



VALUTAZIONE DELLE PUBBLICAZIONI ED EFFETTI SUL SETTORE EDITORIALE

di Piero Attanasio

L'articolo propone alcuni spunti critici sul dibattito italiano sulla valutazione della ricerca, assumendo esplicitamente un punto di vista editoriale. La valutazione delle pubblicazioni deve tenere in conto il contesto editoriale di riferimento in cui queste sono prodotte. Un esercizio valutativo è sempre riferito a modelli predefiniti, come il *peer review*, ed è utile renderlo esplicito, per considerare anche le conseguenze, in termini di incentivi, sugli specifici contesti editoriali. Un punto critico è dato dalle misure dell'impatto, in genere misurato dalle citazioni: quando riferite alla sede editoriale (come l'*impact factor*) invece che alle singole pubblicazioni o agli autori (come l'*indice h*) l'effetto sul mercato è di creare barriere all'ingresso e alla mobilità competitiva. Inoltre, il tentativo di sviluppare misure quantitative in assenza di dati, come nelle discipline umanistiche, può introdurre distorsioni molto evidenti. Considerato il contesto editoriale italiano, dove le pubblicazioni svolgono un ruolo di trasferimento culturale non distinto da quello interno al dibattito scientifico, misure concentrate sull'impatto citazionale possono indurre a privilegiare la seconda funzione, creando incentivi ad un cambiamento di paradigma editoriale.

This paper analyses some key issues in the Italian debate about the research assessment, starting from a publishing viewpoint. Using the publications in research assessment should consider the publishing context where such publications are produced. Any assessment exercise is referred to a pre-defined publishing model, such as that based on *peer review*. It is useful to make it explicit, in order to consider the effects on the publishing market in terms of incentives. A key element is in the impact measures, usually based on citations: when such measures are referred to the journal (such as the *impact factor*) or the publishing house, rather than to specific publications or authors (such as the *h index*), the effects on the market consist of the creation of barriers to entry and to competitive mobility. Furthermore, the creation of quantitative measure when there is a lack of data, like in humanities, may introduce very relevant biases. When considering the Italian publishing context, where publications have a role in supporting the cultural transfer to broader audiences that is not separated to the role within the scientific debate, measures that mainly focus on the second function tend to create incentives towards a change of the publishing paradigm.

1. UN PUNTO DI VISTA EDITORIALE

Poiché i risultati della ricerca sono spesso (non sempre) resi pubblici tramite *pubblicazioni scientifiche* e poiché le pubblicazioni sono prodotte dagli editori, ciascun processo di valutazione ha un impatto, diretto o indiretto, sul mondo editoriale.

Parlare di valutazione può voler dire molte cose. Qui mi limiterò ad affrontarne alcune, assumendo esplicitamente il punto di vista del settore editoriale. Spero che ciò sia utile sotto due profili: da un lato la conoscenza delle modalità organizzative ed economiche in cui le pubblicazioni sono prodotte in un dato contesto (quello italiano, in questo caso, nel raffronto con quello internazionale) può aiutare a definire in modo più consapevole le scelte di metodo che si è chiamati a fare; dall'altro lato un ragionamento sugli effetti delle stesse scelte sulle dinamiche competitive nel settore editoriale può ampliare l'orizzonte di analisi verso terreni che sono di interesse, credo, dello stesso mondo universitario, considerata la relazione bidirezionale di influenza che esso ha con l'editoria scientifica.

In modo particolare, mi concentrerò sul tema dell'uso di indici quantitativi all'interno del processo di valutazione. Ho qualche remora a dire che questo articolo parla di *bibliometria*, perché piuttosto cerca di leggere alcuni elementi del dibattito in corso oggi in Italia nelle diverse comunità accademiche nell'attuale, particolare congiuntura. Nel farlo, provo a discutere alcuni assunti di questo dibattito, spesso considerati scontati e che invece, a mio avviso, scontati non sono.

Vi è il rischio, in simili esercizi critici, che la *pars destruens* prevalga su quella *construens*. Rischio che vorrei prevenire con una dichiarazione iniziale di metodo: partendo dalla constatazione che non esiste un modello di valutazione ideale, ogni scelta concreta in questo ambito è una scelta tra *second best* possibili, in un dato contesto. L'esercizio critico su ciascuna metodologia è fondamentale per ordinarle in relazione a obiettivi prefissati ed effettuare una scelta. Con la consapevolezza che *nessuna metodologia* e tanto più *nessuna valutazione*, sono opzioni non disponibili.

2. QUALITÀ DELLE PUBBLICAZIONI E MODELLI EDITORIALI

La qualità, di per sé, non è un elemento misurabile secondo *criteri e parametri* oggettivi. Si tratta allora di ricorrere a *indicatori*¹ che forniscano elementi misurabili connessi (ma non coincidenti) con una *idea* di qualità². Il modello più consolidato è quello che segue il seguente ragionamento: *Un elemento della qualità di una pubblicazione è dato dal suo impatto sulla comunità scientifica di riferimento; l'impatto è a sua volta rilevabile, sia pure con approssimazione, dal numero di citazioni in altre pubblicazioni.* Ciò non significa dire, ovviamente, che la qualità coincide con le citazioni ricevute, ma solo che le citazioni sono un indicatore (indiretto, parziale, imperfetto, ma migliore di altri) della qualità.

Torneremo sul punto. Prima mi preme analizzare invece un diverso metodo che sembra emergere da numerosi documenti prodotti negli ultimi anni da gruppi di lavoro e società scientifiche di diverse discipline³, quello di misurare il livello di aderenza di una

¹ Adotterò nel corso dell'articolo la terminologia proposta dal CUN in *Proposta su «criteri e parametri per la valutazione» ai fini di cui all'Art. 16 comma 3 lettere a) e h) della Legge 30 dicembre 2010, n. 240, 9 giugno 2011*, che riporto di seguito per comodità di lettura: *Per criteri si intendono elementi di giudizio suscettibili di una valutazione di carattere qualitativo e non necessariamente misurabile). Per parametri si intendono quei particolari elementi di giudizio che sono suscettibili di una quantificazione Per indicatori si intendono gli strumenti operativi mediante i quali è resa possibile la quantificazione e quindi la misurazione dei parametri. Per valori di riferimento si intendono quei particolari valori degli indicatori da adottare come soglie il cui superamento appare necessario ... ai fini del conseguimento di una valutazione positiva.*

² Metodologicamente, è inevitabile che un indicatore di questo genere, proprio perché riferito a elementi qualitativi, lavori sulla base di *modelli ideali di qualità*, con ragionamenti indiretti. Spesso, come vedremo, si propongono strumenti che *misurano* il grado di adesione del caso concreto a tale modello ideale. In questi casi, rendere esplicito il modello di qualità sottostante è il primo strumento di validazione della tecnica che si propone di adottare.

³ Ciò accade più spesso in documenti che riguardano le scienze umane e sociali, dove le altre tecniche – generalmente definibili come bibliometriche – non sono sviluppate. Si dice in sostanza: *ci fossero i dati useremmo tecniche diverse, in assenza dobbiamo inventarci qualcosa.* Il tutto sembra guidato più dall'ansia di avere un numero, quasi che ciò di per sé oggettivasse la valutazione.

pubblicazione a un modello editoriale prefissato. Con un certo livello di approssimazione, il più delle volte il modello editoriale è quello delle riviste scientifiche che adottano un criterio di selezione tramite *peer review*. La casa editrice (la rivista, la collana) è concepita come una *sede* cui l'autore sottometta un lavoro aspettando il giudizio dei propri pari per essere ammesso. L'editore gestisce il processo di selezione ma non assume in prima persona la responsabilità della scelta. In questo senso è *al servizio* della comunità scientifica, che tuttavia si autoregola, accettando solo il giudizio di chi è giudicato come *pari* (non per forza un altro accademico, di certo un altro studioso).

È un modello nobilissimo, tipico dell'editoria scientifico-academica anglosassone. Ma non è l'unico possibile. La migliore editoria italiana ha alcune caratteristiche peculiari e affatto diverse, particolarmente evidenti nelle scienze umane. Ne ha fornito una descrizione, iperbolica ma efficace, Gian Arturo Ferrari in un recente scritto su *Il Mulino*, utilizzando la definizione largamente condivisa nel nostro paese di *editoria di cultura*:

L'editoria di cultura italiana, o all'italiana, è una creazione originale di Giulio Einaudi, la cui grandezza [sta] nella determinazione lucida e feroce con cui seppe perseguire un progetto grandioso, smisurato e forse insensato ... Einaudi aveva osato concepire un simile megalomane progetto, quello di fare di una casa editrice il ponte di comando, lo stato maggiore, la guida della cultura nazionale. Altro che University Press! La casa editrice non è al servizio dell'università, è l'università ad accodarsi alla casa editrice.

... L'idea che il soggetto della cultura fosse la casa editrice [ha trovato] altri e fertili terreni su cui impiantarsi ... La più ambiziosa di queste imprese fu, agli inizi degli anni Cinquanta, il Mulino ... Ma poi vi furono anche la Laterza di Vito Laterza, il Saggiatore fenomenologico sotto l'influenza di Enzo Paci, la Feltrinelli storico e filosofico-scientifica di Paolo Rossi e Ludovico Geymonat, la Vallecchi utopico-hegeliana al seguito di Arturo Massolo, e via via Bompiani, Edizioni di Comunità e infiniti altri ...

*Ma i casi più interessanti furono le due costole staccatesi dalla Einaudi, Boringhieri e Adelphi, che applicarono con più rigore la ricetta einaudiana, soprattutto per quanto riguarda la posizione antiaccademica o, per meglio dire, anticipatrice e ispiratrice dell'accademia.*⁴

Se l'università italiana deciderà di adottare una *misura* della qualità data dall'aderenza a un modello diverso da quello storicamente prevalente nel nostro paese si possono prevedere due effetti: dal lato dell'università si corre il rischio di introdurre errori sistematici nella valutazione delle pubblicazioni, nel senso che quella che è la migliore produzione editoriale nazionale può essere invece considerata negativamente; dal lato dell'editoria si incentiva un cambio di paradigma che, beninteso, potrebbe anche essere auspicato (non si vuole qui sostenere che il modello *all'italiana* è migliore, solo che è diverso) ma, in questo caso, lo dovrebbe essere in modo consapevole ed esplicito.

Vi è un ulteriore elemento nel dibattito italiano sulla valutazione che sembra non tener conto della realtà editoriale del nostro paese. Si dice: *devono esser giudicate solo le pubblicazioni scientifiche*, intese come quelle che diffondono i risultati delle ricerche, distinte – si suppone – dalle pubblicazioni *divulgative* così come da quelle *didattiche*. Si dà per scontato che tale distinzione sia possibile, e addirittura semplice.

Leggiamo, tuttavia, quanto ancora scrive Ferrari sul punto, così da approfittare di nuovo della chiarezza dell'iperbole.

L'editoria di cultura è stata, e per certi versi ancora è, una delle poche invenzioni italiane di grande successo internazionale. Nell'ambito editoriale forse l'unica. Negli altri Paesi la nozione stessa di editoria di cultura era, e in larga misura ancora è, incom-

⁴ G.A. Ferrari, *Editoria di cultura e cultura dell'editoria*, «Il Mulino», 2010, pp. 181-190 (corsivi miei). Lo scritto, appunto per i toni iperbolici, ha suscitato polemiche che qui tuttavia non interessano: per il presente discorso resta l'efficacia di una contrapposizione netta tra il modello di servizio e quello italiano *di cultura* nei rapporti tra editoria e università, per altro proposto in un contesto che nulla aveva a che vedere con il dibattito sulla valutazione.

*prensibile ... Esiste un'editoria accademica, buona per gli accademici ... Esiste infine un'editoria divulgativa, fatta da bravi divulgatori e compilatori che rielaborano quel che di meglio accademici e specialisti hanno fatto. Ma non c'era, e forse non c'è, il fascino dell'editoria di cultura italian style, il glamour, lo chic del bianco di Einaudi o dei pastelli di Adelphi, congiunti beninteso a impervie vette del pensiero.*⁵

Detto in altri termini: la migliore editoria di cultura italiana sta esattamente nel mezzo tra l'editoria *scientifica* e quella divulgativa. Ambisce ad ampliare il numero di destinatari dei prodotti che nascono in ambito di ricerca (ciò è più evidente nelle scienze umane e sociali) verso un pubblico colto, curioso, o professionale.

3. DEL PUBBLICO E DELLE TIRATURE (O DELLE MISURE DELL'IMPATTO)

Se si accetta che può essere considerata *di qualità* una pubblicazione che risponde a logiche editoriali diverse da quelle tipiche dell'editoria scientifica anglosassone, si dovrebbe provare a ragionare su come valutare anche queste opere.

L'occasione per questo articolo è stata un convegno nelle aule dell'Università Statale di Milano, luogo dove è nata la migliore scuola italiana di analisi critica dell'editoria rispetto al suo *pubblico*. Mi piace allora rendere omaggio a questa scuola proponendo un ragionamento basato sui due termini che negli anni Vittorio Spinazzola ha utilizzato come titolo degli annuari, da lui curati, sul tema: *Pubblico e Tirature*⁶. La chiave, a mio avviso, sta appunto nell'analisi

⁵ *Ibidem* (corsivi miei).

⁶ *Pubblico. Rassegna annuale di fatti letterari* è uscito dal 1977 al 1987 prima per il Saggiatore, dal 1982 al 1986 per Milano Libri e nel 1987 per Rizzoli; *Tirature* esce annualmente (con qualche salto) dal 1991: il primo anno per Einaudi, dal 1992 al '95 per Baldini & Castoldi e dal 1998 per il Saggiatore, in coedizione con la Fondazione Mondadori negli ultimi anni. Per inciso: le due pubblicazioni sono paradigmatiche di altri problemi che possono nascere: il Servizio Bibliotecario Nazionale indica sia *Pubblico* sia *Tirature* talvolta come *Periodico*, talvolta come *Monografia*, categorie che negli

del *pubblico* dei prodotti che si vogliono valutare, e forse le *tirature* (o meglio, la diffusione) sono un elemento quantitativo di non trascurabile valore.

Occorre però preliminarmente ragionare su *cosa* si intende misurare. Vi sono elementi diversi: *i) la qualità interna; ii) l'importanza e l'impatto nella comunità degli scienziati; iii) l'impatto al di fuori dell'accademia*⁷. La variabile più utilizzata in bibliometria è il numero di citazioni che un'opera riceve in altre pubblicazioni. È evidente che ciò misura in primo luogo il secondo elemento: l'impatto nella comunità degli scienziati. Assunto che la qualità interna è certamente valutabile ma non *misurabile* (intendendo con ciò l'attribuzione di valori numerici secondo criteri oggettivi), per mancanza di una metrica possibile, l'impatto al di fuori dell'accademia delle pubblicazioni scientifiche è un elemento su cui ci si sofferma poco.

Ciò attiene appunto al pubblico di riferimento: è composto dalla sola *comunità degli studiosi*, o – secondo una felice espressione utilizzata dal CUN⁸ – comprende le *Comunità di coloro che praticano la disciplina*? Interessa valutare la rilevanza di una pubblicazione solo in relazione al dibattito scientifico o anche nel dialogo con un pubblico colto, o con il mondo delle professioni o delle imprese (di *coloro che praticano la disciplina*, appunto)?

In ambito giuridico la cosa può essere di particolare evidenza. Le pubblicazioni scientifiche *svolgono appieno la loro funzione soltanto ove siano in grado di dialogare efficacemente con la pratica*; basarsi solo sull'impatto interno alla comunità scientifica rischia di *incentivare l'assunzione di un atteggiamento autoreferenziale*⁹. Ma allora: l'efficacia

schemi di valutazione prevedono trattamenti diversi; il vagare della pubblicazione tra cinque diversi editori rende invece evidente come l'uso della *sede editoriale* come *proxy* della qualità della pubblicazione può produrre effetti paradossali: che uno stesso annuario riceva *punteggi* diversi da un anno all'altro per ragioni che nulla hanno a che vedere con le sue caratteristiche intrinseche.

⁷ A. Baccini, *Valutare la ricerca scientifica*, Il Mulino, Bologna 2010, p. 40.

⁸ Parere del 19-11-2009, reperibile su www.cun.it.

⁹ *Proposta per la definizione comune dei criteri e l'adozione di soluzioni unitarie per l'area giuridica*, s. i. d ma 2009, mimeo.

del dialogo *con la pratica* può essere *misurato* dalle citazioni¹⁰? E, ancor più, ha senso valutare la qualità puramente sulla base dell'aderenza a modelli di *peer review*? L'editore, in questo caso, seleziona (e lo fa con l'attenzione e il rigore di chi investe il proprio) valutando appunto il potenziale pubblico che un'opera è in grado di conquistarsi. È un criterio da valutarsi negativamente?

Per approfondire il discorso qualche esempio può aiutare, e mi si perdonerà se a mia volta lo traggo dalla *letteratura scientifica* (ma sarà poi tale?) che mi capita di praticare, sia pure da non accademico: quella degli studi sull'editoria. Il primo esempio è proprio nel già citato annuario *Tirature*. La pubblicazione, anno dopo anno, ha l'ambizione di creare un ponte tra l'accademia e i professionisti del mondo editoriale. Il progetto sottostante – culturale (scientifico?) ancor prima che editoriale – considera un valore la capacità di parlare alla *comunità di coloro che praticano la disciplina*, anche e soprattutto a quelli che, in questa comunità, vivono fuori dall'accademia.

Ciò condiziona la forma della pubblicazione e le procedure con le quali viene alla luce. Probabilmente, non passerebbe i criteri di *qualità* dettati da molti dei documenti sulla valutazione nelle materie umanistiche elaborati negli ultimi anni. Dal che deduco – spero senza che ciò sia una forzatura – che quei criteri, applicati nel contesto editoriale italiano, considerano un disvalore l'apertura del discorso scientifico verso i non accademici.

Da *Tirature* si può trarre un secondo esempio, leggendo questa volta un articolo in essa pubblicato. Parlando degli studi sulla musica pop in Italia, U. Fiori scrive:

*Nei libri sulla canzone si possono grosso modo distinguere due categorie: 1) testi celebrativi, parapubblicitari; 2) Testi critici ... Assegnare una pubblicazione a questa o quella categoria, in Italia, non è sempre facile.*¹¹

¹⁰ Più che dalle citazioni in altre pubblicazioni rileverebbero, a quel punto, le citazioni in sentenza, specie di Cassazione, a sanzione di una autorevolezza dottrina-ria con forte impatto sulla pratica.

¹¹ U. Fiori, *Canta che ti studio*, in «*Tirature*» 2010, a cura di V. Spinazzola, Il Saggiatore, Milano 2010, pp. 95-99.

Soprattutto, spiega l'autore (senza alcun riferimento al dibattito sulla valutazione), è difficile farlo guardando semplicemente alle procedure di produzione editoriale.

Un terzo esempio riguarda la difficoltà di distinguere tra pubblicazioni scientifiche e didattiche. Ancora nel ristretto ambito degli studi sull'editoria, il recente volume di Gino Roncaglia¹² sugli *ebook* si presenta, nella forma, come un manuale. Allo stesso tempo è testo utile per i professionisti del libro, e tra questi – nella mia esperienza – molto diffuso. Eppure è un testo che ha certamente un contenuto *scientifico*. Ma come sarebbe giudicato applicando le griglie di cui si va discutendo per le materie umanistiche?

Un'ipotesi che qui potrebbe farsi è quella di considerare la diffusione editoriale come una misura di impatto nelle comunità più vaste, anche extra-accademiche. Nei documenti sulla valutazione che ho avuto modo di analizzare, la diffusione viene a volte considerata, ma solo in relazione a quella nel mondo accademico, attraverso misure come la *presenza nelle biblioteche universitarie*, in Italia o all'estero. Ci si può chiedere se la diffusione extra-accademica può essere un valore, non per suggerire che i docenti universitari debbano trasformarsi in scrittori di best-seller, ma per sottolineare che la capacità di fare trasferimento culturale può essere giudicata positivamente. In termini pratici si tratterebbe di capire se esistono dati utilizzabili e quali sistemi di ponderazione – rispetto ad altri – possono essere adottati. Ma in via preliminare si dovrebbe ragionare sul *se* la funzione di queste pubblicazioni, che non è di *divulgazione*, ma di *trasferimento culturale*, sia da considerarsi nel processo di valutazione¹³.

¹² G. Roncaglia, *La quarta rivoluzione. Sei lezioni sul futuro del libro*, Laterza, Bari 2010.

¹³ In linea generale, ritengo utile tenere distinti i problemi di cosa si vuole valutare, quali parametri possono essere di supporto alla valutazione e infine della disponibilità di dati per l'utilizzo in concreto delle variabili scelte. In questo caso vi sono tre domande diverse: (i) è giusto valutare anche la funzione di trasferimento culturale di una pubblicazione? (ii) le vendite (o le tirature) sono una buona variabile *proxy* di questa funzione? (iii) sono disponibili dati affidabili sulle vendite (o sulle tirature)?

4. LEGGERE O FAR DI CONTO?

O DELL'USO DI MODELLI QUANTITATIVI NELLA VALUTAZIONE

Il tema ci porta a un altro problema: come fare *valutazione quantitativa* nei diversi contesti disciplinari e in particolare laddove dati sull'impatto non sono presenti?

L'esercizio richiesto è dopo tutto sempre lo stesso: si tratta (a) di contare le pubblicazioni e (b) di pesarle. Contare è il primo e più elementare esercizio in ambito quantitativo, ma non privo di insidie. Perché bisogna mettersi d'accordo su cosa *conta* e va dunque *contato*.

Proviamo a guardare i pareri elaborati da CUN e ANVUR in vista dell'emanazione del regolamento ministeriale (previsto dall'art. 16 della legge 240/2010, la cd *Riforma Gelmini*) per definire i *criteri e parametri* di valutazione ai fini delle nuove procedure di abilitazione¹⁴. Il numero minimo di pubblicazioni è una variabile presente sia nel documento ANVUR, sia in tutti i 14 documenti per area disciplinare del CUN. Il punto però è cosa sia una pubblicazione scientifica.

Un criterio usato è contare solo pubblicazioni in determinate sedi editoriali, per esempio – specie nelle discipline scientifico tecnico mediche (STM) – in riviste indicizzate in banche dati internazionali quali ISI Thomson o *Scopus* Elsevier. È la scelta secca dell'ANVUR per tutte le aree STM e alcuni settori disciplinari nelle scienze sociali (economia politica, statistica, ecc.). È però interessante constatare come, lavorando nel dettaglio di ogni singola area disciplinare, il CUN suggerisca una serie di eccezioni. Persino per la fisica, area principe di applicazione delle tecniche bibliometriche, è individuato un settore disciplinare in cui si chiede di derogare al criterio. Per fare un altro esempio, nell'area 8 (Ingegneria civile e architettura) il criterio delle *riviste ISI* viene temperato contando allo stesso modo altre riviste selezionate dalla Conferenza dei presidi di

¹⁴ Cfr. CUN, *Proposta su «criteri e parametri per la valutazione» ai fini di cui all'Art. 16 comma 3 lettere a) e h) della Legge 30 dicembre 2010, n. 240*, 9 giugno 2011 e ANVUR, *Criteri e parametri di valutazione dei candidati e dei commissari dell'abilitazione scientifica nazionale*, 22 giugno 2011.

Architettura, precisando che debbono avere *elevato contenuto scientifico*; nell'area 9 un articolo deve essere uscito, più genericamente, in una rivista che pratichi la *peer review*.

Interessante è la perplessità mostrata dai matematici (area 1): forse proprio perché per loro natura più attenti a valutare il significato dei numeri, non forniscono alcuna definizione escludente e non si fidano delle risorse informative esistenti, così che stabiliscono soglie minime nel numero di *pubblicazioni riconosciute*, senza dare indicazioni su come riconoscerle.

La delicatezza metodologica di questa fase sta nella scelta di tipo 0/1 che comporta. Qualunque sia la definizione adottata, se una pubblicazione non vi rientra semplicemente conta zero nel processo di valutazione, senza possibilità di recupero nella fase successiva. Per questo è un elemento delicato del processo. Gli effetti sul mercato possono essere molto importanti. Se una sede editoriale non è considerata in questa fase, non potendo offrire *valore* a chi vi pubblica, sarà non appetibile per gli autori e tenderà a svalutarsi ulteriormente.

La scelta di fidarsi di una selezione affidata completamente a terzi (come nel caso delle riviste ISI o *Scopus*) ha il limite dell'assenza di controllo nella selezione stessa. Tale limite può essere considerato un pregio quando la ricerca di oggettività è dettata in primo luogo dal tentativo di rendere non controllabile (e quindi non truccabile) il processo di selezione, perché non ci si fida dei soggetti chiamati a gestirlo. Non ci si può nascondere che questa preoccupazione ha grande peso nel dibattito italiano, ma anche che ciò crea una distorsione di sistema: poiché non ci si fida, non si possono calibrare le tecniche di misurazione su obiettivi condivisi. Si procede allora in modo rudimentale. L'assunto sembra essere: meglio la distorsione introdotta per questa via che quella derivante dai comportamenti di valutatori disonesti.

Pressoché tutte le altre definizioni di *pubblicazione scientifica* presuppongono una valutazione della sede editoriale: la rivista o l'editore. E qui ricompaiono i problemi già discussi sopra: qual è il criterio e poi la procedura per *includere* o *escludere* una sede? La presenza di una *peer review*, si ripete molto spesso. Ma la *peer review* è

per sua natura una procedura formale, che ha senso laddove nasce da esigenze editoriali precise. Se la si impone, l'incentivo per le imprese non è quello di costruire procedure formali rigorose, ma formalità procedurali create al solo fine di *passar l'esame*. Il punto è che non ci si può esimere da un giudizio qualitativo nell'attribuzione di una qualifica di *scientificità* a una pubblicazione. Quando lo si basa su formalismi rischia di essere controproducente: le forme consolidate e di qualità sono le più difficili da modificare, perché hanno un loro senso. Sono invece le pubblicazioni più povere a potersi vestire di nuovi abiti con minor difficoltà.

Meglio allora, credo, evitare meccanismi 0/1 come quello della definizione di *scientificità* e *contare* in questa fase tutte le pubblicazioni, se si ha la capacità – successivamente – di pesarle opportunamente. Il che può includere il caso di pubblicazioni con un peso pari a zero, al termine però di un'analisi più articolata, non a priori.

Quando si passa a *pesare* le pubblicazioni, l'esercizio è quello di trasformare un giudizio qualitativo in un numero. Occorre innanzi tutto esser consapevoli che non si sta trasformando il giudizio in *oggettivo*. Ce lo spiega bene proprio uno statistico – che è ben consapevole della differenza tra mutabili (qualitative) e variabili (quantitative) ed è abituato a maneggiarle –, Vittorio Frosini, quando scrive a proposito dei valori dell'*Impact Factor*, che pure sono elaborati sulla base di gran quantità di dati:

*Il giudizio da soggettivo (fatto a seguito di lettura da parte di un referee) diventa oggettivo (il numero dato dagli if). Se a oggettivo si sostituisce formale, si descrive meglio la procedura, che non ha nulla di oggettivo.*¹⁵

Quando mancano dati sulle citazioni, sono stati proposti meccanismi di trasformazione della qualità in numeri più o meno elaborati, in genere basati su una valutazione della sede editoriale e sempre,

¹⁵ B.V. Frosini, *Valutazione della ricerca e valutazione delle riviste scientifiche in ambito statistico*, in «*Statistica & Società*» 6 (2008), n. speciale, pp. 39-47.

per loro natura, opinabili. Due esempi tipici sono rintracciabili in documenti elaborati per le aree umanistiche, in ambito CUN e da un gruppo di lavoro CNR¹⁶. Il meccanismo che si vuole costruire è il seguente:

1. si definiscono alcune caratteristiche di un modello editoriale ideale¹⁷;
2. si decidono i meccanismi con cui costruire gli indicatori di aderenza di una singola sede editoriale al modello predefinito; essenzialmente si tratta di metodi di trasformazione di un giudizio qualitativo in un numero, tramite un meccanismo di punteggi;
3. si stabiliscono i pesi relativi delle diverse caratteristiche;
4. si calcola un punteggio per ogni sede editoriale;
5. nel caso del documento CUN – in omaggio a una prassi che si sta diffondendo in numerosi nuclei di valutazione degli atenei o nelle società scientifiche di discipline che si cimentano sul tema – si crea un numero limitato di classi di qualità delle sedi editoriali (tipicamente quattro), attraverso un'analisi della distribuzione dei punteggi;
6. si definisce una scala di valori tra le classi per stabilire ciò che è cruciale: il rapporto tra i punteggi delle classi, ovvero quanto vale una pubblicazione appartenente a una classe rispetto a quella di ciascuna altra classe;
7. a ogni singola pubblicazione si attribuisce il valore della classe della sede editoriale in cui è stata pubblicata.

¹⁶ Mi riferisco rispettivamente alla già citata *Bozza finale del Gruppo di lavoro CUN sulla valutazione in area umanistica (aree 10 e 11)*, e ai *Criteri per la valutazione della ricerca nel campo delle scienze umane e sociali* elaborati da un gruppo di lavoro CNR nel dicembre 2009.

¹⁷ Ad esempio, per le riviste nel documento CUN si parla di presenza di *peer review* anonima, livello di internazionalizzazione, presenza nei più importanti repertori nazionali e internazionali, presenza in biblioteche italiane e straniere, anzianità, regolarità e continuità di pubblicazione, presenza in rete; in quello CNR di tradizione e diffusione nei rispettivi ambiti disciplinari, di puntualità d'uscita, di riconosciuta autorevolezza del direttore, del comitato scientifico, della *organizzazione scientifica* che promuove o pubblica la rivista, e di nuovo della presenza di *peer review*, e di *abstract* in due lingue.

Un processo di *quantificazione* quindi molto lungo, in sette passi, ciascuno dei quali ha le sue criticità, a partire appunto dalla definizione del modello editoriale di riferimento. In questa sorta di *bibliometria senza dati* tutto si fonda sulla capacità delle comunità scientifiche di definire il miglior modo di far gli editori. È probabilmente questa la ragione per cui talvolta gli editori vedono con fastidio questo tipo di esercizio.

D'altro canto, nell'intero processo si possono individuare ulteriori elementi critici, ciascuno dei quali meriterebbe un approfondimento. Le modalità di attribuzione dei punteggi per ogni singola caratteristica (passo 2), il peso relativo tra le caratteristiche (3), il numero di classi (5) e il peso tra le classi (6). Rendere esplicite le conseguenze nel processo di valutazione di ciascuna di queste scelte richiederebbe approfondimenti che rischiano di essere infiniti. Forse per questo la tendenza è il più delle volte quella di semplificare.

Un primo esempio è nell'ingenuità del citato documento di fonte CNR, pur frutto di un gruppo di lavoro ampio. Dopo aver definito le caratteristiche ideali di una sede di pubblicazione in termini di *tradizione* della rivista, *autorevolezza* del direttore o del comitato editoriale o della società scientifica che la promuove, il documento presuppone che per l'implementazione del modello sia sufficiente il *lavoro di spoglio e schedatura di periodici italiani del settore con attribuzione di un codice corrispondente al grado di aderenza (nessuna, parziale, totale) ai criteri stabiliti*. E che per far ciò basti *1 unità di personale per la schedatura e l'immissione dei dati*. Dunque, per trasformare il giudizio qualitativo su tradizione e autorevolezza in un valore numerico è sufficiente l'opera di un impiegato di concetto per qualche mese. L'illusione per i cui i numeri possano portare da soli a una oggettività auto-evidente raggiunge qui il suo punto più alto.

Un secondo esempio riguarda il numero delle classi di qualità delle sedi editoriali. È un elemento di grande importanza sia nel distribuire gli incentivi all'adeguamento delle sedi editoriali ai modelli proposti, sia per evitare che sedi editoriali dalle caratteristiche molto simili, ma che si trovino una al confine superiore di una classe e l'altra al confine inferiore della classe più alta, ricevano valutazioni molto distanti. Eppure il numero di classi in cui suddividere

la distribuzione è spesso fissato senza alcuna giustificazione metodologica, e quasi sempre in quattro. Perché¹⁸?

5. VALUTARE LE PUBBLICAZIONI O LE SEDI EDITORIALI

Questa analisi ci riporta a un punto cruciale dei processi di valutazione: se questo debba riguardare le sedi editoriali o direttamente le pubblicazioni. Dal punto di vista degli effetti sul mercato editoriale la distinzione è molto importante¹⁹. Come già accennato, meccanismi che valutano le sedi introducono barriere all'ingresso e alla mobilità competitiva, accentuando tendenze alla concentrazione di mercato già presenti a livello internazionale.

Nel caso estremo, un meccanismo che pone uguale a zero il valore di una pubblicazione in una sede editoriale che, essendo nuova, non può rispondere a criteri fondati sulla storia precedente (in termini di *impact factor* e sulla base di punteggi legati al suo prestigio acquisito, non importa) rende pressoché impossibile fondare una nuova rivista. Più in generale, se il valore delle pubblicazioni è dato dal *punteggio* della rivista, vi sarà un forte incentivo per gli studiosi a sottomettere i migliori lavori a queste riviste, che quindi diverranno sempre migliori, in un circolo che si autoalimenta. Nelle riviste (e in genere in tutti i periodici) un meccanismo di questo genere è già presente: indipendentemente dalle procedure di valutazione, gli

¹⁸ Poiché le classi sono poi denominate facendo ricorso a lettere e a non a numeri, da A a D, il sospetto è che tutto derivi – molto banalmente – dai sistemi di voto del sistema educativo americano. Se la discussione fosse nata in Italia si sarebbe pensato a 10 o 30 classi, a seconda che il riflesso condizionato avesse spinto la memoria verso i voti liceali o universitari.

¹⁹ Pro e contro dei due modelli dal punto di vista bibliometrici sono ampiamente esaminati in letteratura. Qui basti ricordare che la valutazione di una pubblicazione per il tramite della sua sede editoriale introduce un errore tanto più grande quanto è maggiore il rapporto tra la variabilità all'interno di ogni sede editoriale rispetto alla variabilità tra le sedi. D'altro canto, l'uso di valori riferiti alle sedi può consentire di attribuire valori quantitativi a pubblicazioni molto recenti che non hanno ancora avuto modo di produrre un impatto misurabile in termini di citazioni.

autori vogliono pubblicare nei luoghi più prestigiosi, perché sanno che saranno così più letti. La cristallizzazione di questo fattore di inerzia in meccanismi formali finisce per ingessare definitivamente il mercato.

Al contrario, un sistema che valuta direttamente le pubblicazioni crea un incentivo per le imprese a produrre riviste o monografie di maggior qualità.

Influenzato da questo punto di vista, certo molto editoriale, mi capita di sorprendermi per il fatto che strumenti appartenenti ai due sistemi di valutazione, delle sedi o delle pubblicazioni, siano trattati come equivalenti, senza giustificare la scelta verso l'uno o l'altro²⁰. Quando i dati citazionali sono disponibili, come nelle discipline STM, non è affatto la stessa cosa utilizzare l'*impact factor*, misura tipica di valutazione delle sedi, invece che il numero totale di citazioni o l'*indice h*, che invece attengono all'analisi delle pubblicazioni. In generale, per le ragioni dette, riterrei che concentrarsi sulle seconde sia altamente preferibile²¹.

²⁰ Ad esempio, nell'articolazione per area disciplinare del citato parere CUN del 9 giugno 2011, per le aree 2-4, 6 e 9 indicatori diversi sotto questo profilo, quali l'*impact factor* e l'*indice h* o il *profilo citazionale* sono messi assieme senza distinzione, nell'area 5 è privilegiato l'*impact factor*, nelle aree 7 e 8 ci si basa solo sul numero di pubblicazioni ecc.

²¹ È questa la scelta fatta da ANVUR nel citato parere del 22 giugno 2011, anche se in una delle variabili che si suggerisce di utilizzare (il numero delle pubblicazioni) si introduce per le discipline umanistiche e le scienze sociali un sistema di ponderazione così semplificato da poter fungere da esempio dei possibili effetti distorsivi delle scelte: in una prima versione si era proposto che una pubblicazione edita fuori dai confini nazionali per ciò stesso valesse tre volte quella di un editore italiano; successivamente (cfr. *Commenti, osservazioni critiche e proposte di soluzione sul documento ANVUR relativo ai criteri di abilitazione scientifica nazionale*, 25 luglio 2011) si è attribuito lo stesso peso alla lingua (italiano vs. altra lingua). L'argomentazione è più robusta ed è basata, nel linguaggio che ho qui adottato, sull'uso della lingua come *indicatore* di una maggiore diffusione internazionale. Rispetto al ragionamento qui svolto si possono osservare due cose: l'imprecisione dell'indicatore (un articolo in sloveno è più leggibile di uno in italiano?), che avrebbe a mio avviso dovuto suggerire pesi meno elevati (per non enfatizzare gli errori che inevitabilmente si commettono) o piuttosto una scelta tra lingua dell'autore e lingua che, per ogni disciplina, possa esser considerata veicolare; ma soprattutto l'introduzione, per questa via, di una gerarchia di valo-

La scelta delle prime è spesso giustificata con l'assenza di dati e con la parallela necessità di avere dei numeri. Vi è una ragione di economicità, in questo.

Attribuire un numero a un giudizio qualitativo su ogni pubblicazione è possibile solo dopo l'esame di ciascuna di esse. Non è certo qualcosa di nuovo. È lo stesso esercizio che quotidianamente si fa nei processi educativi: metter dei voti. *Io, valutatore, secondo il mio parere e sotto la mia responsabilità, stabilisco che questa pubblicazione vale un 30 e lode o un 18 tirato*, oppure, poiché usare metodi USA è sintomo di maggiore scientificità: *un doppio A o un D striminzito*. Discrezionalità del parere e relativa responsabilità, in un meccanismo che garantisca la trasparenza, sono gli elementi chiave, da sempre, in questo ambito.

Si dice tuttavia: una valutazione quantitativa delle sedi è l'unica possibile quando l'esercizio investe grandi numeri, e le eventuali imprecisioni sui grandi numeri tendono a compensarsi. L'elemento di economicità non va certamente sottovalutato, ma deve essere affrontato nel modo corretto, mettendo a confronto appunto le economie di modelli alternativi. Gli elementi del modello che sembra prevalere sono fondati, come si è visto, sul tentativo di surrogare l'assenza di dati sulle pubblicazioni con sistemi di valutazione delle sedi basati su meccanismi formali più o meno complessi. Proprio questa complessità produce costi, a meno che non si accettino semplificazioni come quelle sopra viste per cui la valutazione qualitativa delle sedi è riducibile a un meccanismo burocratico di facile implementazione.

Stupisce che non si pensi alla soluzione che parrebbe invece più logica: se il problema è l'assenza di dati, non sarebbe opportuno almeno verificare se la soluzione non possa essere proprio la costruzione di tali dati? I modelli, le tecnologie, i sistemi di rilevazione sono consolidati. I costi non sono banali, ma non credo possano essere stimati *ex ante* come maggiori di quelli impliciti nel modello alternativo. La mia impressione è che un sistema di valutazione del-

re che – in assenza di misure di diffusione – guarda solo alla funzione intraccademica delle pubblicazioni e non a quella di trasferimento culturale.

le sedi sulla base di parametri, punteggi, sistemi di accreditamento, o qualsivoglia alternativa ha costi molto elevati, ma spesso poco visibili, perché attengono all'inevitabile coinvolgimento dell'intera comunità scientifica, disciplina per disciplina, come già sta avvenendo. La preferenza per questi sistemi, *per ragioni di costo*, deriva a mio avviso proprio dalla sottovalutazione di questi costi nascosti.

In ogni caso, se la preferenza verso un sistema è dettata da ragioni di economicità non può essere proposta senza un confronto dettagliato tra costi e benefici di modelli alternativi, e non mi risulta che sia mai stato fatto un tentativo di stimare i costi né di un metodo né dell'altro. Di più, sembra che la risposta più banale al problema della mancanza di dati (stimiamo quanto costerebbe produrli) sia completamente assente dal dibattito.

O quasi. Ne trovo traccia, indirettamente, nel già citato parere CUN sui criteri e parametri per la valutazione relativo all'area 1 (Matematica). Dopo aver espresso scetticismo sulla qualità dei dati esistenti, i rappresentanti dei matematici italiani propongono (sia pure per un solo settore disciplinare) di utilizzare l'indice h , ma non ricavandolo dalle banche dati esistenti, bensì chiedendo ai candidati di autocertificare le citazioni che servono al suo calcolo. Un metodo (non so se il più efficiente)²² per creare nuovi dati, appunto.

²² Tra le altre cose, richiede obbligatoriamente che il valore soglia dell'*indice h* sia prefissato senza riferimento alla distribuzione degli stessi indici nella comunità di riferimento, come invece ad esempio proposto da ANVUR, che richiede che ciascun indice dev'essere *superiore alla mediana dello specifico Settore Concorsuale e della fascia (associati o ordinari) per cui si chiede l'abilitazione*. Il caso non è qui citato come esempio da seguire, ma solo perché è l'unico che conosca che in qualche modo segue la logica: *se non ci sono i dati, creiamoli*.