

Maria Teresa Biagetti

*Dagli indicatori bibliometrici all'analisi qualitativa:
la valorizzazione del peer-review nella valutazione della
ricerca scientifica*

Sebbene i criteri quantitativi e gli indicatori bibliometrici continuino ad essere ampiamente utilizzati nella valutazione *ex post* dei lavori scientifici e per la distribuzione dei finanziamenti nei settori tecnologico e bio-medico, e permanga l'attrazione per il loro impiego anche nel campo delle scienze umane e sociali, negli ultimi anni in diversi ambiti si è avvertito il desiderio di superare l'uso degli indicatori bibliometrici e di recuperare l'attività di valutazione tramite peer-review, pur nella consapevolezza dei limiti e delle difficoltà che questa metodologia inevitabilmente comporta.

Dal marzo al novembre 2021 il Directorate-General for Research and Innovation della Commissione europea, consapevole dell'importanza di una equilibrata valutazione dell'attività dei ricercatori, e del danno che gli studiosi appartenenti ai settori ad alto impatto sociale ma scarsamente rappresentati in caso di valutazioni basate su indicatori metrici potrebbero ricevere da un'attività valutativa incentrata sull'uso indiscriminato degli indicatori bibliometrici, come il computo

delle citazioni e l'Impact Factor delle riviste, ha avviato un'indagine sui criteri e sulle metodologie di valutazione della qualità dei lavori di ricerca più appropriati, ha consultato diverse associazioni internazionali che si occupano di ricerca, accademie, società scientifiche e università, e nel novembre del 2021 ha pubblicato un rapporto conclusivo.¹

Alla fine dei lavori è stato proposto un *European agreement*, che potrebbe essere condiviso tra le istituzioni europee che distribuiscono finanziamenti, allo scopo di modificare l'attuale sistema di valutazione.² L'obiettivo è valutare i ricercatori e i progetti di ricerca seguendo la metodologia qualitativa, quindi sulla base dei loro meriti più che sulla base del computo del numero dei lavori pubblicati e delle sedi editoriali presso le quali sono stati pubblicati, abbandonando risolutamente, in quanto inappropriato, l'impiego della misura del fattore d'impatto attribuito alle riviste scientifiche per calibrare il valore dei lavori. Si propone di adottare il tradizionale metodo del peer-review, quindi un giudizio qualitativo, arricchito con l'eventuale supporto dell'uso responsabile degli indicatori quantitativi.

Peer-review è il sistema da ritenersi imprescindibile, secondo il rapporto della Comunità europea, per la valutazione, in quanto in primo luogo permette di considerare tutte le attività espletate da un ricercatore e non limitarsi a computare solo quanti articoli o libri un autore ha pubblicato, e addirittura in quanto tempo, in secondo luogo consente di considerare anche l'attività di *mentoring*, il ruolo di leader in un settore disciplinare, il contributo a ricerche multi-, inter- o trans-disciplinari e permette di considerare e soppesare le diversità dei percorsi di ricerca intrapresi dagli studiosi e di non pretendere da tutti i ricercatori il raggiungimento di tutte le competenze ritenute eccellenti, anzi, valutare positivamente i percorsi che non hanno seguito precisamente i binari delle discipline tradizionali e i ricercatori che si sono inseriti in percorsi multi-disciplinari o trans-disciplinari, aprendo piste di ricerca o rivalutando con nuova linfa settori ritenuti

¹ European Commission 2021.

² Pubblicato il 20 luglio 2022. Si veda *Agreement 2022*.

sorpassati o poco attraenti agli occhi degli scienziati inseriti nel panorama attuale.

Il riconoscimento della funzione positiva ricoperta dalle “diversità”, sia nel settore della produzione scientifica, sia nei percorsi di ricerca seguiti dagli studiosi, sia nel campo delle metodologie di valutazione, è un tema che sta assumendo rilievo in settori anche apparentemente lontani tra di loro.

Un esempio significativo è costituito dalla posizione assunta recentemente dalla Confederation of Open Access Repositories (COAR),³ un'associazione fondata nel 2009 che coordina la comunità dei gestori di *repositories* ad accesso aperto. Nell'aprile 2016 COAR ha istituito un gruppo di lavoro dedicato a “Next Generation Repositories”⁴ allo scopo di studiare nuove tecnologie in grado di sviluppare depositi come infrastrutture per la comunicazione scientifica.

L'ultimo rapporto disponibile di COAR, Annual Report 2020-2021,⁵ relativo all'attività tra il maggio del 2020 e l'aprile del 2021, contiene l'annuncio dei progetti e degli ambiti di lavoro sui quali s'incentrerà l'attività futura di COAR. Uno dei settori di riflessione riguarda la promozione della “bibliodiversità” nella comunicazione scientifica, cioè la decisione di sostenere la diversità nella comunicazione scientifica: diversità nelle tipologie di pubblicazioni, pluralismo linguistico, pluralismo epistemologico delle diverse comunità scientifiche, ma anche la diversità dei meccanismi di richiesta di finanziamenti, diversità delle piattaforme di scambio e differenze tra i metodi di valutazione dei progetti. La diversità riduce il rischio della concentrazione dei servizi nelle mani di un unico fornitore, quindi riduce il rischio del monopolio e dell'incremento dei costi. La bibliodiversità dunque consiste nel riconoscere la necessità di mantenere attivi diversi canali di produzione scientifica, mantenere diversi formati di

³ <<https://www.coar-repositories.org/>>.

⁴ Si veda il rapporto Next Generation Repositories 2017.

⁵ <https://www.coar-repositories.org/files/COAR-Annual-Report-2020-2021_public.pdf>.

edizione, rifiutare l'omogeneizzazione dei prodotti provocata dalla presenza di pochi editori internazionali. Sostenere la diversità nella comunicazione scientifica, quindi la bibliodiversità, richiede un buon coordinamento per evitare la frammentazione e l'isolamento dei diversi settori scientifici. Sostegno della bibliodiversità sono considerate le strategie open access e la creazione di diversi servizi e infrastrutture open destinati alle diverse comunità.⁶

Nel campo della valutazione dei lavori scientifici frutto di ricerche, il riconoscimento del valore che assumono le "diversità" dei percorsi, l'apprezzamento delle personalità originali e dei lavori che si distaccano dai binari consueti, possono verificarsi soltanto attraverso analisi qualitative seriamente condotte e non utilizzando banali computi quantitativi.

Le problematiche legate alla valutazione che si avvale degli indicatori bibliometrici sono ampiamente note e su di esse è stato scritto molto anche in Italia. E' sufficiente ricordare alcune procedure valutative:

- Il computo semplice delle citazioni ricevute;
- Il computo del numero di pubblicazioni su riviste ad alto Impact Factor;
- Le misure dell'impatto scientifico e dell'impatto sociale;
- Le misure d'impatto normalizzate rispetto ai diversi campi scientifici;
- H Index di Hirsch e le misure derivate.

Come ha chiarito Alberto Baccini,⁷ a mio avviso lo studioso che ha analizzato questi problemi con più acume, il computo del numero di pubblicazioni su riviste ad alto Impact Factor è una misura valutativa che produce distorsioni inaccettabili in quanto anche articoli che hanno ricevuto uno scarso numero di citazioni, grazie alla ponderazione con l'Impact Factor della rivista sulla quale sono stati pubblicati, acquistano una valutazione più benevola di quella che riceverebbero

⁶ Si possono avere raggugli in Shearer, Chan, Kuchma, Mounier 2020.

⁷ Baccini 2010.

se valutati singolarmente. L'applicazione dell'H Index è fortemente criticata sia da Baccini che, ad esempio, dal sociologo e storico della scienza canadese Yves Gingras,⁸ come misura inadatta a rendere ragione del valore e della qualità dei lavori scientifici. Henk Moed, uno degli autori che si sono occupati con più attenzione delle distorsioni che l'uso degli indicatori bibliometrici, e in particolare degli indici citazionali, possono provocare, ha segnalato diversi anni fa l'inadeguatezza di questi indici nella valutazione dei prodotti frutto di ricerca scientifica e ne ha suggerito l'impiego solo come indici che possono fornire dati all'attività di valutazione qualitativa.⁹

Ma più che sulle posizioni dei singoli studiosi che nel tempo si sono occupati del problema, che possono avere espresso opinioni personali, è importante richiamare l'attenzione sulle critiche all'uso di indicatori bibliometrici che negli anni sono state rivolte in contesti internazionali da gruppi di persone che si occupano istituzionalmente del problema oppure hanno condiviso posizioni che successivamente sono state rese ufficiali.

Il rapporto della Commissione europea dichiara di riferirsi idealmente ad alcuni principi espressi in documenti emanati in precedenza da istituzioni come All European Academies,¹⁰ e alle raccomandazioni rivolte alla comunità scientifica contenute nella *San Francisco Declaration on Research Assessment (DORA)*,¹¹ nel *Leiden Manifesto for*

⁸ Gingras 2014 e 2014a.

⁹ L'argomento è ampiamente affrontato in Moed 2005, in particolare alle p. 229-234, e ripreso in Moed 2007. Sull'impiego degli indici citazionali, la letteratura internazionale è molto vasta e ne ho dato conto in altri miei lavori, ai quali rinvio. Importante ricordare almeno l'interpretazione che dell'uso della citazione bibliografica fornisce Gingras (2008, 2014a), il quale la considera piuttosto un artificio retorico degli autori per sostenere le proprie posizioni riferendosi ad altri studiosi, schermendosi dietro la loro autorevolezza, al fine di persuadere i lettori della scientificità delle loro ricerche.

¹⁰ ALLEA 2017.

¹¹ <<https://sfdora.org/read/>>.

research metrics,¹² e nel Manifesto di Hong Kong.¹³ È particolarmente utile, quindi, presentare un breve approfondimento su questi documenti ufficiali.

San Francisco Declaration on Research Assessment (DORA) è un documento pubblicato nel 2013 e redatto l'anno precedente da un gruppo di curatori di riviste scientifiche. La Società Americana di Biologia Cellulare (American Society for Cell Biology), durante il convegno tenuto a San Francisco il 16 dicembre 2012, aveva aperto il dibattito sulla necessità di migliorare le modalità di valutazione dei lavori scientifici per la pubblicazione da parte delle riviste, da parte delle istituzioni accademiche e da parte degli enti che finanziano le ricerche. Deriva proprio da quel dibattito la definizione di una serie di raccomandazioni sulla valutazione della ricerca.

Nel documento si assume una posizione molto critica soprattutto sull'impiego della misura dell'Impact Factor delle riviste allo scopo di rendere agile la valutazione della produzione dei ricercatori. Si sottolinea infatti la necessità di preferire il metodo del peer-review all'uso del Journal Impact Factor, anzi, si afferma risolutamente che è necessario eliminare l'uso delle metriche relative alle riviste scientifiche, come appunto il Journal Impact Factor, per decidere a chi distribuire i fondi per la ricerca, o come criterio per le assunzioni di ricercatori e gli avanzamenti in carriera.

Uno dei punti più significativi riguarda il suggerimento di valutare i singoli articoli pubblicati, senza tenere conto della rivista su cui sono pubblicati, focalizzandosi quindi sul valore dell'articolo in sé; un secondo punto riguarda le agenzie che finanziano le ricerche, le quali devono esplicitare i criteri adottati per valutare la produzione scientifica dei richiedenti; un terzo elemento significativo ci riconduce alla valutazione dell'impatto della produzione scientifica, che dovrebbe essere valutato preferibilmente adottando criteri qualitativi, ad esempio considerando l'influenza esercitata sulle politiche e sulle prassi

¹² The Leiden Manifesto 2015.

¹³ Si veda Moher 2019 e 2020.

adottate nel mondo culturale e industriale.

Bisogna inoltre ricordare che molto recentemente, nel 2021, DORA ha lanciato il progetto TARA: *Tools to Advance Research Assessment*,¹⁴ che ha l'obiettivo di rendere più facile la definizione di nuove regole e pratiche innovative per la valutazione delle carriere accademiche, sviluppando strumenti interattivi che permettono di far conoscere i criteri e gli standard usati dalle Università del mondo per prendere decisioni in merito alle promozioni di carriera, strumenti per promuovere il miglioramento delle pratiche e in ultima analisi sottoporre a riforma la procedura della valutazione della ricerca.

Molto più noto, anche in Italia, *Bibliometrics: The Leiden Manifesto for research metrics*.¹⁵ Il documento, firmato da Diana Hicks e da altri bibliometricisti, è rivolto al mondo scientifico e mette in evidenza le distorsioni che l'abuso delle metriche nella valutazione della ricerca può provocare. Si propone anche come sintetico vademecum di buone pratiche per l'attività di valutazione dei lavori degli studiosi, pratiche che contemplano anche l'uso di indicatori di tipo bibliometrico, ma solo come aiuto alla valutazione dei pari.

Si dichiarano 10 principi da seguire: il primo, fondamentale, invita appunto ad usare le metriche solo come supporto alla valutazione qualitativa, basata sul giudizio di esperti, riconoscendo in modo implicito il valore primario di questo tipo di valutazione, qualitativa; il terzo suggerisce ai valutatori di salvaguardare la ricerca di eccellente qualità prodotta in ambito locale, che sfugge all'indicizzazione degli istituti internazionali, e dunque invita a valutarla esclusivamente con metodi qualitativi.

Ancora nel 2015, è rilevante il rapporto firmato da James Wilsdon, sociologo della scienza e della ricerca all'Università di Sheffield, con diversi collaboratori, *The metric tide*.¹⁶ Il rapporto si focalizza sull'analisi del ruolo delle metriche nella valutazione della ricerca e ne evi-

¹⁴ <<https://sfdora.org/project-tara/>>.

¹⁵ The Leiden Manifesto 2015.

¹⁶ The Metric Tide 2015.

denza le distorsioni e gli effetti inattesi, pur nella consapevolezza degli interessi che contribuiscono a far sollevare la “marea metrica”, come, ad esempio, il controllo della spesa pubblica, che inviterebbe a dare la preferenza a procedure che non impegnano né tempo né denaro. La revisione tra pari, nonostante i suoi difetti e limiti, continua ad essere indicata come il sistema più solido e l'uso delle metriche dovrebbe – si afferma nel rapporto – essere limitato al supporto dell'attività di giudizio degli esperti. Si afferma la centralità del peer-review affiancato dall'uso responsabile degli indicatori bibliometrici, l'*informed peer review*,¹⁷ e si caldeggia la riduzione dell'enfasi sull'Impact Factor delle riviste da parte degli editori stessi. Gli editori ricoprono infatti un ruolo importante e possono influenzare la percezione degli stessi lettori e utenti, ma anche degli attori principali in questo contesto, gli stessi studiosi.

Dell'attività di peer-review, nel rapporto firmato da Wilsdon, vengono analizzati gli elementi positivi e le debolezze. Tra queste ultime è particolarmente rilevante, come ampiamente noto, la soggettività dei giudizi, la distorsione pre-concetta (*bias*) e l'influsso della società o dell'accademia stessa.

La sociologa e storica della scienza Michèle Lamont¹⁸ è l'autrice che meglio di altri ha rappresentato le difficoltà dell'attività di peer-review nell'ambito delle Scienze sociali ed umane (SSH). Nel suo volume, pubblicato nel 2009, quindi in un contesto temporale precoce rispetto alla assunzione di consapevolezza recente da parte dei bibliometrici, Lamont ha messo in evidenza come in questo ambito gli studiosi adot-

¹⁷ Nella letteratura internazionale sul tema, anche l'*informed peer-review* ha ricevuto critiche. Paul Wouters, ad esempio, già nel 2014 rilevava la problematicità nella scelta degli indicatori bibliometrici da affiancare alla valutazione tra pari e la sostanziale discrezionalità delle decisioni di volta in volta assunte (si veda Wouters 2014). Alberto Baccini e Giuseppe De Nicolao hanno sottolineato la difficoltà di adottare indicatori bibliometrici per i settori disciplinari STEM e *informed peer-review* per le SSH. Si veda Baccini, De Nicolao 2016.

¹⁸ Lamont 2009.

tino criteri diversi per stabilire la “qualità” dei lavori scientifici anche all’interno della stessa disciplina e come i loro giudizi siano influenzati dall’appartenenza ad una particolare corrente all’interno del proprio settore, e anche dalle interazioni personali tra gli studiosi, molto più di quanto non avvenga nei settori delle scienze esatte e delle tecnologie. Si evidenzia la mancanza di consenso sui criteri da adottare nella valutazione nelle SSH. Inoltre, difficilmente gli studiosi che hanno il compito di valutare trovano accordo sul criterio della “qualità” all’interno di una disciplina. Il loro comportamento è piuttosto influenzato dai rapporti interpersonali ed è guidato dalle emozioni, dalle interazioni tra gli individui, e spesso anche dal desiderio di vedere apprezzate le proprie opinioni da parte dei colleghi di altre discipline. Infine, i valutatori possono essere suggestionati dalla comunità scientifica o sociale e possono essere condizionati dalla reputazione dell’autore da valutare, se non viene utilizzata la modalità in doppio-cieco.

D’altro canto, il computo delle citazioni ricevute e la valutazione fondata sulla verifica dell’Impact Factor delle riviste si prestano a manipolazioni. Come hanno sottolineato Mario Biagioli e Alexandra Lippman,¹⁹ la condotta fraudolenta e la manipolazione o la distorsione dell’attività di valutazione vengono spesso messe in atto durante il processo di pubblicazione dei lavori, e non più, o non soltanto, durante la realizzazione dei lavori stessi. Le strategie comprendono, come evidenziato in alcuni articoli del volume curato da Biagioli e Lippman, ad esempio la presentazione di false revisioni o l’inserimento di nomi di altri autori durante il processo di stampa, in modo da sostenere scientificamente il lavoro, oppure, e in maggior misura, la manipolazione dell’Impact Factor delle riviste e il numero di citazioni ricevute dagli articoli. Sono nuove modalità di “misconduct”, che si differenziano dalle attività fraudolente del passato, che venivano perpetrate durante la scrittura, come la falsificazione e il plagio. La modalità con cui le riviste online si presentano al lettore, inoltre, offrendo in primo

¹⁹ Biagioli - Lippman 2020, articolo introduttivo della monografia edita a loro cura, *Gaming the Metrics* 2020.

piano e in bella evidenza l'Impact Factor della rivista, comporta una implicita valutazione dell'articolo che il lettore ha scelto, e lo induce a ritenere che l'articolo sia di alto livello, se l'IF della rivista è alto. Si realizza così una manipolazione dell'impatto che in futuro l'articolo avrà, e si suggerisce al lettore una implicita valutazione. Come gli autori rilevano, si tratta spesso di attività illegali o comunque non sostenibili dal punto di vista etico.

In anni più recenti, l'uso dell'Impact Factor delle riviste come indicatore di qualità scientifica viene di nuovo messo in discussione con vigore. Paul Wouters, docente di *Scientometrics* presso la Leiden University e firmatario del Leiden Manifesto del 2015, nel 2019 ritorna ad insistere sulle criticità dell'uso del fattore d'impatto delle riviste nella valutazione della ricerca. In un articolo pubblicato su «Nature»²⁰ suggerisce di abolire o almeno di limitare molto l'uso dell'Impact Factor delle riviste come criterio preferenziale per la valutazione degli articoli degli studiosi (nella valutazione *ex post*) e di utilizzare invece un insieme ampio di indicatori riferibili piuttosto alle diverse funzioni ricoperte dalle riviste scientifiche. Invita, ad esempio, a considerare se le riviste su cui gli studiosi hanno pubblicato abbiano utilizzato l'*expertise* dei comitati editoriali, se abbiano o meno dichiarato in modo trasparente i criteri di accettazione e se abbiano permesso di visionare il processo valutativo dell'accettazione. Si propone quindi che il Journal Impact Factor abbia un ruolo limitato nella valutazione del lavoro degli studiosi.

Il documento più significativo, che ha avuto una importante risonanza a livello internazionale, a mio avviso è però il Manifesto di Hong Kong. Nel 2019, durante la 6th World Conference on Research Integrity che si tenne a Hong Kong dal 2 al 5 giugno, venne approvato *The Hong Kong Manifesto for Assessing Researchers: Fostering Research Integrity*.²¹

²⁰ Si veda Paul Wouters 2019.

²¹ Il Manifesto di Hong Kong si può leggere in Moher 2019 e in Moher 2020. David Moher è un epidemiologo canadese, di origini irlandesi.

I principi approvati come parte dei lavori della conferenza, alla quale hanno partecipato in grande maggioranza esponenti del settore bio-medico, mirano ad incentivare le buone pratiche valutative che riconoscono e apprezzano il lavoro di quei ricercatori che mantengono un comportamento scientifico degno di fiducia ed hanno evitato, durante la loro attività, le pratiche di ricerca che sono considerate discutibili.

La rilevanza del Manifesto consiste nel fatto che mira a coniugare la valutazione della ricerca con l'integrità della ricerca. La pratica valutativa apprezzata e incoraggiata è quella che tende a riconoscere e a premiare con l'avanzamento nella carriera quei ricercatori che hanno mostrato un comportamento che è orientato a sostenere e a perseguire l'integrità della ricerca. Raramente, si fa notare, i criteri valutativi includono proprio questo: valutare l'affidabilità della ricerca e la trasparenza e il rigore in ogni stadio dei progetti di ricerca e della loro realizzazione.

I principi sui quali è stato raggiunto un accordo possono essere qui velocemente elencati:

- assess responsible research practices;
- value complete reporting;
- reward the practice of open science;
- acknowledge a broad range of research activities;
- recognise essential other tasks like peer review and mentoring.

Come orientamento generale, si accettano i punti su cui si è espressa DORA, in particolare l'indicazione di limitare l'uso delle metriche nella valutazione, ma si insiste particolarmente su di un valore imprescindibile, sul fatto cioè che la ricerca e i risultati ottenuti come esito siano "trustworthy, robust, rigorous, and complete".

Il primo principio invita a valutare e valorizzare le pratiche di ricerca responsabile. Non è indicativo della bontà della ricerca e del suo rigore né il numero delle pubblicazioni di un ricercatore, né il numero delle citazioni ricevute, influenzate, come è ampiamente noto, dalle diverse pratiche citazionali e dall'uso retorico della citazione, o,

addirittura, l'entità dei finanziamenti ricevuti. È preferibile adottare, invece, altri criteri, ad esempio quanto un ricercatore sia stato capace di coinvolgere nelle sue ricerche membri esterni, ad esempio nel caso di studi bio-medici. Nella valutazione della carriera di un ricercatore si invita a valutare anche se egli abbia fornito i dati delle proprie ricerche, li abbia messi in condivisione dando la possibilità ad altri di verificarli e riutilizzarli. Questa attività richiede tempo ai ricercatori, ma è una buona pratica per la ricerca responsabile, e il tempo impiegato per raggiungere questo obiettivo deve essere considerato come tempo speso per la ricerca.

Il secondo principio si rivolge in particolare al settore bio-medico, nel quale spesso accade che ricevano una pubblicazione solo le ricerche che hanno prodotto i risultati aspettati, positivi, e le ricerche che invece hanno messo in evidenza risultati non positivi non vengano edite. Si suggerisce quindi, come buona pratica, di valutare l'intera produzione dei ricercatori, e di considerare tutti i loro lavori, non solo quelli che hanno ricevuto la dignità di pubblicazione.

Delle pratiche dell'accesso aperto e del raggiungimento degli obiettivi della *open science* – ampiamente note – si occupa il terzo principio, mentre il successivo invita a valutare un insieme di attività scientifiche e di ricerca più ampio possibile, ad esempio l'attività di replicazione dei risultati delle ricerche di altri ricercatori, la traduzione o la sintesi di lavori altrui, la validazione di nuovi strumenti, in particolare nel campo bio-medico o della Fisica. L'invito è a riconoscere un più ampio spettro di prodotti di ricerca, dando dignità ad attività tradizionalmente ritenute minori e non degne di essere considerate. Tra queste, col principio successivo, si suggerisce di prendere in considerazione anche l'attività di *reviewer* per le riviste e il contributo allo sviluppo delle carriere dei giovani ricercatori, cioè l'attività di *mentoring*, che richiede la revisione dei lavori dei giovani collaboratori e la costante supervisione delle loro ricerche.

Il Manifesto si differenzia da DORA perché si focalizza sul concetto di “integrità della ricerca” che non significa solo evitare le cattive

condotte, evitare di falsificare, evitare il plagiarismo, ma si incentra invece sulle buone pratiche e la responsabilità nella ricerca e sul fatto che questo e la metodologia della *open science*, che permettono la trasparenza, il controllo e il riuso, debbano essere premiate. Bisogna che siano valutate queste pratiche perché richiedono tempo ai ricercatori, che quindi meritano di essere premiati per questo.

In conclusione, il Manifesto invita a considerare, nella valutazione delle carriere dei ricercatori, il comportamento “responsabile” nell’attività di ricerca e l’aspirazione all’integrità della ricerca. Per mettere in pratica questi comportamenti virtuosi i ricercatori impiegano parte del loro tempo, e rischiano probabilmente di avere un numero di pubblicazioni minore. È quindi opportuno riconoscere i comportamenti che si ispirano all’ideale dell’integrità della ricerca e attribuire ad essi un riconoscimento sensibile.

La recente assunzione di una posizione responsabile e critica nei confronti dell’uso di indicatori bibliometrici scarsamente affidabili, come l’Impact Factor delle riviste, e la riconsiderazione favorevole dell’impiego della valutazione qualitativa dei pari, inducono a considerare una serie di elementi che convergono nella medesima direzione. Possiamo affermare con una certa sicurezza che le proposte avanzate da Wouters e quelle contenute nel Manifesto di Hong Kong indirettamente si appellano alla necessità di adottare buone pratiche e linee guida per la valutazione che siano fondate in qualche modo su di una visione della valutazione della ricerca che si basi su principi etici.

Il tema dell’impostazione etica della valutazione della ricerca è molto poco conosciuto e poco studiato, tuttavia recentemente ha ricevuto attenzione nell’ambito della Cost Action ENRESSH.²² Un progetto da me proposto e coordinato nell’ambito di ENRESSH nel corso del 2019 ha avuto come obiettivo l’individuazione di un’area specifica

²² European Network for Research Evaluation in the Social Sciences and Humanities (ENRESSH) <<https://enressh.eu/>> è stata una COST Action all’interno di Horizon 2020, alla quale hanno partecipato 125 ricercatori appartenenti a 37 nazioni.

dedicata all'Etica della valutazione della ricerca. Lo spazio euristico può emergere dalla sovrapposizione tra le problematiche specifiche dell'Etica della valutazione considerata nella sua funzione generale, dedicata in gran parte alla valutazione dei progetti internazionali di sostegno e sviluppo sociale, e che può contare su molte linee guida internazionali, e dell'Etica e integrità della ricerca, settore più specifico, per la quale sono disponibili linee guida europee.

Nella letteratura internazionale che si occupa della valutazione della ricerca una porzione significativa è costituita da riflessioni sull'attività di valutazione dei progetti internazionali di sostegno e sviluppo sociale, che possono presentare dilemmi di natura etica che riguardano soprattutto il comportamento professionale dei ricercatori.²³

Sulla base anche dei principi proposti da Karen Kitchner nell'ambito del *counseling*,²⁴ Michael Morris individua 5 principi etici di ordine generale, sui quali tutti possono essere d'accordo, per la valutazione:

- Systematic Inquiry: i valutatori devono realizzare un'analisi scrupolosa, basata sui dati, e devono comunicare i criteri che hanno adottato.

- Competence: devono assicurare di fornire un'analisi competente, cioè devono possedere competenze e abilità per valutare i casi specifici.

- Integrity/Honesty: assicurare un comportamento rivolto all'integrità e all'onestà durante tutta la valutazione. Devono essere dichiarati anche i conflitti d'interesse.

- Respect for People: rispettare la sicurezza personale, la dignità di chi partecipa ad un programma o è interessato alla valutazione; rispettare le condizioni economiche o politiche e le differenze di religione.

- Responsibilities for General and Public Welfare: mostrare responsabilità nei confronti del benessere pubblico e tenere in considerazione tutti i diversi interessi e i valori collegati ad una valutazione.

²³ Michael Morris, professore di Psicologia, è l'autore di riferimento per questo settore. Si veda *Evaluation ethics* 2008.

²⁴ Kitchener 1984.

Nell'ambito del settore che invece si occupa dell'etica e dell'integrità della ricerca, è stato possibile trovare spunti nelle linee guida che si dedicano a tratteggiare il comportamento ideale dei ricercatori e a definire le loro responsabilità. All'European Academies (ALLEA 2017) e Economic and Social Research Council (ESRC 2015, 2022) mettono in evidenza i seguenti punti nel descrivere le responsabilità professionali ed etiche del ricercatore:

- Reliability, affidabilità e assicurazione della qualità della ricerca, della metodologia adottata e dell'uso delle risorse.
- Honesty, onestà e sincerità nella presentazione trasparente e non condizionata da visioni parziali, prevenute e distorte dei risultati.
- Respect, rispetto a tutti i livelli, per i colleghi, per i partecipanti alla ricerca, per l'ambiente e il patrimonio culturale.
- Accountability, responsabilità, con obbligo di rispondere e dare conto, della gestione della ricerca e della gestione dei fondi, nonché del mentoring effettuato.

Dall'intersezione dei criteri stabiliti negli ambiti dell'etica della valutazione in generale e dell'etica e della integrità della ricerca, e tenendo conto degli elementi messi recentemente in evidenza nel Manifesto di Hong Kong, possono emergere i contenuti specifici del settore dell'etica nella valutazione della ricerca, e in particolare l'obiettività nella revisione tra pari, la necessità di evitare le distorsioni disciplinari (*bias*) e la consapevolezza del fatto che accusare i ricercatori dei quali si sta valutando la carriera di "misconduct" senza fornirne le prove, oppure adottare comportamenti sostenuti dalla "malevolence" o rappresentare in modo volutamente errato i risultati ottenuti da chi è sottoposto a valutazione, sono pratiche inaccettabili e riprovevoli anche dal punto di vista etico.²⁵

²⁵ Rinvio ai due lavori realizzati su questo argomento: Biagetti, Gedutis, Ma 2020 e Gedutis, Biagetti, Ma 2022.

Bibliografia

- Agreement 2022 = *Agreement on reforming research assessment*, <<https://www.scienceeurope.org/our-resources/agreement-reforming-research-assessment/>>
- ALLEA 2017 = All European Academies, *The European Code of Conduct for Research Integrity*. Revised edition. Berlin, ALLEA.
- Baccini 2010 = Alberto Baccini, *Valutare la ricerca scientifica: uso e abuso degli indicatori bibliometrici*, Bologna, Il Mulino.
- Baccini - De Nicolao 2016 = Alberto Baccini - Giuseppe De Nicolao, *Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise*, «Scientometrics», 108, 3, p. 1651–1671. DOI 10.1007/s11192-016-1929-y.
- Biagetti - Gedutis - Ma 2020 = Maria Teresa Biagetti - Aldis Gedutis - Lai Ma, *Ethical Theories in Research Evaluation: An Exploratory Approach*, «Scholarly Assessment Reports», 2 (1) 11.
- Biagioli - Lippman 2020 = Mario Biagioli - Alexandra Lippman, *Introduction: Metrics and the New Ecologies of Academic Misconduct in Gaming the Metrics: Misconduct and Manipulation in Academic Research*, edited by Mario Biagioli and Alexandra Lippman, The MIT Press, 2020 DOI: <<https://doi.org/10.7551/mitpress/11087.001.0001>>.
- ESRC 2015, 2022 = Economic and Social Research Council, *Framework for Research Ethics* <<https://www.ukri.org/councils/esrc/guidance-for-applicants/research-ethics-guidance/framework-for-research-ethics/>>.
- European Commission 2021 = European Commission, Directorate-General for Research and Innovation, *Towards a reform of the research assessment system : scoping report*, Publications Office <<https://data.europa.eu/doi/10.2777/707440>>.
- Evaluation ethics 2008 = *Evaluation ethics for best practice. Cases and commentaries*, edited by Michael Morris, New York, The Guilford Press.
- Gaming the Metrics 2020 = *Gaming the Metrics: Misconduct and Manipulation*

- ulation in Academic Research*, edited by Mario Biagioli and Alexandra Lippman, The MIT Press, 2020 DOI: <<https://doi.org/10.7551/mitpress/11087.001.0001>>.
- Gedutis - Biagetti - Ma 2022 = Aldis Gedutis - Maria Teresa Biagetti - Lai Ma, *The challenges for research evaluation ethics in the social sciences* in *Handbook on research assessment in the social sciences*, edited by Tim C.E. Engels and Emanuel Kulczycki, Cheltenham (UK), Edward Elgar Publishing, 2022, p. 367-385.
- Gingras 2008 = Yves Gingras, *Du mauvais usage de faux indicateurs*, «Revue d'histoire moderne et contemporaine», 55-4 bis, 5, p. 67-79 DOI:<[10.3917/rhmc.555.0067](https://doi.org/10.3917/rhmc.555.0067)>.
- Gingras 2014 = Yves Gingras, *Criteria for evaluating indicators*, in *Beyond Bibliometrics: harnessing multidimensional indicators of scholarly impact*, edited by Blaise Cronin and Cassidy R. Sugimoto, Cambridge (Mass.) - London, The MIT Press, p. 109-125.
- Gingras 2014a = Yves Gingras, *Les dérives de l'évaluation de la recherche. Du bon usage de la bibliométrie*, Paris, Éditions Raisons d'agir.
- Kitchener 1984 - Karen Kitchener, *Intuition, Critical Evaluation and Ethical Principles: The Foundation for Ethical Decisions in Counseling Psychology*, «The Counseling Psychologist», Vol. 12, 3, p. 43-55.
- Lamont 2009 = Michèle Lamont, *How professors think. Inside the curious world of academic judgement*, Cambridge-London, Harvard University Press.
- The Leiden Manifesto 2015 = *Bibliometrics: The Leiden Manifesto for research metrics. Use these ten principles to guide research evaluation*, urge Diana Hicks, Paul Wouters and colleagues, «Nature», 520, 7548, p. 429-431.
- The Metric Tide 2015 = Higher Education Funding Council for England, *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management* DOI: [10.13140/RG.2.1.4929.1363](https://doi.org/10.13140/RG.2.1.4929.1363).
- Moed 2005 = Henk F. Moed, *Citation analysis in research evaluation*, Dordrecht, Springer.

- Moed 2007 = Henk F. Moed, *The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer-review*, «Science and Public Policy», 34, 8, p. 575-583.
- Moher 2019 = David Moher, Lex Bouter, Sabine Kleinert, Paul Glasziou, Mai Har Sham, Virginia Barbour, Anne-Marie Coria, Nicole Foeger, Ulrich Dirnagl, *The Hong Kong Principles for assessing researchers: Fostering research integrity*, «OSF Preprints», created: September 17, 2019 | Last edited: June 28, 2021 <<https://osf.io/m9abx/>>.
- Moher 2020 = David Moher, Lex Bouter, Sabine Kleinert, Paul Glasziou, Mai Har Sham, Virginia Barbour, Anne-Marie Coriat, Nicole Foeger, Ulrich Dirnagl, *The Hong Kong Principles for assessing researchers: Fostering research integrity*, «PLOS Biology» Published: July 16, 2020 <<https://doi.org/10.1371/journal.pbio.3000737>>.
- Next Generation Repositories 2017 = *Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group*, November 28, <<https://www.coar-repositories.org/files/NGR-Final-Formatted-Report-cc.pdf>>.
- Shearer - Chan - Kuchma - Mounier 2020 = Kathleen Shearer, Leslie Chan, Iryna Kuchma, Pierre Mounier, *Fostering Bibliodiversity in Scholarly Communications A Call for Action!* April 15, 2020, COAR <<https://www.coar-repositories.org/news-updates/fostering-bibliodiversity-in-scholarly-communications-a-call-for-action/>>.
- Wouters 2014 = Paul Wouters, *The citation: from culture to infrastructure, in Beyond Bibliometrics: harnessing multidimensional indicators of scholarly impact*, edited by Blaise Cronin and Cassidy R. Sugimoto, Cambridge (Mass.) - London, The MIT Press, p. 47-66.
- Wouters 2019 = Paul Wouters, Sugimoto, C. R., Larivière, V., McVeigh, M. E., Pulverer, B., de Rijcke, S., Waltman, L. *Rethink impact factors: find new ways to judge a journal*. «Nature», 569, May, pp. 621-623. Comment 28 May 2019, <<https://media.nature.com/original/magazine-assets/d41586-019-01643-3/d41586-019-01643-3.pdf>>.

Abstract

L'uso degli indicatori bibliometrici e in particolare dell'Impact Factor delle riviste nella valutazione della produzione scientifica dei ricercatori è stato recentemente sottoposto a revisione in ambito internazionale, facendo emergere la necessità di riconsiderare l'attività di peer-review come procedura che permette di riconoscere le diverse attività collegate alla ricerca degli studiosi. Nell'articolo si prendono in considerazione alcuni rapporti pubblicati in ambito internazionale – DORA, The Leiden Manifesto, Rapporto della Commissione europea del 2021 – e si evidenzia la connessione significativa tra le buone pratiche valutative e il Manifesto di Hong Kong, dedicato alla integrità della ricerca e l'Etica della valutazione della ricerca.

Valutazione della ricerca; Indicatori bibliometrici; Impact Factor; Peer-review; Ricerca; Etica

The use of bibliometric indicators, and in particular of the Impact Factor of journals in evaluating the scientific production of researchers, has recently been subjected to international review, highlighting the need to reconsider the peer-review activity as a procedure that allows to recognize the different activities connected to the research of scholars. The article discusses some reports published internationally – DORA, The Leiden Manifesto, Report of the European Commission of 2021 – and highlights the significant connection between good evaluation practices and the Hong Kong Manifesto, dedicated to research integrity and the ethics of research evaluation.

Research evaluation; Bibliometric indicators; impact factor; Peer-review; Research; Ethics