione di coefficienti di riequilibrio e altri correttivi (De Vries, 2001).

Il *peer review*, senza dubbio il più antico e riconosciuto sistema di valutazione, è ancora oggi largamente utilizzato dalle principali e più prestigiose riviste scientifiche di tutto il mondo, come anche dalle più serie ed accreditate pubblicazioni periodiche accademiche del settore umanistico e nel campo delle scienze giuridiche e sociali. Ma c'è di più. Il ricorso al giudizio di *referees* anonimi si è progressivamente esteso, dalle riviste, all'intero complesso dell'editoria scientifica, come testimonia il fatto che, in diversi casi, anche l'accoglimento di monografie all'interno di prestigiose collane editoriali è sottoposto a tale pratica (Guédon, 2004).

Come è stato giustamente osservato da M. Frulli e A. Tonini, il *peer review* è caratterizzato da una serie di limiti che occorre tenere sempre ben presenti. Basterebbe qui accennare all'incidenza, difficilmente eliminabile, del fattore “soggettività” nella valutazione dei *referees*; alla possibilità, da parte degli stessi *referees* coinvolti nel processo di valutazione, di trovarsi in situazioni di conflitto d'interesse; o, più raramente, di registrare atteggiamenti arbitrari, quando non addirittura vere e proprie scorrettezze da parte di coloro che sono chiamati a valutare anonimamente un prodotto di ricerca. Per altri versi, è solo il caso di segnalare che il sistema di *refereaggio* previsto nel *peer review* risulterà adeguatamente funzionale ed efficace solo laddove il reclutamento dei *referees* potrà essere effettuato su un

ampio bacino di specialisti di alto profilo, ovvero di studiosi ben noti alla comunità scientifica internazionale per il loro rigore e la loro competenza e ampiamente motivati a svolgere il delicato lavoro di valutatori; il che, come si può ben immaginare, implica un costante monitoraggio del sistema e la determinazione di regole e di condizioni di esercizio rigorose. Da più parti, ad esempio, sono stati sollevati seri dubbi circa la possibilità di garantire il sistema del “doppio cieco” (*double-blind peer review*), ossia l’anonimato non solo del *referee* chiamato a fornire la propria valutazione, ma anche dell’autore dell’articolo o altro contributo sottoposto a giudizio, laddove ci si trovi ad operare in ambiti scientifico-disciplinari di dimensioni troppo ristrette o nei casi in cui il *referaggio* si attui attraverso il coinvolgimento di soli esperti nazionali (Frulli, Tonini, Casalini, 2009, 6-7).

In realtà, come ha efficacemente rilevato J.-C. Guéron, il *peer review* «non è altro che un biglietto d’ingresso per accedere alla competizione scientifica». Esso, infatti, pur certificando che il contributo di ricerca valutato presenta le qualità essenziali per essere proposto all’attenzione della comunità scientifica (rigore metodologico, originalità, carattere innovativo dei risultati raggiunti, ecc.), non è in grado di valutare *ex ante* l’impatto che tale contributo potrà avere sul progresso degli studi e delle ricerche in un determinato settore e, dunque, il suo valore complessivo in termini di efficacia e incidenza, come anche in termini di potenziamento di nuovi filoni d’indagine e di arricchimento di altri studi condotti su temi contigui (Guéron, 2004).

Se a ciò aggiungiamo che, come già ricordato, in assenza di altri strumenti affidabili (massime nel settore delle scienze umane e sociali, ivi comprese quelle della formazione) il ricorso all’utilizzo del *peer review* anche nella valutazione di singoli ricercatori o di gruppi di ricerca e dipartimenti scientifici ai fini del finanziamento alla ricerca ha prodotto risultati non del tutto soddisfacenti e, in qualche caso, addirittura inaccettabili, si comprende bene come da più parti sia stata segnalata la necessità – anche per il settore delle scienze umane e sociali – di avvalersi di metodi di valutazione più complessi e articolati, ad esempio abbinando la tradizionale pratica del *peer review* all’utilizzo degli indici bibliometrici, le cui potenzialità creative, nell’“era digitale” e in virtù del crescente successo di *internet*, appaiono smisurate anche rispetto a 15-20 anni fa (Harnad, 2008). E non c’è dubbio sul fatto che gli indici bibliometrici, ovvero i metodi di carattere squisitamente quantitativo, possano svolgere un ruolo estremamente importante nell’ambito della valutazione della ricerca, anche di quella a carattere umanistico e sociale.

Ma che cosa intendiamo quando parliamo di indici bibliometrici e quali vantaggi o inconvenienti possono derivare dalla loro applicazione nella valutazione della ricerca in ambito umanistico? Per cominciare a rispondere a tali quesiti, conviene prendere le mosse dall'analisi di quello che risulta essere, tra gli indici bibliometrici, senza dubbio il metodo più conosciuto e tradizionalmente utilizzato nell'ambito della valutazione della ricerca, in particolare (anche se non esclusivamente) sul versante scientifico e tecnologico. Intendo riferirmi al cosiddetto *Impact Factor* (IF), inventato nel 1955 da Eugene Garfield, un esperto di biblioteconomia della University of Pennsylvania (USA) (Pistotti, 2005; Campbell, 2008).

Com'è stato giustamente ricordato, «il problema al quale Garfield cercava una soluzione non aveva a che fare direttamente con la valutazione della ricerca e dei ricercatori: l'obiettivo era individuare un metodo in base al quale confrontare le riviste per poterne selezionare un numero ristretto, considerevole essenziale all'interno di una disciplina; in altri termini, per arrivare a costruire il pacchetto dei c.d. *core journals*. La crescita esponenziale del numero delle riviste e insieme di quello dei prezzi dei periodici, infatti, stava mettendo in gravi difficoltà le biblioteche universitarie» (Frulli, Tonini, Casalini, 2009, 7-8).

L'*Impact Factor*, in sostanza, è un parametro ricavato sui dati raccolti dall'*Institut of Scientific Information* (ISI) *Thomson*, il quale, si badi, non è un'istituzione di ricerca a carattere pubblico, ma un'impresa editoriale privata (un grande editore, in buona sostanza), la quale opera sulla base di una selezione assolutamente discrezionale, oltre che assai parziale (esclusivamente limitata, ad esempio, alle pubblicazioni periodiche in lingua inglese), delle riviste indicizzate (De Robbio, 2008), della quale selezione non è tenuta a giustificare i criteri adottati o le motivazioni alla base delle scelte operate (Frulli, Tonini, Casalini, 2009, 8).

Il principale limite dell'*Impact Factor* è che esso misura principalmente la qualità e il prestigio della rivista indicizzata, non la qualità dell'articolo che il singolo autore ha pubblicato in un determinato anno su quella stessa rivista (Lavorato, Marchetto, 2003, 16). La sostanziale costruzione dell'*Impact Factor* in base all'analisi delle citazioni, d'altra parte, fa emergere ulteriori limiti del modello adottato, che in alcuni casi risultano essere «intrinseci alla logica in base alla quale è stato pensato l'uso delle metriche citazionali» (Frulli, Tonini, Casalini, 2009, 8).

Deve essere, inoltre, sottolineato che l'*Impact Factor* presenta anche altri e non secondari limiti e problematicità. In primo luogo, è ben noto che le

citazioni considerate ai fini della determinazione del fattore d'impatto si limitano a quelle degli articoli delle riviste indicizzate; ma, soprattutto, la determinazione dell'*Impact Factor* non tiene conto delle citazioni presenti in monografie, atti di convegni, tesi di dottorato e altre pubblicazioni di carattere scientifico.

Occorre altresì aggiungere che, oltre al rischio di vedere maggiormente citati lavori qualitativamente non eccellenti, ma che si misurano con temi e filoni di ricerca più celebrati e maggiormente in voga in un determinato periodo, rispetto ad altri, di notevole rilevanza e di portata assolutamente innovativa, ma che approfondiscono scenari e problematiche più "periferici", ovvero assai meno approfonditi dai ricercatori e dagli studiosi (Lavorato, Marchetto, 2003), in qualche caso ci si potrebbe trovare di fronte al paradosso – stante il mero meccanismo di computo delle citazioni, che non entra nel merito della qualità e del valore contenutistico dei lavori citati – di vedere premiati da un elevato *Impact Factor* lavori che vengono citati "in negativo", al solo scopo di segnalarne i limiti, le incongruenze, talora la scarsa o nessuna affidabilità scientifica (De Robbio, 2008; Harnad, 2008; Frulli, Tonini, Casalini, 2009).

Infine, non bisogna dimenticare quanto rilevato da alcuni studiosi riguardo alla vulnerabilità di uno strumento come l'*Impact Factor* in presenza di strategie orchestrate ad arte da determinate "scuole di pensiero" e da appositi gruppi di studiosi e ricercatori e volte ad alterare opportunisticamente, attraverso un uso distorto delle citazioni, il valore di questo o quel contributo: si tratta del fenomeno, tutt'altro che raro o marginale in ambito accademico e scientifico, che A. Figà-Talamanca ha richiamato, qualche anno fa, attraverso l'efficace immagine delle «allegre brigate di entusiasti reciproci citanti» (Figà-Talamanca, 1999), e sul quale, più recentemente, si sono soffermati M. Frulli, A. Tonini e B. Casalini, ricordando quegli studiosi e quei «ricercatori che amano citarsi reciprocamente per rafforzare gli interessi del gruppo a cui appartengono» (Frulli, Tonini, Casalini, 2009, 9).

Del resto, l'*Impact Factor*, come è stato giustamente segnalato da M.C. Lavorato ed E. Marchetto, «influisce sulla tendenza a valorizzare linee di ricerca molto consolidate, a svantaggio di lavori innovativi e pionieristici, solitamente poco citati»; così come è largamente accertato che i contributi apparsi su «riviste di carattere divulgativo, o tendenzialmente divulgativo e multidisciplinare», sono citati più degli altri (Lavorato, Marchetto, 2003, 19).

Si comprende bene allora che, nel caso in cui l'*Impact Factor* venga utilizzato quale strumento di valutazione del lavoro scientifico di un ricer-

catore per l'assegnazione di fondi o per la progressione in carriera, la sua utilizzazione – specie se in forma esclusiva – espone l'intero processo valutativo a forti inconvenienti e ad esiti addirittura paradossali (Guédon, 2004; Harnad, 2008).

L'*Impact Factor* (IF) tuttavia, lo avevamo già ricordato, non è l'unico metodo di carattere quantitativo, l'unico indice bibliometrico esistente e attualmente a disposizione della docimologia della ricerca. Gli enormi progressi compiuti in questi ultimi anni sul versante delle tecnologie digitali e il sempre più largo utilizzo di *internet* consentono, infatti, di impiegare, su questo versante, indici bibliometrici di gran lunga più complessi, versatili e sofisticati (De Robbio, 2008). È chiaro che, in questa sede, non è possibile – e forse neanche utile – passare in rassegna in modo sistematico l'intera gamma delle nuove risorse tecnologiche e delle nuove metodiche in via di sperimentazione; tantomeno è possibile fornire un quadro esaustivo delle diverse procedure e delle complesse problematiche che attengono a questo ambito.

Qui, più semplicemente, s'intende offrire, attraverso il riferimento a taluni tra gli indici bibliometrici e tra i modelli di valutazione quantitativa che recentemente sono stati proposti all'attenzione degli studiosi, una serie di spunti di riflessione che consentano di cogliere il più generale *trend* che caratterizza il recente dibattito internazionale su questa materia. E, nel contempo, si vuole focalizzare l'attenzione su talune criticità e su taluni nodi problematici destinati a riflettersi – talora in modo determinante – sull'evoluzione e sugli stessi destini della ricerca universitaria condotta nel nostro Paese in ambito umanistico.

Un altro indice bibliometrico utilizzato come indicatore della produttività scientifica è il cosiddetto *Indice di Hirsch* o *h-index* (Jacsó, 2008), il quale, rispetto al già ricordato *Science Citation Index*, presenta almeno due vantaggi: il primo è che registra le citazioni ottenute non da una rivista, ma dalla complessiva produzione scientifica di un ricercatore; il secondo vantaggio è che non prende in considerazione solamente gli articoli, ma si ricava sulla base delle citazioni che il complesso delle pubblicazioni di un ricercatore ottiene. In altre parole, l'indice è strutturato per verificare, con un processo relativamente semplice, non soltanto la produzione, ma anche la reale influenza degli studi di un ricercatore, distinguendo coloro la cui produzione ha registrato una vasta eco e contribuito in modo significativo al progresso della ricerca in un determinato settore da coloro che invece, pur avendo pubblicato molto, hanno prodotto solo lavori di scarso interesse (Frulli, Tonini, Casalini, 2009, 9-11).

Esistono indici bibliometrici che funzionano anche con metriche diverse da quella citazionale, come, ad esempio: il numero dei “download online”, gli “indici cronometrici” di crescita e decrescita dei “download” e delle “citazioni”; il “*ranking*” delle pagine di *Google* e altri ancora (De Robbio, 2008). Con molta probabilità, essi saranno sempre più utilizzati per la valutazione della ricerca, poiché risultano essere poco costosi in termini di tempo, oltre che di denaro, e meno complicati e soggettivi del sistema del *peer review* (Guédon, 2003; Pauli, Stergiou, 2005; Goodman, Deis, 2005; 2007; Harzing, Van der Wal, 2007).

Tutto ciò, in realtà, lungi dal semplificare il nostro discorso e dal rassicurarci circa l'evoluzione sul breve e medio termine dei processi di valutazione della ricerca, complica enormemente il quadro e induce ad approfondire tutta una serie di problematiche intrinseche alla costruzione e alle finalità degli indici bibliometrici utilizzati e agli stessi *database* ai quali essi fanno riferimento (Guédon, 2004). Un fattore, questo, considerato sovente solo da un punto di vista meramente tecnico (del quale, cioè, non si valutano appieno le forti valenze di carattere politico e culturale) e, come tale, troppo spesso, e a torto, largamente sottovalutato, specie nel nostro Paese.

Su questo versante – come hanno recentemente sottolineato M. Frulli, A. Tonini e B. Casalini in un contributo di sintesi che abbiamo ampiamente utilizzato in questa sede –, è necessario avere piena consapevolezza delle finalità per cui tali indici bibliometrici «vengono costruiti e pensati, ovvero sapere cosa misurano, per capire che tipo di valutazione e giudizio della ricerca prodotta essi forniscono; e, forse, non meno rilevante, se non più importante, conoscere le caratteristiche dei *database* su cui gli indici vengono calcolati, dei criteri di inclusione ed esclusione che essi adottano e delle logiche in base alle quali sono stati costruiti» (Frulli, Tonini, Casalini, 2009, 10).

Lo *h-index* (con le sue molteplici varianti), ad esempio, può essere calcolato su *database* con caratteristiche molto diverse. In particolare, può essere ricavato da *Google Scholar*, una banca dati gestita da un motore di ricerca (*Google*, appunto) e gratuitamente disponibile in rete, nata con l'ambizione di costruire una biblioteca completa e universale (Pauli, Stergiou, 2005; Harzing, Van der Wal, 2007). Lo stesso indice, tuttavia, è calcolato oggi anche da *Scopus*, una banca dati sviluppata nell'ultimo decennio dal gigante dell'editoria internazionale Elsevier, principale concorrente del *Thomson ISI's Citation Index* (Frulli, Tonini, Casalini, 2009, 10-11).

La principale finalità di *Scopus* sembra essere quella di fare acquisire al colosso editoriale Elsevier un sorta di egemonia sul mercato delle pubbli-

cazioni scientifiche e, dunque, di limitare drasticamente lo strapotere fino ad oggi esercitato su questo versante dallo *Science Citation Index*, di *ISI-Thomson* (Guédon, 2009). Si tratta di una finalità per il cui raggiungimento sono state adottate due differenti strategie: la sostituzione del discusso *Impact Factor* con lo *h-index* nella valutazione della produzione scientifica; e la progressiva apertura del ventaglio delle pubblicazioni inserite nel *database*, nel cui ambito figurano anche pubblicazioni periodiche editate in lingue diverse dall'inglese (*ISI-Thomson* contempla prevalentemente testate in lingua inglese), riviste *open access* e diverse altre tipologie di prodotti editoriali (cfr. Jacsó 2008; Frulli, Tonini, Casalini, 2009).

Il fenomeno *Scopus* è particolarmente interessante, in quanto consente di cogliere taluni processi in atto, a livello internazionale, sul versante della valutazione della ricerca scientifica e di focalizzare l'attenzione su una serie di criticità del sistema di non poco conto (Goodman, Deis, 2005; 2007). Proprio con riferimento a *Scopus*, ad esempio, da più parti sono state espresse forti «perplexità in merito all'ambizione di una grande impresa globale, operante in un mercato chiaramente oligopolistico, di svolgere ad un tempo la funzione di editore (per altro, anche in settori estremamente delicati quali quello della medicina) e di valutatore della ricerca, dei ricercatori, dei gruppi di ricerca e delle istituzioni universitarie, ponendosi come referente per i governi e le istituzioni destinate ad assegnare e a distribuire finanziamenti. Il sospetto che possa profilarsi un qualche conflitto d'interessi è più che lecito» (Frulli, Tonini, Casalini, 2009, 11).

Sotto questo profilo, «sebbene il progetto di *Google* non sia totalmente privo di ombre, l'uso della banca dati *Google Scholar* per la produzione di indici bibliometrici rimane preferibile sotto diversi profili: non solo perché gratuita, e potenzialmente utilizzabile per produrre più indici, ma anche e soprattutto in virtù del fatto che gli algoritmi utilizzati da *Google* per classificare le pagine fanno riferimento all'uso mondiale della rete e non restituiscono un'immagine del sapere scientifico confinato al solo mondo occidentale» (*Ibidem*, 11-12; Guédon, 2003; 2009).

2. Oltre i metodi e le tecniche di valutazione della ricerca: le incognite e i rischi della ricerca italiana nel settore delle Scienze della Formazione

Le questioni di fondo che occorre porci oggi sono: come governare, e non subire più o meno passivamente, tale cambiamento? In quale misura,

e con quali strategie, infrastrutture e strumenti, è possibile gestire in modo efficace la transizione in atto verso un sistema di valutazione della ricerca in ambito umanistico – e segnatamente nel settore delle scienze della formazione – che implica una ridefinizione del modo stesso di fare ricerca e rischia di marginalizzare, di fatto, la ricerca italiana nel settore? Ma soprattutto è opportuno domandarci: su questo versante, a che punto siamo in Italia?

Recentemente, com'è noto, in vista dell'introduzione di un nuovo sistema di reclutamento dei ricercatori e docenti universitari e dell'avvio dell'esercizio di valutazione periodica della ricerca universitaria da parte dell'Agenzia Nazionale di Valutazione degli Atenei e della Ricerca (ANVUR), il Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) ha chiesto al Consiglio Universitario Nazionale (CUN) di formulare «proposte atte ad individuare alcuni indicatori di qualità relativi all'attività scientifica e di ricerca» da definire in relazione a diversi livelli concorsuali e alle peculiarità delle diverse Aree scientifiche. Ha chiesto altresì di determinare *standard* minimi di qualità ai fini delle procedure concorsuali e di proporre una razionalizzazione degli attuali settori scientifico-disciplinari (SSD).

Il CUN ha istituito degli appositi Gruppi di lavoro sulla valutazione della ricerca che hanno operato per aree omogenee, comprendenti in genere una o più aree scientifiche, coinvolgendo direttamente le Società e Consulte scientifiche rappresentative dei diversi settori e delle diverse discipline. In particolare, il Gruppo di lavoro del CUN sulla valutazione in area umanistica (aree 10 e 11), dopo avere concluso i suoi lavori, in data 21 maggio 2009 ha predisposto un documento dal titolo *Raccomandazioni finali*, che è stato sottoposto poi all'analisi e alle osservazioni delle Società e Consulte scientifiche afferenti alle Aree 10 e 11 (CUN, 2009).

Si tratta di un documento particolarmente interessante, il quale, tuttavia, si limita a proporre indicatori e criteri di tipo quantitativo, senza entrare nel merito – stante la pluralità e varietà dei settori umanistici considerati, la complessità dei saperi e delle discipline, le differenti tradizioni delle diverse comunità scientifiche – di possibili indicatori e criteri di tipo qualitativo. Per quello che qui interessa, il documento fornisce un'efficace fotografia dell'esistente e, soprattutto, avanza una serie di indicazioni e proposte di adeguamento minimale e a carattere trasversale – ossia riferite ai SSD e agli ambiti di ricerca di tutta l'area umanistica – agli standard internazionali della ricerca universitaria.

Non è possibile, in questa sede vagliare criticamente le proposte e soluzioni – peraltro provvisorie – formulate nel documento licenziato dal Gruppo

di lavoro del CUN sulla valutazione in area umanistica. Assai più utile, ai fini del nostro discorso, può rivelarsi un'analisi trasversale e in filigrana del documento, che focalizzi l'attenzione sui ritardi e le carenze riscontrate dal Gruppo di lavoro del CUN, nonché sulle più complessive criticità che caratterizzano la ricerca universitaria di stampo umanistico nelle Aree 10 e 11 rispetto agli standard internazionali. Proveremo a farlo in modo estremamente sintetico, attraverso l'individuazione dei principali nodi problematici, con specifico riferimento alla realtà dei saperi pedagogici.

Una prima constatazione: il settore delle scienze pedagogiche risulta essere ancora scarsamente internazionalizzato e, soprattutto, molto arretrato e deficitario dal punto di vista degli standard minimi internazionalmente riconosciuti:

a) una parte indubbiamente consistente della *produzione scientifica* (articoli, monografie, ecc.) è edita in riviste e collane editoriali prive di qualsiasi tipo di accreditamento, non risulta inserita nei principali *database* di carattere internazionale e non figura neppure nelle principali banche dati nazionali; in una parola, essa è inesistente, o quasi, ai fini di qualsiasi tipo di valutazione fondato sui più diffusi indici bibliometrici;

b) la quasi totalità della medesima produzione scientifica è tradizionalmente prodotta in lingua italiana. Ciò pone problemi non indifferenti alla sua *circolazione (e ricezione) in ambito internazionale*. Si tenga presente che portoghesi, spagnoli, tedeschi, svedesi, belgi, olandesi, polacchi, ungheresi, ecc. hanno ormai largamente adottato il cosiddetto “duplice livello linguistico”: a seconda delle finalità e della specifica rilevanza, gli studiosi di questi e di altri Paesi pubblicano i loro lavori in lingua inglese o nella rispettiva lingua nazionale. Nel nostro Paese l'adozione (anche) dell'inglese come lingua veicolare incontra ancora notevoli resistenze;

c) largamente carente, inoltre, è la *partecipazione a programmi di ricerca europei e a gruppi di ricerca internazionali*. Il tasso di partecipazione (e, naturalmente, il livello di successo) dei ricercatori italiani del settore delle scienze pedagogiche ai programmi di ricerca finanziati dall'Unione Europea è del tutto inadeguato, così come assai bassa è la partecipazione di unità di ricerca e di singoli ricercatori alle reti di ricerca internazionali;

d) non meno problematico appare lo scenario relativo alla *formazione dei nuovi ricercatori* attraverso il Dottorato di ricerca. L'analisi dei corsi di Dottorato di ambito pedagogico istituiti e attivati in Italia negli ultimi due decenni rivela il bassissimo livello di internazionalizzazione che caratterizza tali corsi. Su questo versante, peraltro, sussistono ulteriori fattori

di criticità connessi con la peculiare parcellizzazione dei corsi: quasi ogni sede universitaria della penisola ha il suo Dottorato di ricerca, in genere con poche borse disponibili e raramente collegato con quelli di altre sedi, la cui programmazione delle attività formative e di ricerca dei dottorandi risulta essere, in genere, assai eterogenea (dunque non comparabile) e, nella maggior parte dei casi, priva di qualsiasi standard qualitativo riferibile a parametri internazionali. Indubbiamente penalizzante, in questo ambito, è l'assenza di Dottorati di ricerca o Scuole di dottorato di carattere nazionale, coordinati e garantiti, dal punto di vista degli standard di qualità, dalle Società e Consulte scientifiche del settore;

e) Il tasso di *traduzioni all'estero* (monografie, articoli su rivista, ecc.) di lavori di ricerca editi in Italia nel settore delle scienze pedagogiche è anch'esso ancora assai limitato (in alcuni ambiti scarsissimo), mentre si registra un costante andamento del processo inverso (traduzioni in italiano di lavori stranieri).

C'è, tuttavia, un secondo scenario da tenere presente. Con riferimento alla realtà italiana, occorre segnalare che il settore delle scienze pedagogiche presenta un quadro altamente problematico anche sul versante degli strumenti e delle cosiddette infrastrutture per la divulgazione e valorizzazione della ricerca.

a) *La delocalizzazione degli strumenti di diffusione dei risultati della ricerca: il caso delle riviste scientifiche e delle collane editoriali.* Sul versante delle scienze pedagogiche (il problema, naturalmente, riguarda anche gli altri ambiti dell'area umanistica e sociale) non disponiamo ancora, in Italia, di un sufficiente numero di riviste scientifiche e di collane editoriali di carattere autenticamente internazionale, capaci cioè di attrarre studi e contributi di ricercatori stranieri e di proporsi quali autentici punti di riferimento per la comunità scientifica internazionale. Il rischio, in tempi di crescente e irreversibile internazionalizzazione della ricerca, è che, a lungo andare, si verifichi una sorta di forzata delocalizzazione dei processi di divulgazione dei risultati della ricerca (un fenomeno analogo a quello riscontrato negli ultimi decenni sul versante della produzione manifatturiera), capace di influenzare e condizionare in misura significativa la stessa selezione dei prodotti scientifici (e dei filoni d'indagine) da diffondere e/o valorizzare.

b) *Il deficit infrastrutturale della ricerca italiana nel campo delle scienze pedagogiche.* Una serie di fattori e di motivi storici, ai quali, in questa sede, non è possibile neppure far cenno, ha a lungo impedito che le Società e Consulte scientifiche del settore delle Scienze pedagogiche operassero come stru-

mento di coordinamento, valorizzazione e ulteriore qualificazione, nonché come fattore di raccordo internazionale della ricerca. Solo negli ultimi anni si è registrata su questo versante – penso al fondamentale ruolo assunto in questo ambito dalla Società Italiana di Pedagogia (SIPED) – un’opportuna inversione di tendenza. Tutto ciò ha prodotto non pochi ritardi e incertezze su un terreno chiave per lo sviluppo del settore, come, ad esempio, quello della predisposizione e gestione di infrastrutture destinate a sostenere e a qualificare internazionalmente il nostro ambito di ricerca: la costituzione di un’anagrafe nazionale della ricerca del settore; l’introduzione di un sistema di valutazione “interno” e “condiviso” della ricerca, ispirato agli standard internazionali; la creazione di *database* dei prodotti interfacciabili con quelli delle agenzie di valutazione italiane e straniere; infine, l’approntamento di strumenti e di strategie per la selezione e valorizzazione dei prodotti.

Conclusioni

Affrontare seriamente la questione del metodo in una realtà delicata e complessa quale è quella della valutazione della ricerca nell’ambito delle scienze della formazione significa, a mio avviso, interrogarsi sul fenomeno nella sua globalità, e in questo preciso contesto storico, provando certamente a fare il punto sulle metodologie e sui sistemi di valutazione oggi maggiormente utilizzati – e a lumeggiarne caratteristiche e modalità applicative, pregi e difetti –, ma anche non omettendo di riflettere sulle conseguenze che qualsivoglia processo e sistema di valutazione è destinato a produrre sulla ricerca in atto nel nostro Paese, in ambito umanistico e sociale in senso lato e sul versante delle scienze pedagogiche in modo più preciso.

Di qui la necessità di situare la riflessione in uno scenario preciso, che faccia emergere le matrici culturali e le caratteristiche storicamente condizionate del recente dibattito e dell’attuale importanza che la questione della valutazione della ricerca ha assunto nel nostro paese, nonché il complesso delle problematiche irrisolte e delle criticità che caratterizzano la ricerca italiana nel settore delle scienze della formazione e, in modo particolare, l’articolato ambito dei saperi pedagogici: problematiche e criticità che, a mio avviso, non possono essere eluse – come talvolta si è fatto in passato – in nome di una presunta peculiarità del modello italiano o con riferimento ad un’altrettanto presunta tradizione culturale e scientifica nazionale da preservare ad ogni costo.

Proprio i processi di internazionalizzazione e la conseguente globalizzazione della ricerca in atto anche nell'ambito delle discipline umanistiche e sociali inducono a guardare in modo radicalmente nuovo alle sfide che ci attendono.

Presentazione dell'Autore: Roberto Sani è ordinario di Storia dell'educazione all'Università degli Studi di Macerata. È direttore del «Centro di documentazione e ricerca sulla storia del libro scolastico e della letteratura per l'infanzia» e del «Centro di studi e documentazione per la storia dell'Università di Macerata» del medesimo Ateneo. Ha fondato e dirige la rivista scientifica internazionale *History of Education & Children's Literature*.

Bibliografia

- BRINN, T., JONES, M.J., PENDLEBURY, M. (2001), «The Impact of Research Assessment Exercises on UK Accounting and Finance Faculty», in *British Accounting Review*, 33, 333-355.
- BUTLER, L. (2003), «Explaining Australia's increased share of ISI publications. The effect of a funding formula based on publication counts», in *Research Policy*, 32, 143-155.
- COLUCCI, F.P. (2004), «Riflessioni dissonanti sull'internazionalizzazione della ricerca in psicologia», in *Giornale italiano di psicologia*, 2, 219-224.
- CAMPBELL, P. (2008), «Escape from the impact factor», in *Inter-Research Ethics in Science and Environmental Politics*, 103-107, 5-7.
- CONSIGLIO UNIVERSITARIO NAZIONALE (2009), *Gruppo di lavoro CUN sulla valutazione in area umanistica (aree 10 e 11). Raccomandazioni finali 21 maggio 2009*, in http://www.cun.it/media/104168/informacun_2010_01_056.pdf.
- DE ROBBIO, A. (2008), *Analisi citazionale e indicatori bibliometrici nel modello Open Access*, <http://eprints.rclis.org/archive/00011999/02/valutazione-23gennaio2008.pdf>.
- DE VRIES, J. (2001), «Peer Review: The Holy Cow of Science», in E.H. Fredriksson (ed.), *A Century of Scientific Publishing. A Collection of Essays*, Amsterdam-Berlin-Oxford-Tokyo-Washington, IOS Press, 231-244.
- DI DONATO, F. (2005), «Verso uno "European Citation Index for the Humanities". Che cosa possono fare i ricercatori per la comunicazione scientifica», in *Bollettino telematico di filosofia politica*, <http://purl.org/hj/bfp/50>.
- (2007), «Come si valuta la qualità nella Repubblica della Scienza? Una riflessione sul concetto di peer review», in *Bollettino telematico di filosofia politica*, <http://purl.org/hj/bfp/164>.
- ETZKOWITZ, H., WEBSTER, A., GEBHARDT, C., CANTISANO TERRA, B.R. (2000), «The future of the university and the university of the future: evolution of the ivory tower to entrepreneurial paradigm», in *Research Policy*, 29, 313-330.
- FIGÀ-TALAMANCA, A. (1999), «Come valutare "obiettivamente" la qualità della ricerca: Il caso dell'Impact Factor», in *Bollettino U.M.I. La matematica nella*

- società e nella cultura*, pp. 249-281, http://siba2.unile.it/ese/issues/7/71/1999-3_2.pdf.
- FRULLI, M., TONINI, A., CASALINI, B. (2009), *Riflessioni sullo stato dell'Università. 3. La valutazione della ricerca e della didattica*, Firenze, Facoltà di Scienze Politiche "Cesare Alfieri" dell'Università degli Studi di Firenze, 1-16, in <http://www.scpol.unifi.it/upload/sub/Gruppi.../Valutazione.Didattica.pdf>.
- GOODMAN, D., DEIS, L. (2005), «Web of Science (2004 version) and Scopus», in *The Charleston Advisor*, 1, 5-21.
- (2007), «Update on Scopus and Web of Science», in *The Charleston Advisor*, 1, 15-18.
- GUÉDON, J.-C. (1996), *Le planete cyber*, Paris, Gallimard.
- (2001), *Internet. Le monde en reseau*, Paris, Gallimard.
- (2003), «Open Access: from scientific plutocracy to the republic of science», in *IFLA Journal*, 2, 140.
- (2004), *La lunga ombra di Oldenburg: i bibliotecari, i ricercatori, gli editori e il controllo dell'editoria scientifica*, Pisa, Methexis-Plus.
- (2009), *Open access. Contro gli oligopoli nel sapere*, Pisa, ETS.
- HARNAD, S., CARR, L., BRODY, L., OPPENHEIM, C. (2003), «Mandated online RAE CVs linked to university eprint archives: enhancing UK research impact and assessment», in *Ariadne*, Issue 35, <http://www.ariadne.ac.uk/issue35/harnad/intro.html>.
- HARNAD, S. (2008), «Validating Research Performance Metrics Against Peer Rankings», in *Inter-Research Ethics in Science and Environmental Politics*, 103-107, <http://eprints.ecs.soton.ac.uk/15619/2/ese-harnad.pdf>.
- HARZING, A.W., VAN DER WAL, R. (2007), *Google Scholar: the democratization of citation analysis?*, in <http://www.harzing.com/download/gsdemo.pdf>.
- JACSÓ, P. (2008), «The Pros and cons of computing the h-index using Scopus», in *Online Information Review*, 4, 524-535.
- JOINT, N. (2008), «Bemused by bibliometrics: using citation analysis to evaluate research quality», in *Library Review*, 5, 345-357.
- LAVORATO, M.C., MARCHETTO, E. (2003), «Il giudizio degli psicologi italiani sulle riviste nazionali e internazionali», in *Giornale italiano di psicologia*, 1, 15-36.
- LAWRENCE, P.A. (2008), «Lost in publication: how measurement harms science», in *Ethics in science and environmental politics*, 1, 9-11.
- PAULI, D., STERGIOU, K.I. (2005), «Equivalence of results from two citation analyses: Thomson ISI's Citation Index and Google' Scholar service», in *Ethics in Science and Environmental Politics*, 3, 33-35.
- PISTOTTI, V. (2005), «The Impact of the "Impact Factor": A Critical Assessment», in *Economia politica. Journal of analytical and institutional economics*, 3, 317-324.
- REALE, E., BARBARA, A., COSTANTINI, A. (2006), *Peer review for the evaluation of the academic research. The Italian experience*, CERIS-CNR, in <http://www2.ceris.cnr.it/ceris/workingpaper/abstract/wp2006-15.htm>.
- UNIVERSITÀ DEGLI STUDI DI MACERATA (2009), *Se questa vi sembra una valutazione*, Macerata, EUM.