



**UNIVERSITA' DEGLI STUDI DI PADOVA**  
**DIPARTIMENTO DI SCIENZE ECONOMICHE E**  
**AZIENDALI "MARCO FANNO"**

**CORSO DI LAUREA IN ECONOMIA INTERNAZIONALE**  
*L-33 Classe delle lauree in SCIENZE ECONOMICHE*

Tesi di laurea

**Produzione scientifica e bibliometria in economia**  
*Scientific production and bibliometris in economics*

Relatore:  
Prof. CAINELLI GIULIO

Laureando:  
MORETTO NICOLO' FABIO

Anno Accademico 2016-2017

## **1. Introduzione**

La valutazione della ricerca scientifica negli ultimi anni ha assunto un ruolo fondamentale nel dibattito politico-istituzionale italiano. Essa è la principale strategia utilizzata dal governo delle università e degli enti di ricerca non solo per misurare il merito dei singoli ricercatori e docenti ai fini della loro selezione, ma anche per la distribuzione di fondi per la ricerca il cui scopo è produrre benessere al Paese. Gli investimenti in ricerca scientifica costituiscono infatti il principale strumento attraverso il cui il governo può alimentare la crescita nel medio-lungo periodo. La necessità di misurare l'attività di ricerca di una istituzione, quale si deduce dai lavori scientifici pubblicati, è stato il presupposto per l'affermarsi di indicatori capaci di dare un punteggio alla rivista che li diffonde. Questi indicatori sono quelli utilizzati dalla Bibliometria, per valutare quantitativamente, per mezzo di modelli matematici e statistici, la produzione dei documenti, il loro uso e peso - o impatto - all'interno delle comunità scientifiche. Prende il nome di Scientometria, perché essenzialmente orientata, sin dalla sua nascita alla fine dell'Ottocento, alla valutazione della documentazione scientifica. Tra gli indicatori bibliometrici la famiglia più numerosa è composta da quelli citazionali. La tesi si apre con un'analisi dei principali strumenti della valutazione della ricerca, cercando di comprendere al meglio i punti di forza ed evidenziandone i limiti strutturali. L'attenzione poi si sposta in Italia, dove nella Valutazione della Qualità e della Ricerca (VQR) gli strumenti di valutazione della ricerca vengono utilizzati per dare un vero e proprio punteggio alle università ed enti di ricerca, fondamentale per la distribuzione del Fondo di Finanziamento Ordinario, che costituisce la principale fonte di entrate per le università italiane. Nel secondo capitolo è dedicato alla rassegna dei principali indicatori bibliometrici, dagli indicatori semplici a quelli derivati, con lo scopo di comprendere quali informazioni possono essere estrapolate da ognuno di essi. Il lavoro si conclude con un approfondimento sulla VQR 2011-2014, evidenziando la struttura e la metodologia utilizzata per svolgere il rapporto, riportando i principali risultati ottenuti nelle varie aree di ricerca con una focalizzazione finale ai risultati ottenuti dall'Università degli Studi di Padova.

## **2. Valutazione della ricerca scientifica**

### **2.1 Storia degli indici citazionali**

I primi tre indici citazionali Science Citation Index (SCI), Social Sciences Citation Index (SSCI), Art & Humanities Citation Index (A&HCI) furono pubblicati da Eugene Garfield

dal 1964 al 1978 ed editi dall'Istituto per l'Informazione Scientifica (ISI), da lui fondato. L'idea motrice della banca dati dell'ISI (Institute of Scientific Information), attribuibile al suo fondatore, Eugene Garfield, è stata di pubblicare un repertorio delle pubblicazioni scientifiche, che includesse l'elenco dei lavori citati, da ciascun lavoro preso in esame. Lo scopo che si proponevano gli ideatori era di facilitare ricerche bibliografiche che, a partire da un lavoro importante del passato, consentissero di identificare i lavori recenti che ne sviluppavano i risultati. Un indice aggiornato delle citazioni scientifiche può fornire quindi uno strumento molto utile per seguire gli sviluppi di una problematica o di un settore molto particolare di una disciplina. Quest'idea ha portato alla creazione di un archivio elettronico delle citazioni che è stato denominato appunto "Science Citation Index" (SCI). L'archivio fu subito molto popolare nella Chimica e nella Biologia, dove ha costituito un utile repertorio per ricerche bibliografiche. Molto meno nella fisica dove, a livello mondiale, si era creato un sistema informale di comunicazione mediante "preprints" che rendeva meno essenziale l'accesso ad un repertorio della produzione scientifica. Ancor meno nella matematica dove esistevano tre riviste (una tedesca, una sovietica ed una americana) dedicate esclusivamente alla pubblicazione di recensioni di articoli di matematica. Queste riviste, ed in particolare quella edita dalla American Mathematical Society, "Mathematical Reviews", costituivano e costituiscono la fonte principale per ricerche bibliografiche (ma, ovviamente, non consentivano le ricerche "in avanti", se non attraverso l'indice degli argomenti.) E, in effetti, per trovare lo SCI in una buona università americana negli anni sessanta bisognava andare a cercarlo nella biblioteca (generale o dipartimentale) utilizzata dai chimici. Questi indici cominciarono a interessare agli editori commerciali: fu così che Thomson Reuters acquistò ISI nel 1992 e lo portò nel web come Web of Science (WoS). Nel 1997, le potenzialità commerciali di un repertorio quale lo SCI diventarono una realtà concreta per i grandi editori. Il fattore di impatto delle riviste – Impact Factor – il primo e più famoso di questi indicatori, insieme ad altri dal 1992 verrà conteggiato in archivi di proprietà di grossi editori, come Wos e dal 2004 Scopus dell'Elsevier, rivale di Thomson Reuters.

“L'Impact factor nasce con lo scopo di indicare il peso (il fattore di impatto) di una rivista all'interno del suo settore disciplinare specifico ed è [...] in pratica il numero medio di citazioni ricevute da un singolo articolo pubblicato da una determinata rivista in un certo lasso di tempo.” (Piazzini, 2010, p.68).

## 2.2 Strumenti per la valutazione della ricerca

Per Baccini (2010) il processo di valutazione della ricerca consiste, sempre, nel riconoscimento da parte di uno o più soggetti del contributo che il prodotto della ricerca apporta nell'area scientifica di riferimento. In Italia, il CIVR individua le due fondamentali metodologie per la valutazione della ricerca: (a) la valutazione di tipo qualitativo, basata sulla revisione dei pari (peer review). Questo sistema, chiamato comunemente paper, è per definizione soggettivo e viene svolta da un ricercatore con le stesse competenze disciplinari del tema in esame; (b) la valutazione quantitativa, strutturata su indicatori volti a misurare l'impatto di una ricerca sulla comunità scientifica. E' basata anch'essa su una valutazione "soggettiva", non sotto forma di giudizio da parte di un valutatore come nel caso della peer review ma attribuendo alla citazione la funzione di quantificare l'impatto che quella pubblicazione ha avuto nel settore di riferimento.

La revisione dei pari è definita come "un insieme di pratiche eterogenee e non standardizzate, attraverso le quali un gruppo di individui esprime un giudizio sul lavoro scientifico di altri per determinarne la qualità" (Baccini, 2010, p.52). E' il sistema per cui viene valutata la qualità di una ricerca, per decidere se pubblicarla su una rivista o finanziarla. La revisione avviene prima dell'accettazione ufficiale dell'elaborato presentato alla rivista o come progetto di ricerca. Per garantire l'imparzialità, l'identità dei referee rimane segreta.

Nella revisione, il valutatore è chiamato ad esprimere un giudizio soggettivo sulla qualità, originalità e potenzialità che la produzione scientifica può avere nello sviluppo della ricerca. Solo il valutatore è in grado di entrare in merito al valore "interno" della ricerca, il contributo di ogni singolo autore, plagi, errori e, se si pensa al caso dei revisori delle riviste scientifiche, l'assenza di eventuale manipolazione dei risultati.

Nonostante questo criterio di valutazione sia largamente predominante nella comunità scientifica ne sono ben noti i limiti in letteratura. Sembra abbastanza logico pensare al ruolo determinante che sia il giudizio sia la soggettività del referee può avere nell'esito della valutazione. Può essere quindi che la valutazione sia influenzata quindi, dalla personalità dei referees e dai loro pregiudizi culturali. De Bellis individua altri due limiti che rendono la valutazione dei pari fallibile: "Il primo è conseguente alla mancanza di standard oggettivi nella comparazione delle performance individuali o aggregate che rendono gli esiti delle valutazioni solo parzialmente confrontabili. Il secondo limite è connesso alla possibilità che il potere discrezionale del giudizio dei pari sia guidato oltre che da impostazioni culturali o metodologiche anche da atteggiamenti opportunistici dettati ad esempio da conflitti di interessi.

A fronte dei limiti della peer review e soprattutto grazie alla grande diffusione delle banche dati citazionali, prima fra tutte l'ISI, negli ultimi anni è andato ad affermarsi l'uso di indicatori numerici atti a misurare, in modo equo e sintetico, la quantità e l'impatto della produzione scientifica. Tali indici, chiamati indici bibliometrici, se da una parte non entrano in merito alla qualità interna del prodotto (pubblicazione), o competenze del produttore (ricercatore), forniscono indicazioni accurate circa il grado di "utilizzo" (impatto) che quella pubblicazione ha avuto nel progresso della conoscenza.

Come per la valutazione dei pari anche l'analisi citazionale presenta dei limiti e l'utilizzatore finale deve procedere con cautela. L'attendibilità dell'analisi citazionale dipende da alcuni fattori, tra cui: (1) la dipendenza del risultato dall'area scientifica che si analizza. In particolare è noto che l'analisi citazionale sia particolarmente efficace per le valutazioni nell'area delle scienze naturali mentre è ritenuta meno indicata per le scienze sociali; (2) dipende anche in modo eccessivo dal grado di copertura delle banche dati citazionali. Ci sono inoltre altri limiti di cui tenere conto quando si parla di indici citazionali, e la letteratura ne è piena. Baccini (2010) propone un interessante rassegna dei principali fattori che indeboliscono l'ipotesi di correlazione tra quantità delle citazioni e qualità della ricerca data.

1. Tempo. Sulla base della letteratura empirica esistente, si può affermare che la probabilità di essere citati non si distribuisce nel tempo in modo uniforme ma segue un andamento asimmetrico (Baccini,2010). Tahai e Rigsby (1998) stimano, con riferimento alle citazioni presenti in riviste di area aziendalistica, una distribuzione di frequenza che si approssima ad una variabile Gamma (generalizzata) il cui valore modale si realizza dopo tre anni dalla pubblicazione, mentre mediana e media si collocano rispettivamente dopo sette e nove anni dalla data pubblicazione della rivista. Più recentemente Adler et al. (2008) confermano, con riferimento alle pubblicazioni del 2003 presenti nella banca dati ISI e classificate in "Economics", l'andamento asimmetrico della distribuzione temporale delle citazioni. In particolare, il numero di citazioni annuo è crescente per il primo quadriennio, rimane pressoché costante dal quarto al sesto anno, per poi decrescere progressivamente dal settimo anno. Questo risultato è indirettamente confermato dall'analisi di De Bellis (2005) secondo cui circa il 50% delle citazioni presenti nelle pubblicazioni (anche se vi è una rilevante eterogeneità tra le aree scientifiche) usualmente cita articoli con età compresa tra uno e sei anni. In questo senso, Il valore degli indici bibliometrici sarà distorto (verso il basso) per quelle unità di ricerca (ad es. ricercatori) la cui produzione scientifica è recente. Per tenere ridurre tale

distorsione la letteratura ha suggerito l'introduzione di correttivi per l'età di pubblicazione dei prodotti valutati.

2. Dimensione della comunità scientifica. Moed et al. (1985) mostrarono che la probabilità di essere citati dipende in modo rilevante dal numero di ricercatori che operano nella comunità scientifica di riferimento. In questo senso, il punteggio citazionale potrà essere più o meno elevato in funzione della dimensione della comunità di studiosi a cui si rivolge quella ricerca.
3. Prassi redazionale prevalente nella disciplina. Questa critica evidenzia come il numero di citazioni dipenda dal numero di co-autori delle pubblicazioni, dalla tipologia di editoriale prevalente (articoli piuttosto che monografie); dalla lingua di riferimento (nazionale piuttosto che inglese); dalla vita citazionale media delle pubblicazioni. Su questo ultimo punto Adler et al. (2008) mostrano come, per le scienze "forti" la distribuzione delle citazioni nel tempo ha un valore modale anticipato rispetto alle scienze sociali, ma a questo corrisponde anche una loro più veloce obsolescenza scientifica.
4. Etica citazionale. Gli indici bibliometrici possono essere manipolati da un uso strategico delle citazioni. Tra la letteratura che si occupa di questo problema, Beaver (2004) stima una correlazione positiva tra numero di citazioni e numero di coautori dell'articolo citato. La spiegazione è da rinvenirsi nell'effetto che gli autori avrebbero nell'influire sulle citazioni dell'articolo (ad es. per le autocitazioni). In questa ipotesi si muove la ricerca di Mahlck e Persson (2000), secondo cui, più ampia è la rete di conoscenze dell'autore maggiore sarà il numero di citazioni che riceveranno le sue pubblicazioni.
5. Collocazione editoriale. Moed et al. (1985), Van Dalen e Henkens (2005) evidenziano una correlazione positiva tra visibilità, intesa come grado di internalizzazione, prestigio e/o impact factor, e numero di citazioni. Laband e Piette (1994) e Smart e Waldfogel (1996) segnalano che, nelle riviste di economia e finanza, gli articoli collocati nella parte iniziale della rivista riceverebbero più citazioni delle altre.
6. Tipologia di contenuto scientifico. Gli articoli che propongono nuove metodologie di indagine o le rassegne della letteratura, ricevono mediamente più citazioni rispetto alle altre tipologie di contenuti (Cano e Lind, 1991). La ratio di questa distorsione (verso

l'alto) dell'indice citazionale è nella maggiore dimensione del bacino dei potenziali fruitori della pubblicazione.

7. Errori nelle bibliografie. Schulmeister (1998) e Evans et al. (1990), controllando le citazioni riportate in un campione di articoli in ambito medico, evidenziarono come il numero di citazioni errate (rispettivamente del 32% e 48% rispetto al totale) era tale da inficiare profondamente l'affidabilità dei dati bibliometrici.
8. Difficoltà tecniche con cui l'analizzatore-valutatore bibliometrico deve confrontarsi. Come sa bene chi si cimenta in analisi citazionali, il maggiore problema operativo consiste nel depurare i risultati da omonimie, sinomie, cognomi e nomi multipli (in particolare per gli autori di origine spagnola, cinese, coreana, ecc.) e l'inclusione nel cognome dell'autore di segni ortografici (apostrofi, trattini, pedici, ecc.). L'esperienza di questa ricerca conferma quanto rilevato da Baccini (2010).”

### **2.3 Valutazione della ricerca in Italia**

Pur con la consapevolezza che non sia perfetta e priva di errori strutturali, l'analisi citazionale rimane il metodo più funzionale per governi e organizzazioni per stimare un'analisi quantitativa per quanto riguarda la produzione scientifica. Il governo Italiano ne è consapevole e nel 1994 cambia il sistema di finanziamento dell'università italiana con la legge finanziaria approvata nel dicembre 1993. Con la nuova legge finanziaria le risorse non saranno più distribuite agli atenei in singoli capitoli di spesa, ma riunite in un unico fondo. Viene inoltre introdotta la “funzione valutativa”, anche se non sono chiariti i parametri con cui effettuare questa valutazione. L'obiettivo finale è quello di distribuire parte dei fondi ministeriali in modo più meritocratico, premiando gli atenei che soddisfano meglio una serie di parametri. La stessa legge prevede anche l'istituzione dell'Osservatorio per la valutazione del sistema universitario, un organo tecnico del Ministero composto da esperti esterni. Istituito formalmente nel 1996, ha il compito di “valutare l'efficienza e la produttività delle attività di ricerca e di formazione e verificare i piani di sviluppo e di riequilibrio del sistema universitario”. In questo modo si introduce una valutazione esterna svolta da esperti nominati dal Ministero. Negli anni successivi nascono anche il Civr (Comitato di indirizzo per la valutazione della ricerca) e il Cnvsu (Comitato nazionale per la valutazione del Sistema universitario), che saranno poi soppiantati dall'ANVUR (Agenzia Nazionale Valutazione del sistema Universitario e della Ricerca), istituita formalmente con la

Finanziaria del 2007 per “razionalizzare il sistema di valutazione della qualità delle attività delle università e degli enti di ricerca pubblici e privati destinatari di finanziamenti pubblici, nonché dell’efficienza ed efficacia dei programmi statali di finanziamento e di incentivazione delle attività di ricerca e di innovazione” (Marco Scanu).

L’ANVUR è un’agenzia nazionale di diritto pubblico autonoma, che risponde al ministero dell’istruzione dell’Università e della Ricerca (Miur) e nell’idea dei suoi fondatori dovrebbe fungere da contraltare all’autonomia dell’università, essere un sistema di bilanciamento che permetta di valutare il merito di ciò che viene fatto. L’ANVUR fissa i parametri per la valutazione e da essi stende relazioni periodiche dove da esse viene distribuita la quota premiale del Fondo di Finanziamento Ordinario (Ffo). Tra gli elementi presi in considerazione dall’ANVUR c’è la qualità della ricerca (VQR). Si tratta di report svolti in un arco temporale variabile (la prima VQR comprendeva un’analisi dal 2004-2010, quella successiva 2011-2014) in cui viene chiesto a università e centri di ricerca di presentare i loro ‘prodotti di ricerca’.

### **3 Gli indicatori bibliometrici**

Una efficace definizione di indicatore bibliometrico è stata proposta da Baccini che scrive (Baccini 2010 p.68): «un indicatore bibliometrico è un indicatore di qualità o impatto costruito, con opportune tecniche statistiche, a partire da informazioni elementari ricavate da riferimenti bibliografici contenuti in pubblicazioni scientifiche o in archivi creati appositamente». Gli indicatori possono essere suddivisi principalmente in due grandi categorie: (1) indicatori bibliometrici primari, semplici misure quantitative come il conteggio delle pubblicazioni di un autore o di una rivista o del conteggio delle citazioni ottenute in un intervallo di tempo; (2) indicatori bibliometrici secondari o derivati, dove partendo da una base quantitativa come ad esempio il conteggio delle citazioni in un determinato anno degli articoli pubblicati in una rivista nei due anni precedenti diviso il numero di pubblicazioni nello stesso periodo dalla stessa rivista (Impact Factor). Alcune considerazioni di carattere generale sono necessarie prima di procedere. Rispetto alla tradizionale peer review, la valutazione della ricerca attraverso l’uso di indicatori bibliometrici è più veloce, più economico e per certi aspetti anche più obbiettivo, ma rimane sempre relativo al database preso in considerazione per calcolarne l’indice. Significa che l’*output* bibliometrico sarà tanto più accurato quanto più esaustivo sarà l’archivio bibliografico preso in considerazione.



### 3.1 Indicatori Primari

Gli indicatori bibliometrici primari sono misure quantitative elementari poiché consistono nel semplice conteggio nelle pubblicazioni prodotte da un autore o da un gruppo di ricerca, o dal conteggio delle citazioni ottenute in un determinato intervallo di tempo o da un autore o da una rivista. Gli indicatori bibliometrici primari inoltre possono essere la base di partenza per altri indicatori cosiddetti secondari o derivati. Essi sono essenzialmente due: (1) il conteggio del numero di pubblicazioni e (2) il *Citation Index*.

- (1) Il numero totale di pubblicazioni è un indicatore di produttività scientifica calcolabile sia a livello di singoli autori o di riviste sia a livello di istituzioni o di interi sistemi nazionali di ricerca. Come si può dedurre dal nome, è un indicatore piuttosto semplice e fornisce informazioni di carattere quantitativo. Non esprime quindi alcuna informazione riguardo l'impatto della pubblicazione nella comunità scientifica, a differenza dell'indicatore costituito dal conteggio delle citazioni ricevute, ma si tratta di un elementare conteggio delle pubblicazioni del soggetto di riferimento. La sua semplicità tuttavia non lo rende privo di problematiche in relazione alla sua costruzione. Il primo problema che possiamo incontrare è quello relativo al database bibliografico su cui viene effettuato il conteggio. Se da una parte sempre più spesso sono le stesse istituzioni di ricerca a costruire dei database bibliografici in cui il ricercatore può inserire le proprie pubblicazioni, rendendo molto più facile il calcolo di questo indicatore, dall'altra potrebbe aumentare la difficoltà aumenta nel caso di istituzioni molto complesse o di gruppi di ricerca appartenenti a paesi diversi. Nel caso in cui i gruppi di ricerca i soggetti appartengono a paesi diversi, si ricorre generalmente ad archivi bibliografici citazionali come ad esempio Web of Science o Scopus, di cui parlerò in seguito. Il secondo problema riguarda il tipo di pubblicazioni da includere nel conteggio. Per le discipline scientifiche è più semplice stabilire quali tipologie di pubblicazioni siano più adatte per il calcolo di questo indicatore, ovvero l'articolo su rivista, mentre risulta più complicato per le scienze umane dove ciò non accade. Il terzo aspetto problematico riguarda le modalità di conteggio delle pubblicazioni con più di un autore. Come ben spiegato da Baccini (2010, pp 102-103) sono tre le possibili soluzioni: attribuire l'intera pubblicazione a ciascuno degli autori; attribuire solo una frazione del contributo a ciascuno dei co-autori; o attribuire l'intera pubblicazione solo al primo autore. Infine, anche la definizione della base temporale su cui calcolare l'indicatore può creare qualche difficoltà soprattutto se

si pensa che la produttività di un ricercatore è direttamente proporzionale alla sua età anagrafica, anche in relazione a un determinato intervallo di tempo.

- (2) Il secondo indicatore fondamentale primario è il Citation Index, costruito semplicemente contando il numero di citazioni ricevute in una determinata pubblicazione, da un autore, da un gruppo di ricerca o istituzione in un intervallo di tempo definito. Come il conteggio di numero di pubblicazioni, anche il Citation Index è un indicatore di carattere quantitativo, ma a differenza del primo con questo possiamo ottenere importanti informazioni circa l'impatto che una determinata pubblicazione può avere nella comunità scientifica. Un Citation Index elevato quindi significa che quella determinata pubblicazione ha avuto un grande impatto nella comunità scientifica di riferimento. Il conteggio del numero di citazioni può essere applicato a qualsiasi livello di applicazione. Tuttavia è da tenere presente che a livello micro, corrispondente al conteggio del numero di citazioni della singola pubblicazione o del singolo autore, pur essendo più facile la reperibilità del dato, "inevitabilmente si perdono i vantaggi legati all'omogeneità e alle dimensioni del campione da cui ricavare le redistribuzioni probabilistiche, e ci si trova a fare i conti con l'ambiguità dell'individuo, di ciò che rende l'individuo unico e per certi versi inclassificabile" (De Bellis, 2005 p. 129). Se si procedesse ad un'analisi comparativa tra un ricercatore "giovane" e uno "anziano" (inteso come anni di carriera), come anche ad una comparazione tra due gruppi di ricerca di dimensioni diverse, il risultato sarebbe distorto, poiché sapendo che il numero di pubblicazioni è direttamente proporzionale agli anni di carriera, la probabilità di essere citati è sicuramente maggiore. Un tentativo per ovviare a questo problema è l'impiego di un indicatore noto come *Citations Per Publications* (CPP) che non è altro che il rapporto tra il numero di citazioni ottenute da un set di pubblicazioni, per il numero totale di pubblicazioni. Entrano in gioco altre problematiche, legate alla costruzione dei database citazionali, sbilanciate per quantità verso gli anni più recenti alla pubblicazione, che potrebbero distorcere l'analisi comparativa basata sull'uso di questo indicatore specie se calcolati in intervalli di tempo distanti tra loro, con la conseguenza che il dato relativo a intervalli di tempi più recenti può risultare più elevato. Per questo il *Citations Per Publications* viene accompagnato con altri due indicatori citazionali (Van Raan 2000; Van Leeuwen et al. 2003; Baccini 2010 p.159-160): l'impatto citazionale atteso delle riviste (JCS, Journal Citation Score), che non è altro che il tasso medio di citazioni ricevute dagli articoli contenuti nelle riviste su cui il ricercatore o i membri del gruppo di ricerca hanno pubblicato nell'arco

temporale di riferimento; e l'impatto citazionale atteso nel campo di ricerca (FCS, Field Citation Score), ossia il tasso medio di citazioni ottenute dalle pubblicazioni appartenente il settore disciplinare del gruppo di ricerca oggetto di valutazione. Insieme questi indicatori riescono a fornire un quadro molto più concreto agli occhi del valutatore, soprattutto con i valori medi nel campo di ricerca, proprio perché contestualizzato.

## 3.2 Indicatori Secondari o Derivati

### 3.2.1 L'Impact Factor (IF)

L'Impact Factor è il più noto indicatore elaborato dalla *Thomson-Reuters* e allo stesso tempo il più discusso. Come già accennato nel capitolo precedente, l'indicatore è stato elaborato nel 1955 da Eugene Garfield in fase di realizzazione del *Science Citation Index* come strumento utile per l'individuazione del nucleo di periodici da inserire nello stesso indice citazionale.

Il *Journal Citation Reports* (JCR), la pubblicazione in cui è contenuto l'Impact Factor, è distribuito in due edizioni annuali, la Science Edition e la Social Science Edition, partendo dai dati delle riviste indicizzate nello *Science Citation Index* e nel *Social Science Index* (ma non nell'*Arts & Humanities Citation Index*) contenuti in Web Of Science, stila una classifica annuale della performance delle riviste ivi contenute sulla base non solo dell'*Impact Factor* ma anche di diversi altri parametri (*Total Cites*, *Immediacy Index*, *5Year Impact Factor*, *Citable Items*, *Cited HalfLife*). L'algoritmo per calcolare l'Impact Factor è piuttosto semplice. Esso si ottiene dividendo il numero di citazioni ricevute in un determinato anno dagli articoli pubblicati dalla rivista nei due anni precedenti, per il numero totale degli articoli pubblicati nello stesso periodo di tempo e dalla stessa rivista.

Per esempio, se nel 2011 gli articoli pubblicati nel 2009 e 2010 su una rivista (X) hanno ricevuto rispettivamente 77 e 110 citazioni, e se il numero complessivo degli articoli usciti su quella stessa rivista nel 2009 e 2010 è 52, l'IF calcolato per il 2011 per la rivista X è pari a 3,59:

$$IF(X) = \frac{(77 + 110)}{52} = 3,59$$

Come si evince dalla sua costruzione l'IF è sostanzialmente un indicatore della performance citazionale delle riviste scientifiche presenti nel database ISI, molto utile quindi in campo

biblioteconomico come supporto alle decisioni di politica documentaria (Altmann - Gorman 1996), ma utilizzato in modo del tutto improprio come strumento di valutazione dei singoli ricercatori e dei singoli dipartimenti e gruppi di ricerca. E questa è soltanto una delle tante critiche che, dal lancio dell'indicatore, si è levata contro l'Impact Factor. Riassumerle tutte sarebbe impossibile e, vista la ridondanza spesso riscontrata, forse anche inutile. Schematicamente possono tuttavia essere indicati i punti di maggior dissenso che emergono più frequentemente nella letteratura del settore (Figà-Talamanca 2000; Tammaro 2001; Di Cesare 2002 p. 145; De Bellis 2005 p. 118; Morriello 2007; Cassella – Bozzarelli 2011 p.71).

1. la copertura dei database Web of Science non è adeguata e omogenea per tutti gli ambiti disciplinari (vedi paragrafo 4.3.1. e 4.5), in particolare per le scienze umane e sociali. A ciò si deve poi aggiungere il fatto che, pur contenendo Web of Science un indice specificatamente dedicato alle riviste di arte e discipline umanistiche (Arts & Humanities Citation Index), non esiste un'edizione specifica di JCR dedicata a tali discipline. JCR riporta l'IF delle riviste indicizzate in Science Citation Index e in Social Science Citation Index ma non di quelle contenute in Arts & Humanities Citation Index, lasciando di fatto priva del dato bibliometrico un'importante fetta di letteratura pur indicizzata nel database;
2. Non considera il comportamento citazionale diffuso nelle diverse aree disciplinari, impedendo di fatto un affidabile confronto tra l'Impact Factor non solo di riviste appartenenti ad ambiti differenti,<sup>5</sup> ma anche di quelle considerate trasversali, che coprono cioè ambiti disciplinari diversi (Tammaro 2001 p. 105), e di quelle che, pur riconducibili allo stesso ambito disciplinare, sono incentrate su specifiche aree tematiche (per esempio la biologia teorica rispetto a quella sperimentale e applicata) (Morriello 2007 p. 25);
3. non c'è nessuna segnalazione circa eventuali scostamenti rispetto ai valori medi e, di conseguenza, mancando un termine di confronto, diventa difficile per il valutatore saper attribuire il giusto significato al semplice dato che, in tal modo, non risulta contestualizzato nella sequenza statistica;
4. la base temporale di due anni su cui è calcolato l'IF è penalizzante per alcune discipline la cui performance citazionale, se si considera il tempo medio che passa tra la pubblicazione di un articolo e il picco del numero di citazioni ricevute, si presenta più precoce o più tardiva; alcuni studi (Adler – Ewing – Taylor 2008; Figà-Talamanca 2000)

- hanno registrato nelle riviste di matematica un picco altissimo (90%) di citazioni anteriore all'intervallo di due anni utilizzato per il calcolo dell'indicatore;
5. l'impossibilità di controllare i dati utilizzati per il calcolo dell'indicatore soprattutto alla luce di evidenti e documentati casi di errore;
  6. rischio di distorsione dei risultati dovuto all'incertezza nella definizione delle grandezze, cioè delle tipologie documentarie citabili, al numeratore e al denominatore (Seglen 1997 p. 500; Amin 2000 p. 5). Nel numeratore, che consiste nella somma delle citazioni ricevute da una rivista, si prendono infatti in considerazione, senza esercitare alcun tipo di selezione, tutte le tipologie di articolo da cui proviene la citazione. Nel denominatore invece si tiene conto soltanto degli articoli scientifici, degli atti di convegni e delle rassegne. Non vengono quindi conteggiati al denominatore alcune tipologie come le note, gli editoriali e le lettere agli editori, che tuttavia in molti casi contribuiscono massicciamente all'aumento del numero delle citazioni del numeratore. Ne consegue che le riviste con molte note, editoriali e lettere, magari anche molto citati, vedranno aumentato il valore del numeratore, senza però un corrispondente aumento del denominatore, situazione questa che chiaramente comporterà un conseguente e ingiustificato aumento dell'Impact Factor (Van Leeuwen – Moed 2002);
  7. non si tiene conto né della grandezza del periodico, ovvero del numero medio di articoli per rivista, né delle tipologie di scritti che vi appaiono. Sembra infatti quasi scontato che più alto è il numero di articoli pubblicati da una rivista, più probabilità avrà quest'ultima di venire citata. Lo stesso vale per le tipologie di scritti: le rassegne sono più citate degli articoli e quest'ultimi, se sono più lunghi, sono citati di più di quelli più brevi (Seglen 1997). Le riviste quindi con un numero medio di articoli più alto, che pubblicano rassegne o articoli mediamente più lunghi riportano un valore d'IF più elevato;
  8. l'uso pilotato delle autocitazioni, cioè delle citazioni ad articoli della rivista provenienti da altri articoli della rivista stessa, da parte dei comitati editoriali delle riviste allo scopo di gonfiare i dati citazionali e ottenere così un valore più alto dell'IF, è un problema ancora presente ma che è stato in parte risolto grazie alla possibilità offerta dalla ThomsonReuters di calcolare il valore dell'indicatore escludendo le autocitazioni;
  9. la mancata pulizia di tutti quegli articoli considerati errati dalla comunità scientifica o dagli stessi autori, è un elemento di grande distorsione dell'indicatore poiché tali articoli continuano ad essere citati e, di conseguenza, contribuiscono anch'essi, pur essendo destituiti di ogni credenziale, a formare il valore dell'IF; Nonostante i punti di criticità appena descritti, l'IF presenta anche molte caratteristiche positive che hanno contribuito

a farlo diventare uno degli indicatori citazionali più utilizzati. Baccini (2010 p. 167) ha individuato tra i fattori di maggior successo:

- a) la semplicità del suo algoritmo rende chiari e quindi facilmente comprensibili i meccanismi del suo funzionamento;
- b) la stabilità della misura nel tempo. I valori dell'indicatore non registrano mai grandi variazioni da un anno all'altro e ciò permette una minore obsolescenza del dato, che quindi può essere utilizzato anche se vecchio di uno o due anni senza particolari rischi d'errore;
- c) il costante aggiornamento del dato;
- d) l'idea della facile controllabilità del dato, ovvero la sicurezza purtroppo solo apparente, che sia possibile controllare in ogni momento i dati su cui si basa il calcolo dell'indicatore;
- e) la possibilità di confrontare riviste con periodicità di pubblicazione e dimensioni diverse.

### **3.2.2 Altri indicatori contenuti nel *Journal Citation Reports***

L'Impact Factor non è l'unico indice contenuti nel Journal Citation Reports. Sono pubblicati altri indicatori, sempre calcolati sulla base degli archivi ISI e in parte rispondono alle critiche mosse contro l'indicatore sopra riportato. Dal 2007 è stato aggiunto il 5-Year Impact Factor, che altro non è che lo stesso algoritmo dell'Impact Factor, calcolato non più prendendo in considerazione gli articoli pubblicati nei 5 anni precedenti e non due. Questo indicatore sembra una risposta alle critiche sulla ristrettezza temporale dell'Impact Factor, che a detta di molti sembra non cogliere la specificità disciplinari del comportamento citazionale.

Troviamo poi l'Immediacy Index che, al contrario, calcola il numero medio di citazioni ricevute in un determinato anno dagli articoli pubblicati nello stesso anno dalla rivista presa in esame. Come obiettivo ha quello di misurare la velocità con cui l'articolo viene citato mediamente.

Infine il Cited Half-Life misura praticamente l'obsolescenza degli articoli citati o la stabilità nel tempo delle citazioni ricevute in una rivista. Esso calcola il numero di anni, partendo dall'anno in corso e andando a ritroso nel tempo, in cui si raggiunge la metà del numero di citazioni ricevute dalla rivista nell'anno corrente.

### 3.2.3 *H-index* e indici derivati

H-Index o indice di Hirsh viene ufficialmente presentato dal suo ideatore fisico George E. Hirsh, docente presso l'Università della California a San Diego, nel preprint di un articolo pubblicato il 3 agosto 2005 sul *repository ArXiv* (Hirsh, 2005).

La peculiarità dell'*H-Index* è che è in grado di esprimere in unico valore numerico sia la produttività (numero di pubblicazioni) sia l'impatto nella comunità scientifica (numero di citazioni) riferiti a un singolo ricercatore. Hirsh definisce l'indicatore rivolgendosi direttamente al ricercatore che "ha un indice  $h$  se  $h$  dei suoi articoli pubblicati in  $n$  anni ( $N_p$ ) hanno ricevuto ciascuno almeno  $h$  citazioni, e gli altri ( $N_p-h$ ) articoli hanno ricevuto non più di  $h$  citazioni ciascuno" (Hirsh 2005 p.1).

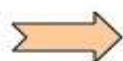
Per calcolare l'indice di Hirsh per un determinato autore è necessario: definire l'intervallo temporale di riferimento e attribuire alle pubblicazioni un ordine decrescente basato sul numero di citazioni ricevute. A questo punto, l'*H-Index* di quell'autore sarà determinato dall'ultima pubblicazione in tale classifica che riuscirà a far corrispondere il numero di posizione occupato in tale elenco con una quantità almeno pari di citazioni ricevute.

Per capire meglio come venga individuato *H-Index* di un autore è necessario fare un esempio più concreto che faccia risaltare le fasi operative di calcolo. Supponiamo di avere un autore che ha pubblicato 8 articoli, i quali hanno riportato complessivamente un totale di 45 citazioni.

Set dati di partenza

P	NC
A	4
B	7
C	6
D	7
E	9
F	1
G	11
H	0

P= pubblicazioni  
NC= numero di citazioni



formiamo un elenco disponendo le pubblicazioni in ordine decrescente in base alle citazioni ricevute

Set dati ordinati

P	POS.	NC
G	1	11
E	2	9
B	3	7
D	4	7
C	5	6
A	6	4
F	7	1
H	8	0

POS.= posizione  
P= pubblicazioni  
NC= numero di citazioni

**H-Index = 5**

**H-Index** è pari a 5 perché ci sono 5 pubblicazioni che hanno ricevuto almeno 5 citazioni ciascuna

L'*H-Index*, non essendo un prodotto proprietario, è inserito tra gli strumenti bibliometrici sia in database citazionali come Web of Science e Scopus, sia in motori di ricerca specialistici come Google Scholar. E' utile ricordare che l'*H-Index* di un determinato autore potrebbe risultare diverso se calcolato sulla base di database citazionali differenti.

Come gli altri indicatori inoltre, anche l'*H-Index* presenta punti di forza e punti di debolezza.

I punti di forza sono (Baccini 2010, p.176):

- a) La semplicità dell'algoritmo;
- b) La capacità di cogliere entrambi gli aspetti fondamentali dell'attività scientifica: produttività e impatto della ricerca;
- c) Può essere applicato facilmente a livello di indagine sia a livello macro sia a livello micro: dal singolo ricercatore alla nazione;
- d) Gli eventuali piccoli errori presenti nel database citazionale di riferimento non influenzano, o comunque, non modificano in modo determinante, i risultati espressi dall'indicatore;
- e) I picchi citazionali di una o poche pubblicazioni non modificano in nessun caso il valore dell'indicatore;
- f) Le pubblicazioni che hanno ricevuto poche o nessuna citazione non condizionano il calcolo dell'indicatore.

Uno degli svantaggi invece più evidenti sta proprio nel suo punto di forza. Se da una parte l'indicatore è facilmente calcolabile e applicabile, dall'altra possono venire a mancare alcune informazioni presenti nel dato citazionale del singolo ricercatore, con la conseguenza che la valutazione può risultare distorta. Prendiamo come esempio due ricercatori con un valore dell'indice  $h$  pari a 10. Significa che entrambi i ricercatori hanno almeno 10 pubblicazioni con 10 citazioni ciascuna. Uno dei due ricercatori ha altre 80 pubblicazioni con 9 citazioni ciascuna, mentre l'altro ricercatore ha 15 pubblicazioni con 100 citazioni ciascuna. Sebbene possa risultare un esempio inverosimile, si può capire come l'indice  $h$  non riesca ad estrapolare delle informazioni che permetterebbero una valutazione più efficace, specie se si tratta di una comparazione tra due o più autori. Tra le critiche riguardo questo indicatore possiamo trovare: la penalizzazione dei ricercatori più giovani o in formazione e di tutti quelli che pubblicano pochi articoli di grande impatto scientifico, dunque molto citati; l'impossibilità di calcolare l'indice al netto delle autocitazioni; non tiene conto di eventuali cambiamenti, in termini di produttività e impatto, di unità di ricerca poiché non può



decreocere; nel caso di pubblicazioni con più di un autore, attribuisce l'intera pubblicazione a ciascuno degli autori, trascurando quindi totalmente il contributo individuale al lavoro; influisce negativamente sul comportamento del singolo ricercatore perché se da una parte è incoraggiato a pubblicare lavori di qualità che facciano guadagnare molte citazioni per contributo, dall'altra però lo spinge, una volta raggiunto un valore dell'indice elevato, a "cullarsi sugli allori" visto che il suo indice può crescere anche senza pubblicare (Baccini 2010, p. 180), o peggio, può essere portato a frammentare la produzione scientifica in unità informative minime che avranno conseguenza tante brevi pubblicazioni ingiustificatamente slegate le une dalle altre con il rischio di far perdere organicità al disegno complessivo del contributo scientifico. Tutti questi punti critici nell'indicatore hanno contribuito a ideare altri strumenti bibliometrici derivanti dall'indice h ma capaci di correggere in parte i punti deboli appena affrontati. Il primo tenta di rimediare al problema che l'indice h tende a favorire i ricercatori più longevi in ragione del fatto che l'indice h cresce proporzionalmente al passare del tempo, penalizzando i soggetti più giovani. Si tratta dello stesso Hirsh (2005) a proporre un indice normalizzato che prenda in considerazione l'età scientifica degli studiosi.

Tale variante è conosciuta come quoziente di m e, nella sostanza, è il rapporto tra H-Index calcolato per un determinato ricercatore sommando gli anni trascorsi dalla prima pubblicazione valida per il calcolo dell'indice stesso.

Un'altra variante dell'indice h è quella proposta da Leo Egghe nota come *G-Index*, che nasce allo scopo di correggere l'incapacità dell'indice di partenza di tener conto degli eventuali cambiamenti sia in positivo (pubblicazioni h-core con molte citazioni) sia in negativo (pubblicazioni con poche o nessuna citazione) rispetto al numero medio di citazioni ricevute dalle pubblicazioni di un ricercatore. L'obiettivo di Egghe è quello di far emergere tutte quelle pubblicazioni che hanno un numero di citazioni molto più elevato rispetto al valore calcolato con l'indice di Hirsh. Secondo la definizione data dallo stesso Egghe (2006a) "l'indice g è il numero d'ordine più grande (quanto gli articoli sono classificati in ordine decrescente per numero di citazioni ricevute) tale che i primi g articoli abbiano ricevuto [cumulativamente] almeno  $g^2$  citazioni". [Esempio]

*H-B-Index (Hirsh-Banks Index)* è definito dallo stesso ideatore, Michael Banks del Marx Planck Institute for Solid State Research, un'estensione dell'indice h (Banks 2006). Banks non ha fatto altro infatti che riprendere *H-Index*, così come sviluppato da Hirsh, e applicarlo, invece che al singolo ricercatore, ai composti usati in fisica dello stato solido o, più in generale, a un determinato argomento scientifico.

L'*A-Index*, creato dalla studiosa cinese Jin Bihui, utilizza le stesse informazioni necessarie al calcolo dell'indice *h* e consiste nel numero medio di citazioni ricevute dagli articoli inclusi nell'*h-core* (Bihui 2006, p.8).

Esistono anche varianti dell'*AIndex*: l'indice *r* e l'indice *ar*.

Il primo, proposto da Bihui e alcuni suoi collaboratori, nasce con lo scopo di misurare la portata delle citazioni ricevute dalle pubblicazioni di un'unità di ricerca che hanno ottenuto più citazioni. Esso consiste nel quadrato del numero complessivo delle citazioni ricevute dalle pubblicazioni che formano *hcore* (Bihui – Liming – Rousseau – Egghe 2007).

Il secondo, proposto dalla stessa Bihui in una breve nota (Bihui 2007 p. 6), ha l'obiettivo di correggere l'incapacità dell'indice *h* di tener conto delle eventuali variazioni della performance citazionale di un ricercatore o unità di ricerca. Come già visto infatti l'*H-Index*, oltre a non prendere in considerazione il numero effettivo di citazioni, non può mai decrescere. L'*ARIndex* è definito come la radice quadrata della somma del numero medio di citazioni per anno delle pubblicazioni incluse nell'*hcore*.

Altri indicatori derivati dall'indice *h* sono: *HCIndex*, *HLIndex* e *HL norm*.

*HCIndex* (contemporary *HIndex*) sviluppato da Sidiropoulos, Katsaros e Manolopoulos (Sidiropoulos – Katsaros – Manolopoulos 2006) allo scopo di dare più rilevanza, rispetto all'*HIndex*, alle pubblicazioni recenti in modo tale da far emergere l'impegno di tutti quei ricercatori che lavorano e pubblicano con continuità.

*HLIndex* (individual *HIndex*), sviluppato da Batista, Campiteli, Kinouchi e Martinez (Batista – Campiteli – Kinouchi – Martinez 2006), corregge il difetto dell'indice *h* di attribuire, nel caso di pubblicazioni con più di un autore, l'intero contributo a ciascuno degli autori.

Esso consiste nel rapporto tra il valore dell'*HIndex* e la media del numero di autori delle pubblicazioni che formano la base per il calcolo dell'indice *h*.

*HL norm* (*Normalized individual HIndex*), costruito e implementato dal tool Publish or Perish

(PoP), è una variante del precedente. Esso prima normalizza il numero di citazioni per ogni pubblicazione mediante una divisione tra il numero di citazioni e quello degli autori che hanno contribuito alla pubblicazione, poi individua l'indice *h* utilizzando però come base per il calcolo il numero di citazioni normalizzate.

## **4 La Valutazione della Qualità della Ricerca 2011-2014**

Nel secondo capitolo abbiamo parlato della valutazione della ricerca in Italia, della nascita dell'ANVUR e della VQR. L'obiettivo di questo ultimo capitolo è quello di analizzare i risultati della VQR 2011-2014 (denominata anche VQR2) per vedere come è cambiata la qualità della ricerca rispetto alla precedente valutazione (2004-2011) e la metodologia utilizzata per arrivare a tali risultati, con un particolare focus sui risultati dell'ateneo di Padova.

### **4.1 Struttura della VQR2**

La valutazione ha riguardato obbligatoriamente le università e gli enti di ricerca vigilati dal MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca), con la possibilità di permettere anche ad altri enti esterni di sottoporsi volontariamente alla valutazione. Alla VQR hanno partecipato 96 università, 18 fra enti di ricerca vigilati dal MIUR (12) e enti di ricerca assimilati (6) che si sono sottoposti volontariamente alla valutazione e 21 altre istituzioni (9 consorzi universitari e 12 enti di ricerca).

Il processo valutativo si articola in 16 Aree di Ricerca identificate dal Comitato Universitario Nazionale (CUN), due in più rispetto alla VQR1, dove per ogni area l'ANVUR ha costituito un Gruppo di Esperti della Valutazione (GEV) con il compito di valutare i prodotti della ricerca secondo criteri di originalità, rigore metodologico e impatto attestato o potenziale. I prodotti di ricerca presi in considerazione per la valutazione sono di vario tipo: dai contributi in rivista alla monografia scientifica, il significato a 'prodotto di ricerca' che da ANVUR è il più generale possibile, comprendendo anche banche dati, progetti architettonici, software ecc. Sta al GEV tuttavia decidere i criteri alla base della valutazione per ogni rispettiva area di ricerca, condividendo delle caratteristiche che possiamo riassumere in seguito:

- la responsabilità finale in capo ai GEV della valutazione dei prodotti con l'attribuzione delle classi di merito;
- la scelta di utilizzare per la valutazione la tecnica della informed peer review, che consiste nel tenere conto di più elementi di valutazione per la classificazione finale di merito. A seconda delle caratteristiche peculiari del GEV, gli elementi vanno dall'uso di due basi di dati per la valutazione bibliometrica, alla combinazione di valutazione peer e bibliometrica;
- l'uso della informed peer review per la valutazione delle monografie e dei capitoli di libro;

- la procedura per l'individuazione dei revisori esterni;
- la responsabilità finale in capo ai GEV della valutazione dei prodotti con l'attribuzione delle classi di merito;
- la scelta di utilizzare per la valutazione la tecnica della informed peer review<sup>9</sup>, che consiste nel tenere conto di più elementi di valutazione per la classificazione finale di merito. A seconda delle caratteristiche peculiari del GEV, gli elementi vanno dall'uso di due basi di dati per la valutazione bibliometrica, alla combinazione di valutazione peer e bibliometrica;
- l'uso della informed peer review per la valutazione delle monografie e dei capitoli di libro;
- la procedura per l'individuazione dei revisori esterni;

I lavori presentati sono oltre 118 mila, con una percentuale dei prodotti mancanti del 5,9% (6,2% contanto solo le università). Il dato è aumentato rispetto la VQR1, dove la percentuale dei prodotti mancanti è del 5,2% (4,7% per le università).

Il processo di valutazione è avvenuto seguendo, singolarmente o in combinazione, due metodologie: la prima (a) si tratta di una valutazione diretta da parte del GEV, anche utilizzando un'analisi bibliometrica (laddove possibile) basata sul numero di citazioni e su indicatori di Impact Factor della rivista ospitante il prodotto; (b) peer review, effettuata da esperti esterni indipendenti tra loro, scelti dal GEV (almeno due per prodotto), a cui è affidato il compito di esprimere una valutazione qualitativa del prodotto, in modo totalmente anonimo.

Il giudizio, come anticipato sopra, va effettuato seguendo 3 criteri:

- a) *originalità*, che va intesa come la capacità di introdurre nuove metodologie di lavoro e nuovi modi di pensare in relazione all'oggetto scientifico della ricerca, che di dovrà distinguere dai precedenti approcci aventi lo stesso oggetto;
- b) *rigore metodologico*, da intendersi come il livello al quale il prodotto presenta in modo chiaro gli obiettivi e la metodologia adottata per il perseguimento dell'oggetto della ricerca e la dimostrazione che gli obiettivi sono stati raggiunti;
- c) *impatto atteso o potenziale nella comunità scientifica di riferimento*, da intendersi come il livello al quale il prodotto ha esercitato, o è suscettibile di esercitare in futuro, un'influenza teorica e/o applicativa su tale comunità anche in base alla sua capacità di rispettare standard internazionali di qualità della ricerca.

Il risultato finale della valutazione consisteva nell'attribuzione a ciascun prodotto di una delle seguenti classi di merito e del relativo peso:

- Eccellente: il prodotto si colloca nel 10% superiore della scala di valore condivisa dalla comunità scientifica internazionale (peso 1);
- Elevato: il prodotto si colloca nel segmento 10% - 30% (peso 0,7);
- Discreto: il prodotto si colloca nel segmento 30% - 50% (peso 0,4);
- Accettabile: il prodotto si colloca nel segmento 50% - 80% (peso 0,1);
- Limitato: il prodotto si colloca nel segmento 80% - 100% (peso 0);

## **4.2 Il processo di valutazione**

### **4.2.1 La valutazione peer review**

La valutazione dei prodotti è stata effettuata dal GEV utilizzando la bibliometria e la peer review. Ogni prodotto è stato assegnato a due membri del GEV responsabili del relativo processo di valutazione. A febbraio 2015 è iniziata la procedura di valutazione tramite peer review con la costituzione di un albo di revisori ANVUR-VQR2. Partendo dall'albo di esperti revisori per la valutazione scientifica della ricerca *REPRISE* (affiliato al MIUR) è stata fatta una prima selezione in base a criteri di merito scientifico (indice h di Hirsh, numero di citazioni, produzione scientifica recente) e successivamente lo hanno integrato con un numero elevato di esperti scelti sulla base degli stessi criteri e interpellati individualmente al fine di valutarne la disponibilità a partecipare alla VQR2. I due membri GEV responsabili di ogni prodotto hanno scelto separatamente i due revisori, ponendo una particolare attenzione a evitare i conflitti d'interesse sulla base delle indicazioni contenute nei documenti sui criteri di valutazione. Complessivamente, la VQR2 ha impegnato 16.969 revisori di cui 13.546 italiani e 3.423 con affiliazione estera.

Ad ogni revisore è stato chiesto di valutare il prodotto sulla base di tre domande a risposta multipla, una per ognuno dei criteri descritti nel punto 4.1, ovvero i tre criteri di originalità, rigore metodologico e impatto atteso o potenziale. La somma dei tre punteggi era confrontata con quattro soglie per generare una classificazione finale in cinque classi. La valutazione peer richiedeva inoltre di dare un giudizio qualitativo scritto sul prodotto da valutare.

### **4.2.2 Valutazione bibliometrica**

Per quanto riguarda la valutazione bibliometrica invece, non tutte le Aree di Ricerca sono state soggette a tale valutazione, in quanto, per tipologia di prodotto da presentare, un'analisi

bibliometrica non porterebbe ad un risultato significativo. In particolare le Aree che non sono state prese in considerazione sono: Area 8a (Architettura), Area 10 (Scienze dell'antichità, filologico-letterarie e storico-artistiche), Area 11a (Scienze storiche, filosofiche e pedagogiche), Area 12 (Scienze giuridiche), Area 14 (Scienze politiche e sociali).

La valutazione bibliometrica dei GEV 2, 3, 4, 5, 6, 7, 8b, 9 e 11b ha riguardato gli articoli pubblicati su riviste indicizzate nelle basi di dati WoS e Scopus. L'ANVUR ha acquisito le informazioni bibliometriche dagli archivi WoS sia dagli archivi Scopus per gli anni 2011-2014 e per la produzione scientifica mondiale. Con l'utilizzo di entrambe le banche dati più diffuse, l'obiettivo è quello di sfruttare meglio le caratteristiche di parziale complementarità delle due basi di dati.

L'algoritmo di valutazione dei GEV 2, 3, 4, 5, 6, 7, 8b, 9 e 11b si basa sul calcolo di due indicatori per ogni prodotto: il numero di citazioni del prodotto e il fattore di impatto della rivista di pubblicazione. Alle istituzioni è stato chiesto, nella fase di conferimento dei prodotti, di indicare nella scheda prodotto la base di dati (WoS oppure Scopus) e l'indicatore di impatto preferiti (IF5Y, Article Influence Score per WoS e IPP, SJR per Scopus). Nel caso di riviste appartenenti a più di una Subject Category le istituzioni hanno espresso una preferenza, sottoposta al giudizio dei GEV per conferma. Alla coppia di valori dei due indicatori caratteristici del prodotto considerato, con regole lievemente diverse da GEV a GEV, si è associata una fra sei classi: le cinque classi della VQR2 e una sesta classe (IR) ottenuta nel caso di indicazioni divergenti dei due indicatori (ad esempio, un prodotto con elevato numero di citazioni pubblicato su una rivista con impatto molto basso o viceversa). I prodotti IR sono stati sottoposti a valutazione peer.

Il GEV1 ha adottato un algoritmo di valutazione bibliometrico lievemente diverso, che non si basa direttamente sulle Subject Categories di ISI WoS e Scopus, ma che ha identificato delle categorie di riferimento, una per ogni SSD di competenza del GEV, che integrano le Subject Categories (SC) usate in WoS e la classificazione All Science Journal Classification (ASJC) usata in Scopus. Inoltre, il GEV1 ha utilizzato, oltre alle basi di dati WoS e Scopus, e limitatamente all'indicatore d'impatto della rivista, MathSciNet dell'American Mathematical Society (MathSciNet). Per i dettagli si veda il Rapporto di Area del GEV1.

Il GEV13 ha utilizzato un algoritmo bibliometrico sensibilmente diverso dagli altri GEV bibliometrici, privilegiando la sede di pubblicazione e utilizzando il numero di citazioni per "premiare" con un salto di classe i prodotti con un numero di citazioni significativo.

Mentre nella VQR1 le autocitazioni erano incluse senza distinzioni nelle citazioni al fine di calcolare l'indicatore citazionale, nella VQR2 sono stati segnalati i casi in cui il numero di autocitazioni superava la soglia del 50% delle citazioni complessive, chiedendo in tali casi ai membri GEV responsabili della valutazione del prodotto una attenzione particolare.

#### **4.3 Risultati della valutazione della ricerca**

I risultati della valutazione della ricerca sono suddivisi per Area di Ricerca, dove per ogni Area di Ricerca le graduatorie sono suddivise per grandezza dell'ateneo. La grandezza dell'ateneo a sua volta prende un connotato diverso per ogni Area di ricerca: prendiamo come esempio l'Area 2, scienze fisiche: rientrano nella categoria Università *piccole* quelle istituzioni con un numero di prodotti attesi minore o uguale a 50; rientrano nella categoria Università medie quelle istituzioni con un numero di prodotti attesi compreso tra 51 e 140; rientrano nella categoria Università grandi quelle istituzioni con un numero di prodotti attesi superiore a 140. Per capire la differenza, la categorizzazione dell'Area 13 (Scienze economiche e statistiche) i prodotti attesi per le Università piccole sono compresi tra 5 e 100, compresi tra 101 e 250 per le Università medie e superiori a 251 per le Università grandi. Ad ottenere i risultati migliori sono le Università di Padova, Bologna e Torino. In 11 aree su 16 l'ateneo di Padova risulta tra i migliori. In particolare il Nord ha avuto dei risultati stupefacenti, che in 12 aree su 16 risultano essere i migliori (per quanto riguarda i grandi atenei) con il triangolo Veneto – Lombardia – Piemonte. In alcuni settori continuano a primeggiare accademie con vocazioni specifiche: in economia a guidare la graduatoria è la Bocconi di Milano, mentre il primato in matematica va all'Università di Pisa. Il Politecnico di Torino, Venezia Iuav e il Politecnico di Milano risultano essere le migliori in Architettura. Sorprende la prestazione di Napoli Federico II, con il primato in ingegneria civile, seguito dal Politecnico di Torino. Per scienze politiche il primo posto appartiene all'Università di Milano, seguita da Bologna e Torino. Tra le materie giuridiche invece il primato va ancora a Milano, seguita da Bologna e Torino. Infine, solamente un podio per l'Università La Sapienza di Roma in matematica e informatica, sotto Pisa e Padova. Dopo il terzetto guidato da Padova, Bologna e Torino, seguono Firenze e Milano Bicocca, ognuna delle quali si è distinta in ben 9 aree, e poi Scuola Superiore Sant'Anna di Pisa, Scuola Normale di Pisa, Statale di Milano e Politecnico di Torino, ciascuna distinta in 5 ambiti. In generale, i risultati della valutazione mostrano un certo miglioramento per la ricerca in tutta Italia, pur mantenendo un certo divario tra Nord e Sud, con il secondo in recupero.

#### 4.4 Risultati per la valutazione della ricerca nell'ateneo di Padova

L'università degli Studi di Padova risulta essere una delle migliori Università nell'analisi delle singole istituzioni: in seguito la tabella riportano le statistiche sui prodotti attesi e conferiti nelle 16 aree (fonte: ANVUR).

Area	# Prodotti attesi ( <i>n</i> )	% sul totale di area	# Prodotti conferiti	% conferiti su attesi	# Prodotti mancanti	% mancanti su attesi
1	230	4,04	201	87,39	29	12,61
2	210	5,29	206	98,10	4	1,90
3	217	4,14	210	96,77	7	3,23
4	95	4,97	94	98,95	1	1,05
5	322	3,70	309	95,96	13	4,04
6	637	3,70	622	97,65	15	2,35
7	356	6,42	355	99,72	1	0,28
8a	30	0,85	30	100,00	0	0,00
8b	94	3,32	94	100,00	0	0,00
9	434	4,53	416	95,85	18	4,15
10	254	2,79	242	95,28	12	4,72
11a	213	3,45	201	94,37	12	5,63
11b	237	10,54	222	93,67	15	6,33
12	238	2,69	228	95,80	10	4,20
13	215	2,46	213	99,07	2	0,93
14	110	3,58	99	90,00	11	10,00
<b>Totale</b>	<b>3892</b>	<b>3,80</b>	<b>3742</b>	<b>96,15</b>	<b>150</b>	<b>3,85</b>

Si può notare come la percentuale dei prodotti conferiti sui prodotti attesi è del 96,15, superiore alla media delle università. La variabilità va dal 87,39% dell'Area 1 al 100% dell'area 8a e 8b.

Segue una seconda tabella dove riporta la sintesi della valutazione della produzione scientifica dell'Università degli Studi di Padova nelle varie aree.

La tabella contiene i valori dei due indicatori di qualità media della ricerca  $I_{i,j}$  e  $R_{i,j}$  della Sezione 4.2 del Rapporto, dei due parametri necessari al loro calcolo, la posizione nella graduatoria di area rispetto ai due indicatori, sia assoluta (nell'insieme delle istituzioni omogenee), sia, in parentesi, nell'insieme delle istituzioni della stessa classe dimensionale (Piccole, Medie, Grandi), un terzo indicatore,  $X_{i,j}$ , calcolato come rapporto tra la frazione di prodotti eccellenti ed elevati dell'ateneo nell'area e la frazione di prodotti eccellenti ed elevati dell'area. Valori di  $X_{i,j}$  superiori a 1 indicano che l'ateneo ha presentato una percentuale di prodotti valutati come eccellenti ed elevati superiore alla media di area.

L'Università degli Studi di Padova è presente in tutte le sedici aree scientifiche, collocandosi tra le grandi in undici aree, tra le medie in quattro aree e tra le piccole solo nell'Area 8a.

L'indicatore  $R$  è maggiore di uno in tutte le aree, mostrando che la valutazione media è sempre superiore alla media nazionale di area. L'indicatore  $X$  è anch'esso superiore a uno



in tutte le aree, mostrando che la frazione di prodotti eccellenti ed elevati è sempre superiore alla media di area. (fonte: ANVUR)

Area	$\pi$ Prodotti A	$\pi$ Prodotti B	$\pi$ Prodotti C	$\pi$ Prodotti D	$\pi$ Prodotti E	$\pi$ Prodotti F	Somma punteggi (v)	$\pi$ Prodotti attesi (n)	Voto medio (l-v/n)	R	(n/N) x 100	IRAS1 x 100	Pos. grad. compl.	Num. istituzioni compl.	Classe dimensionale	Pos. grad. classe	Num. istituzioni classe	% prodotti A+B	X
1	85	70	31	12	3	29	147,60	230	0,64	1,06	4,04	4,29	21	59	G	2	7	67,39	1,11
2	154	29	15	6	2	4	180,55	210	0,86	1,10	5,29	5,85	11	55	G	1	9	87,14	1,10
3	122	72	13	2	0	8	177,80	217	0,82	1,09	4,14	4,53	17	56	G	3	7	89,40	1,13
4	57	20	8	9	0	1	74,60	95	0,79	1,25	4,97	6,21	5	43	G	2	9	81,05	1,28
5	161	98	37	12	1	13	245,60	322	0,76	1,17	3,70	4,34	14	62	G	2	10	80,43	1,21
6	385	144	59	27	2	20	512,10	637	0,80	1,32	3,70	4,89	6	52	G	1	13	83,05	1,38
7	148	129	51	23	4	1	261,00	356	0,73	1,23	6,42	7,89	4	40	G	1	6	77,81	1,29
8a	0	18	8	3	1	0	16,10	30	0,54	1,15	0,85	0,98	8	44	P	7	29	60,00	1,45
8b	42	28	12	8	4	0	67,20	94	0,71	1,13	3,32	3,76	15	51	M	2	7	74,47	1,17
9	213	111	56	34	1	19	316,50	434	0,73	1,13	4,53	5,12	15	63	G	1	6	74,65	1,16
10	40	125	56	18	3	12	151,70	254	0,60	1,04	2,79	2,90	25	66	M	9	23	64,96	1,08
11a	37	96	58	9	1	12	128,30	213	0,60	1,09	3,45	3,75	22	74	G	3	6	62,44	1,12
11b	104	56	30	27	4	16	157,90	237	0,67	1,25	10,54	13,15	8	55	G	1	6	67,51	1,33
12	20	100	75	26	5	12	122,60	238	0,52	1,04	2,69	2,80	33	82	M	12	24	50,42	1,08
13	101	46	35	14	11	8	148,60	215	0,69	1,50	2,46	3,69	5	82	M	1	33	68,37	1,55
14	11	35	35	15	2	12	51,00	110	0,46	1,07	3,58	3,82	23	69	G	4	7	41,82	1,10

Un' ultima analisi può essere il confronto dell'istituzione tra VQR1 e VQR2. Seguirà un grafico dove saranno riportati gli indicatori  $A_{i,j,v}$ ,  $A_{i,j,N}$  e  $B_{i,j}$ . In sintesi, un valore di  $B_{i,j}$  uguale a 2 significa un progresso dell'istituzione nella VQR2 rispetto alla VQR1, uguale a 1 una situazione di sostanziale stabilità dell'istituzione nella VQR2 rispetto alla VQR1, e uguale a 0 un peggioramento dell'istituzione nella VQR2 rispetto alla VQR1. La tabella evidenzia come l'Università degli Studi di Padova sia migliorata in nove aree, rimasta stabile in cinque e peggiorata in due. (fonte: ANVUR)

Area	Classe dimensionale	Presenza nei due esercizi VQR	$A_{i,j,v}$	$A_{i,j,N}$	$B_{i,j}$	# Istituzioni nell'area presenti nella VQR1 e nella VQR2	Posizionamento dell'Istituzione nell'ordinamento derivante dalla distribuzione di R nella VQR1	Posizionamento dell'Istituzione rispetto alla mediana della distribuzione di R nella VQR2
1	G	presente in entrambe	2	2	2	7	estremo superiore	
2	G	presente in entrambe	7	8	1	9	estremo superiore	
3	G	presente in entrambe	3	3	2	7	estremo superiore	
4	G	presente in entrambe	6	4	2	9	estremo superiore	
5	G	presente in entrambe	7	6	2	10	estremo superiore	
6	G	presente in entrambe	10	8	2	13	estremo superiore	
7	G	presente in entrambe	4	3	2	6	estremo superiore	
8a	P	presente in entrambe	16	-5	2	41	tra gli estremi	
8b	M	presente in entrambe	2	3	1	7	estremo superiore	
9	G	presente in entrambe	4	3	2	6	estremo superiore	
10	M	presente in entrambe	5	13	0	23	estremo superiore	
11a	G	presente in entrambe	0	3	0	6	estremo superiore	
11b	G	presente in entrambe	4	4	2	6	estremo superiore	
12	M	presente in entrambe	-2	-4	1	24	tra gli estremi	
13	M	presente in entrambe	31	32	1	33	estremo superiore	
14	G	presente in entrambe	0	2	1	7	estremo superiore	

## Bibliografia

ANVUR. (2017) Valutazione della Qualità e della Ricerca 2011-2014, Rapporto finale ANVUR, (disponibile da: <http://www.anvur.org/rapporto-2016/>)

ANVUR. (2017) Valutazione della Qualità e della Ricerca 2011-2014, Parte terza: Analisi delle singole istituzioni (disponibile da: <http://www.anvur.org/rapporto-2016/>)

Baccini, Alberto (2010) *Valutare la ricerca scientifica. Uso e abuso degli indicatori bibliometrici*, Bologna, Il Mulino, ISBN 978-88-15-13760-9

CIVR (2006), Linee guida per la valutazione della ricerca del CIVR, MIUR, Roma (disponibile da: [http://vtr2006.cineca.it/documenti/linee\\_guida.pdf](http://vtr2006.cineca.it/documenti/linee_guida.pdf)).

De Bellis N. (2005). La citazione bibliografica nell'epoca della sua riproducibilità tecnica: bibliometria e analisi delle citazioni dallo Science Citation Index alla Cybermetrica, (revisione del 31/05/2005. (disponibile da: <http://www.bibliotecheoggi.it/content/CITAZIONE.pdf>)

Figà Talamanca A. (2000), L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica, SINM 2000: un modello di sistema informativo nazionale per aree disciplinari. (disponibile da: <http://siba2.unile.it/sinm/4sinm/interventi/fig-talam.htm>)

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output, «Proceedings of the National Academy of Sciences of the United States of America», 102(46), 16569–16572. <http://arxiv.org/abs/physics/0508025>

Morriello, R. (2007). L'indice di Hirsch (h-index) e altri indici citazionali dopo l'impact factor: uso nella valutazione della ricerca scientifica e nelle politiche documentarie delle biblioteche, «Biblioteche oggi», 25(1), 23-32. <http://www.bibliotecheoggi.it/2007/20070102301.pdf>

Piazzini, T. (2010). Gli indicatori bibliometrici: riflessioni sparse per un uso attento e consapevole, «JLIS.it», 1(1), 63-86. <http://leo.cilea.it/index.php/jlis/article/viewFile/24/38>