

GLI EFFETTI DELLA VALUTAZIONE SULLA RICERCA ACADEMICA

Produzione della conoscenza
e problemi metodologici

a cura di Emanuela Reale

Presentazione di Mauro Palumbo

FrancoAngeli 

AV
*Associazione
Italiana
di Valutazione*

Valutazione

Collana dell'Associazione Italiana di Valutazione

L'Associazione Italiana di Valutazione (AIV) ha tra i suoi obiettivi principali la diffusione della cultura della valutazione. Per realizzare la sua mission AIV produce una rivista quadriennale, Rassegna Italiana di Valutazione (RIV) - edita da FrancoAngeli, diretta da Mita Marra e il cui Comitato editoriale è composto da Francesco Mazzeo Rinaldi e Gabriele Tomei - e promuove una collana editoriale, che qui si presenta.

Direttore scientifico: Mauro Palumbo.

Comitato scientifico editoriale: Stefano Campostrini, Domenico Patassini, Guido Pellegrini, Nicoletta Stame.

La collana si prefigge la diffusione della cultura della valutazione in Italia. Si articola in tre sezioni, cui i testi sono assegnati anche sulla base del giudizio di referee anonimi:

- *Teoria, metodologia e ricerca* comprende testi di carattere teorico e metodologico, manuali di valutazione di carattere generale o settoriale, antologie di autori italiani e stranieri.
- *Studi e ricerche* accoglie rapporti di ricerca, selezioni di contributi a Convegni, altre opere, monografiche o antologiche, che approfondiscono la valutazione all'interno di un contesto specifico.
- *Strumenti* ospita testi più brevi, dedicati a una tecnica o a un tema specifico, orientati all'utilizzo diretto da parte del fruitore, rivolti di norma a un pubblico di professionisti e operatori.

Comitato scientifico (della Collana e della RIV)

Maria Carmela Agodi	Università di Napoli "Federico II"
Gianluca Argentin	Università Cattolica del Sacro Cuore di Milano
Giovanni Bertin	Università di Venezia
Ferruccio Biolcati Rinaldi	Università di Milano
Raffaele Brancati	MET-Monitoraggio Economia e Territorio
Daniele Checchi	Università di Milano
Costantino Cipolla	Università di Bologna
Guido Corbetta	Università Bocconi
Ugo De Ambrogio	IRS, Milano
Bruno Dente	già Politecnico di Milano
Antonio Fasanella	Sapienza Università di Roma
Osvaldo Feinstein	Valutatore indipendente
Kim Forss	International evaluator
Dora Gambardella	Università di Napoli "Federico II"
Giampietro Gobo	Università di Milano
Sebastian Lemire	University of California at Los Angeles, UCLA
Andrea Lippi	Università di Firenze
Andrea Mairate	Commissione europea - Institute for European Studies
Giuseppe Moro	Università di Bari
Daniela Oliva	IRS, Bologna
Katia Papazian	Institut Polytechnique de Lausanne
Aline Pennisi	MEF
Carlo Pennisi	Università di Catania
Donatella Poliandri	INVALSI
Laura Polverari	University of Strathclyde, Glasgow
Claudio Radaelli	University of Exeter
Emanuela Reale	CNR, Roma
Maria Stella Righettini	Università di Padova
Barbara Romano	ASVAPP
Paolo Severati	INAPP
Claudio Torrigiani	Università di Genova
Giancarlo Vecchi	Politecnico di Milano
Alberto Vergani	Valutatore indipendente



OPEN ACCESS la soluzione FrancoAngeli

Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più: [Pubblica con noi](#)

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "[Informatemi](#)" per ricevere via e-mail le segnalazioni delle novità.

GLI EFFETTI DELLA VALUTAZIONE SULLA RICERCA ACADEMICA

**Produzione della conoscenza
e problemi metodologici**

**a cura di Emanuela Reale
Presentazione di Mauro Palumbo**

FrancoAngeli®

Questo volume è stato supportato dal Progetto di Ricerca di Interesse Nazionale (PRIN 2017) dal titolo “Gli effetti della valutazione sulla ricerca accademica: produzione della conoscenza e problemi metodologici”, Prot. 2017NKWYFC.

Isbn e-book Open Access: 9788835179092

Copyright © 2025 by FrancoAngeli s.r.l., Milano, Italy.

Pubblicato con licenza *Creative Commons
Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale*
(CC-BY-NC-ND 4.0).

Sono riservati i diritti per Text and Data Mining (TDM), AI training e tutte le tecnologie simili.

*L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore.
L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni
della licenza d'uso dell'opera previste e comunicate sul sito
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>*

Indice

Presentazione, di Mauro Palumbo	pag.	9
Introduzione, di Emanuela Reale	»	15
1. Le ragioni di questo volume	»	15
2. Quadro teorico e dimensioni di studio	»	18
3. Aspetti metodologici generali	»	20
4. I contributi	»	22
1. La Valutazione della Qualità della Ricerca e i suoi effetti sul lavoro accademico: differenze disciplinari e di contesto, di Antonio Zinilli, Andrea Orazio Spinello, Emanuela Reale	»	25
1. Introduzione	»	25
2. Letteratura di riferimento e ipotesi di lavoro	»	27
3. Dati e metodi	»	30
4. La survey	»	33
5. Risultati: analisi descrittiva	»	35
6. Risultati: modello LOGIT	»	39
7. Conclusioni	»	46
2. Le scelte collegate alla carriera dei giovani ricercatori: storie di vita e narrazioni sugli effetti della valutazione, di Serena Fabrizio, Ugo Finardi, Lucio Morettini, Emanuela Reale	»	50
1. Introduzione	»	50
2. Contesto di riferimento dello studio e obiettivi conoscitivi	»	51

3. Dimensione narrativa degli effetti della valutazione: Storie di Vita “informate”	pag. 54
4. Dalla definizione delle dimensioni analitiche ai “temi di conversazione”: il protocollo delle Storie di Vita	» 56
5. Principali risultati	» 60
6. Discussione e conclusioni	» 70
3. La formazione dell'agenda di ricerca e gli effetti della valutazione esterna, di Valentina Carazzolo	73
1. Introduzione	» 73
2. Valutazione della ricerca e professione accademica nel contesto italiano	» 74
3. Sulla <i>research agenda</i> dei ricercatori: concetti chiave, domande di ricerca e metodologia	» 81
3.1. Concetti chiave e domande di ricerca	» 81
3.2. Metodologia	» 83
4. Risultati della ricerca	» 87
4.1. Risultati della <i>survey</i>	» 88
4.2. Risultati delle interviste informate	» 91
5. Conclusioni	» 95
4. Collaborazioni di ricerca e mobilità degli accademici: sono visibili effetti della valutazione?, di Andrea Orazio Spinello, Antonio Zinilli, Emanuela Reale	96
1. Introduzione	» 96
2. Quadro teorico sulle dimensioni di studio	» 97
2.1. Collaborazioni di ricerca e determinanti individuali	» 98
2.2. Motivazioni per la mobilità internazionale degli accademici	» 99
3. Obiettivo e metodo della ricerca	» 101
3.1. Rilevazione del dato	» 101
3.2. Modalità d'analisi	» 102
4. I risultati della <i>survey</i>	» 103
4.1. Collaborazioni di ricerca	» 104
4.2. Mobilità internazionale	» 112
5. Riflessioni conclusive	» 118
5. Parliamo tutti la stessa lingua? Riflessioni sulla valutazione della ricerca e i suoi meccanismi, di Eliana Alessandra Minelli, Maria Rucsandra Stan	121
1. Una sfida ontologica	» 121

2. Evoluzione del concetto valutativo e relative prassi	pag. 123
3. Il dibattito accademico	» 126
3.1. Prima componente principale e suo sviluppo	» 129
3.2. Seconda componente principale e suo sviluppo	» 133
4. Uno sguardo d'insieme	» 135
5. Produttività: ma a quale prezzo?	» 136
5.1. Metodo	» 137
5.2. Evidenze emerse dallo studio del settore Organizzazione Aziendale	» 139
6. Conclusioni	» 145
 6. Effetti della macchina editoriale sullo sviluppo delle discipline accademiche: un approccio neo-istituzionalista, di Maria Rucsandra Stan, Eliana Alessandra Minelli	 » 147
1. Università di massa e produttività accademica	» 147
2. Lo sviluppo del sistema delle pubblicazioni	» 149
2.1. La <i>peer review</i> nella letteratura	» 151
2.2. Un approccio neo-istituzionalista per studiare il ruolo degli <i>editor</i>	» 152
2.3. La metodologia e il campione	» 154
2.4. I risultati dell'analisi	» 156
3. Le implicazioni delle logiche sottese alla revisione dei pari nell'avanzamento delle carriere accademiche	» 158
4. Logiche istituzionali e finanziamento nella ricerca scientifica	» 162
4.1. Metodologia, raccolta dati e individuazione delle logiche sottese	» 165
4.2. Processo di ricerca e influenza delle logiche istituzionali	» 168
5. Conclusioni	» 172
 7. La risposta delle università italiane all'introduzione della Valutazione della Qualità della Ricerca, VQR 2004-2010, di Giovanni Abramo, Ciriaco Andrea D'Angelo, Flavia Di Costa	 » 176
1. Introduzione	» 176
2. Analisi della letteratura	» 179
3. Dati e metodologia	» 181
3.1. Dati	» 181
3.2. Misurare la produttività della ricerca	» 182
4. Risultati dell'analisi	» 186
5. Conclusioni	» 198

8. L'efficacia del reclutamento dei docenti nelle università italiane: gli effetti dell'Abilitazione Scientifica Nazionale , di <i>Giovanni Abramo, Ciriaco Andrea D'Angelo</i>	pag. 201
1. Introduzione	» 201
2. Discriminazione nel reclutamento dei docenti	» 203
3. Il framework di riferimento determinato dalle policy recenti	» 205
4. Specifiche metodologiche	» 207
4.1. Misurare la performance della ricerca	» 208
4.2. Misurare l'efficacia del reclutamento	» 210
4.3. Il dataset di analisi	» 212
5. Risultati	» 214
5.1. Performance di ricerca dei reclutati	» 214
5.2. Efficacia del reclutamento	» 217
6. Conclusioni	» 221
9. Gli effetti dell'Abilitazione Scientifica Nazionale sulla propensione degli accademici ad auto-citarsi , di <i>Giovanni Abramo, Ciriaco Andrea D'Angelo, Leonardo Grilli</i>	» 226
1. Introduzione	» 226
2. Comportamento auto-citazionale e problemi di misurazione	» 229
3. La letteratura sul comportamento auto-citazionale	» 232
3.1. Fattori che influenzano i tassi di auto-citazione	» 234
3.2. Effetti della valutazione bibliometrica sulle auto-citazioni	» 235
4. Abilitazione Scientifica Nazionale e “incentivi” nei settori bibliometrici	» 237
5. Ipotesi di ricerca e specifiche metodologiche	» 239
5.1. Il dataset di analisi	» 240
5.2. Modello statistico	» 243
6. Risultati	» 247
6.1. Statistiche descrittive	» 247
6.2. Inferenze dal modello statistico	» 249
6.3. Diffusione del cambiamento del comportamento auto-citazionale	» 254
7. Conclusioni	» 254
Conclusioni , di <i>Emanuela Reale</i>	» 261
Riferimenti bibliografici	» 269
Gli Autori	» 309

Presentazione

di *Mauro Palumbo*

Questo volume colma una lacuna importante nell'ambito, peraltro densamente popolato, delle pubblicazioni italiane dedicate alla valutazione della ricerca e del sistema universitario italiano. Riporta infatti i risultati di una ricerca decisamente ampia, condotta da équipe qualificate e con uso combinato di approcci qualitativi e quantitativi; dunque riflessioni ancorate alle evidenze empiriche, che in questo ambito sono spesso sopravanzate da considerazioni derivanti da presupposizioni teoriche o valoriali, assolutamente legittime, ma non sempre utili per l'avanzamento della conoscenza scientifica¹.

D'altro canto, una ricerca empirica ampia e articolata, come quella di cui si dà conto in queste pagine, richiede risorse importanti, attribuite, in questo caso, da un Progetto di Ricerca d'Interesse Nazionale (PRIN). E che ci sia un interesse nazionale su questo tema è del tutto evidente, in quanto la Valutazione della Qualità della Ricerca e l'Abilitazione Scientifica Nazionale – i due ambiti privilegiati dalle analisi condotte – incidono fortemente sul futuro degli Atenei e sulla qualità del lavoro accademico, nonché sui meccanismi di reclutamento dei giovani ricercatori. Come emerge da diversi spunti della ricerca, l'incidenza è riconducibile ad almeno tre dimensioni, tra loro connesse. La prima e spesso la più citata (anche dal sottoscritto) è la funzione premiale della valutazione, stabilita anche normativamente. È infatti proprio l'articolo 1 della legge 240/2010 che indica, nei commi 4, 5 6, la necessità di valutare a quella di premiare, secondo una logica seguita anche in altri Paesi (in primo luogo dal Regno Unito); una funzione già indicata, peraltro, nel DL 262/2006, istitutivo dell'ANVUR, che al 2° comma già affermava “I risultati delle attività di

1. Tra le poche ricerche empiriche in materia si ricorda il (relativamente) recente volume di Fontana e Valentini (a cura di, 2020), che utilizza non solo dati secondari, ma anche rilevazioni empiriche dirette.

valutazione dell’Agenzia costituiscono criterio di riferimento per l’allocazione dei finanziamenti statali alle università e agli enti di ricerca”.

La seconda concerne la connessione tra i risultati della valutazione della ricerca e le carriere dei ricercatori accademici, dal momento che l’Abilitazione Scientifica Nazionale (ASN), istituita anch’essa dalla legge 240/2010 (art. 16), affida all’ANVUR la definizione di importanti elementi per l’ammissione e la valutazione dei candidati.

Entrambe le valutazioni hanno polarizzato l’attenzione degli studiosi, ma, come premesso, non sono molte le evidenze empiriche che hanno consentito di fare luce su quanto le ipotesi soggiacenti alle norme fossero corroborate dai fatti² e quanti effetti inintenzionali o non dichiarati sarebbero stati prodotti da queste due innovazioni (oltre che da quelle connesse delle procedure di accreditamento del sistema AVA e di valutazione della Terza Missione). Questo lavoro consente di dare almeno una parte di risposte ai temi centrali nel dibattito: la valutazione con funzioni premiali ha prodotto miglioramenti? E se sì, li ha prodotti secondo le logiche implicite nella volontà del legislatore o ha seguito altre strade? Ad esempio, ha prevalso l’effetto “strumentale”, che corrisponde a un’ipotesi di homo economicus (una buona valutazione produce risorse aggiuntive), ovvero una leva “espressiva”, legata alla reputazione dei ricercatori che emerge da una buona valutazione? E quale che sia il “meccanismo” che ha operato (nel senso assegnato al termine da Elster, Coleman e naturalmente Pawson), questo ha funzionato allo stesso modo per giovani e meno giovani in carriera, per sedi “marginali” o centrali nel sistema? Il reclutamento dei ricercatori sulla base di criteri di qualità almeno in parte sottratti all’arbitrio delle Commissioni giudicatrici ha prodotto un miglioramento della qualità dei loro prodotti scientifici (e per tal via anche un miglioramento del sistema)? Oppure il peso del cosiddetto “favor loci” ha creato profezie che si autoadempiono (facendo promuovere più facilmente chi aveva la possibilità di essere chiamato nella propria sede di lavoro)?

Sono solo i principali interrogativi cui i risultati della ricerca pubblicata consentono di fornire prime interessanti risposte, che lascio al lettore di scoprire addentrandosi nel volume. Anche se voglio anticipare che alcune conclusioni controintuitive andrebbero enfatizzate fin da subito, per esempio quella secondo la quale gli Atenei peggio piazzati nella VQR precedente hanno segnato i miglioramenti più consistenti, in relazione a quelli dei “campioni” che dovevano sfidare: una evidenza empirica contrastante con argomentazioni da me sostenute anche in queste pagine!

2. Un tentativo in ambito locale viene operato da Colarusso e Giancola, 2020.

In sede di presentazione di un lavoro che contribuisce certamente all'avanzamento della conoscenza in un ambito così rilevante, mi limito a due considerazioni.

La prima riguarda il metodo: la ricerca presentata consta in realtà di una pluralità di ricerche, condotte con l'uso di approcci sia qualitativi che quantitativi, che vanno dalle storie di vita a sofisticate analisi statistiche di vaste basi dati, a rilevazioni dirette degli effetti sui ricercatori della VQR e dell'ASN. Ricerche presentate con apprezzabile rigore espositivo, evidenziato dall'analisi della letteratura che precede ogni ambito esplorato e dal rigore con cui sono prima presentati e poi discussi i risultati, riferendoli agli interrogativi di ricerca dai quali gli estensori erano partiti. Anche l'uso dell'approccio contributivo sembra vincente in un ambito in cui le variabili che influenzano la qualità della ricerca e delle carriere sono tali e tante da scoraggiare qualsiasi semplificazione delle ipotesi iniziali. Dall'altro lato, la compresenza di analisi su dati "oggettivi" e di rilevazioni del punto di vista degli attori (storie di vita e survey) fornisce un panorama davvero ricco del tema trattato.

Mi permetto di aggiungere che uno dei pregi consentiti dal pluralismo metodologico della ricerca è costituito dall'aver potuto svelare alcuni dei "meccanismi" soggiacenti ai comportamenti attivati dai ricercatori e dalle ricercatrici in risposta alle policies sottese a VQR e ASN. Il termine "meccanismi", com'è noto nella comunità dei valutatori, è stato introdotto da Pawson e Tilley ormai quasi 30 anni fa (1997), ed ha avuto da allora vasto eco e soprattutto vasto impiego in ambito valutativo. Secondo questa impostazione, ogni policy che utilizza un "meccanismo" (ossia una ipotesi causale secondo la quale facendo A – attivando un servizio, un incentivo, un disincentivo – si dovrebbe ottenere B in termini di razione comportamentale dell'attore "bersaglio") ottiene uno specifico effetto in base al contesto nel quale il meccanismo è chiamato ad operare³. E infatti si nota che l'effetto apparentemente controiduttivo per cui ricercatori o sedi "svantaggiate" nella VQR precedente hanno ottenuto un miglioramento superiore a quello delle sedi in allora meglio piazzate risulta controiduttivo se si ragiona in termini di azione strumentale rispetto all'obiettivo di conseguire maggiori finanziamenti nazionali (perché il divario è tale da non permettere un significativo riposizionamento), ma è invece comprensibile se si ragiona in termini reputazionali, ossia con l'obiettivo di "non sfigurare" a fronte di colleghi o Atenei considerati eccellenti. Weberianamente, po-

3. Si tratta della nota formula "M+C=O", dove M sta per meccanismo, C per contesto e O per outcome, ossia risultato o effetto.

tremmo dire che l'agire razionale rispetto allo scopo (la premialità) viene sopravanzato dall'agire razionale rispetto al valore (il prestigio, la reputazione⁴). Vorrei anzi aggiungere che il richiamo a Pawson e Tilley non ci deve far dimenticare la lezione di studiosi come Coleman (1990/2005) e Elster (2007/10), che hanno ampiamente sviluppato una concezione di spiegazione sociologica ancorata ai meccanismi, sottolineando il carattere causale della spiegazione sociologica⁵. Una causalità generativa e non lineare, dal momento che l'operato dei meccanismi è sempre filtrato dal contesto e non può quindi mai assumere caratteristiche costanti (e dunque pienamente prevedibili) che sarebbero proprie della causalità lineare (cfr. Palumbo, 2003). Una posizione già ben affermata da Weber e implicita nelle “teorie di medio raggio” di Merton, che tuttavia è stata persa per strada quando si è cercato di inseguire le certezze delle scienze esatte scimmiettandone le semplificazioni modellistiche (Palumbo, 2020).

La seconda riguarda la funzione della valutazione. Il volume mostra che alcune delle (criticate) ipotesi soggiacenti alle politiche che hanno spinto sulla valutazione con finalità premiali sono state confermate, altre meno, altre hanno portato a risultati controveintivi rispetto alle attese di miglioramento. Ci sono dunque evidenze favorevoli sia ai critici che ai sostenitori dell'utilità della funzione premiale della valutazione, al netto del rilevantissimo dibattito attorno al “come” questa viene fatta: basti pensare alle molteplici spinte in favore della Open Science e di un maggior – o esclusivo – uso della *peer review* a fronte degli indicatori bibliometrici; quanto poi CoARA (2022) sia una moda passeggera o una svolta decisiva nel dibattito tra bibliometria e *peer review* sarà solo il futuro a dircelo. Emerge tuttavia, e questo mi pare un aspetto assai importante, che vincolare la valutazione ad una funzione premiale in chiave esclusiva di accountability e di logica “sommativa” sia davvero limitativo e avvilente. Limitativo per le molte dimensioni che l’accountability può avere, prendendo a riferimento non solo l’autorità politica o la comunità scientifica, ma anche l’ampia gamma di stakeholder verso i quali gli Atenei detengono responsabilità sociale:

4. Mi piace pensare che un agire sociale orientato alla premialità (azione razionale rispetto allo scopo) rientra perfettamente nella logica di funzionamento del mercato, sia per Weber che per i cultori nel New Public Management ispiratori della riforma del 2010, mentre un agire sociale che mette al centro la reputazione del ricercatore (per Weber una forma di agire razionale rispetto ai valori) sia riconducibile alla dimensione dello status e del prestigio, che porta anche a svolgere azioni non prevedibili nella pura logica dell’agire strumentale.

5. Afferma Elster che “i meccanismi sono quei *modelli causali osservabili di frequente e facilmente riconoscibili, avviati sotto condizioni generalmente non note o seguiti da conseguenze indeterminate*” (corsivo nel testo, 2010, p. 11).

non solo studenti e famiglie, ma anche territorio, imprese, istituzioni locali. Un aspetto bene sottolineato da Gambardella e Lumino (2020), che proprio a fronte di una pluralità di soggetti rispetto ai quali viene svolta la funzione di accountability della valutazione identificano diverse possibili reazioni comportamentali, che leggono in chiave di learning. La tipologia che propongono, di sapore mertoniano, *full adhesion, compliance, critical agency, refusal*, oltre a rispecchiare molte delle posizioni nel campo del dibattito scientifico, rende ragione delle diverse reazioni dei colleghi a VQR e ASN, documentate anche dalla ricerca presentata in questa sede. Mi soffermo in particolare sulla *compliance*, che secondo le autrici “entails here a sort of bureaucratic ritualism aimed at feeding the demands of the evaluation system, ‘corresponding to the numbers’ or targets specified by ANVUR” (p. 159). Ancora un richiamo a Merton che appare decisamente opportuno, perché configura uno dei rischi più rilevanti che corre oggi la valutazione universitaria: tradursi in una logica dell’adempimento a regole imposte dall’alto cui ottemperare senza convinzione e senza la speranza di un utilizzo fruttuoso. Ma proprio richiamando la pluralità di forme di accountability della valutazione della ricerca, rispetto a pubblici diversi e dunque anche con apprendimenti diversi da parte dei valutati, riteniamo che la funzione di learning della valutazione non debba limitarsi a quella di favorire comportamenti di compliance od opportunistic (come può accadere forzando la funzione premiale e l’accountability verso l’alto), ma possa anche (invece?) incentivare la capacità delle comunità scientifiche e degli Atenei di assolvere al meglio le missioni che sono loro assegnate e di rispondere alle esigenze di sviluppo economico, sociale e civile del Paese. Funzioni legate ai valori fondanti delle Università e che la valutazione deve poter favorire, uscendo da logiche di più breve respiro che possono immiserirne il contributo e vanificarne gli effetti.

Riferimenti bibliografici

- CoARA (Coalition for Advancing Research Assessment) (2022), *Agreement on reforming research assessment*, https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf.
- Colarusso, S., & Giancola, O. (2020). *Università e nuove forme di valutazione. Strategie individuali, produzione scientifica, effetti istituzionali*. Roma: Sapienza Università Editrice.
- Coleman, J. (2005). *Fondamenti di teoria sociale*. Bologna: Il Mulino (ed. or. 1990).
- Elster, J. (2010). *La spiegazione del comportamento sociale*. Il Mulino: Bologna (ed or. 2007).

- Fontana, R., & Valentini, E. (a cura di) (2020). *Conseguenze della valutazione. Idee e pratiche dei docenti universitari nelle scienze sociali*. Milano: Angeli.
- Gambardella, D., & Lumino, R. (2020). Re-framing accountability and learning through evaluation: Insights from the Italian higher education evaluation system. *Evaluation*, 26(2), pp. 147-165.
- Palumbo, M. (2003). Il piacere della spiegazione senza l'incubo della causalità. *Sociologia e Ricerca Sociale*, XXIII, 68/69, pp. 180-195.
- Palumbo, M. (2020). Sociologia e Ingegneria gestionale di fronte alla valutazione. Convergenze parallele? *Sociologia Italiana – AIS Journal of Sociology*, 15, pp. 165-180.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.

Introduzione

di *Emanuela Reale*

1. Le ragioni di questo volume

Il volume presenta alcuni risultati derivanti dal Progetto PRIN su *Gli effetti della valutazione sulla ricerca accademica: produzione della conoscenza e problemi metodologici*¹.

L'obiettivo del progetto è stato investigare gli effetti della valutazione della ricerca realizzata dal governo, attraverso un esercizio esteso di valutazione, sul lavoro scientifico, sulla diffusione dei risultati prodotti e sulla qualità del medesimo, prendendo come unità di analisi gli accademici che lavorano nelle università italiane.

La valutazione presenta diverse categorie di effetti potenziali riassumibili in tre gruppi principali: a) effetti sulla *governance*, ossia sulle relazioni gerarchiche tra attori, con la possibilità di far fronte a una verticalizzazione della distribuzione di potere; b) effetti sulla *comunità scientifica*, in particolare sulla creazione di nuove élite e sulla riconfigurazione del lavoro accademico; c) effetti sull'*organizzazione*, sulla gestione e sulle prestazioni delle università attraverso la razionalizzazione delle risorse disponibili. Nonostante molti lavori siano diretti ad approfondire *governance* e gestione a livello istituzionale e di sistema, i cambiamenti prodotti dalla valutazione sul lavoro accademico sono ancora poco esplorati (Fontana e Valentini, 2020). Il progetto ha quindi cercato di contribuire a colmare questa lacuna.

Le domande di ricerca generali che hanno guidato i lavori dei team coinvolti nel progetto sono: quali cambiamenti può produrre la valutazione esterna della ricerca sulle modalità di ricerca e sulle strategie individuali? In che modo è possibile capire gli effetti della valutazione a livello indivi-

1. PRIN2017 2017NKWYFC.

duale, utilizzando approcci misti che combinano dati quantitativi bibliometrici e non bibliometrici e dati qualitativi derivanti da interviste e storie di vita?

Il nostro punto di partenza per comprendere gli effetti prodotti dalla valutazione è la constatazione che il potere di condizionamento della valutazione è legato ai suoi usi specifici. La letteratura (Molas-Gallart, 2012) evidenzia tre usi potenziali: i) usi distributivi, principalmente riferiti all'assegnazione delle risorse disponibili a diversi beneficiari, sulla base della valutazione delle prestazioni o del merito degli individui e dei gruppi (il cd. finanziamento basato sulla performance); ii) usi di miglioramento, dove l'attenzione è rivolta alle lezioni che possono essere apprese dall'esperienza passata, usando la valutazione per chiarire le ragioni e le spiegazioni di determinati effetti; iii) controllo sull'utilizzo delle risorse, per esaminare come individui, organizzazioni o gruppi svolgono le attività programmate e la misura in cui hanno raggiunto i risultati previsti.

Il focus della nostra indagine è sulla valutazione della ricerca implementata dal governo, diretta a giudicare fino a che punto un obiettivo specifico o una prestazione assegnata a un'organizzazione siano stati raggiunti. Essa è utilizzata prevalentemente per scopi distributivi e di controllo. In Italia il governo ha implementato negli anni 2000 due strumenti per la valutazione della ricerca, che svolgono un ruolo preminente e meritano quindi un'attenzione particolare:

- a) la Valutazione della Qualità della Ricerca (VQR, istituita nel 2001), che affronta il rendimento della ricerca delle università incentrandosi principalmente sui suoi prodotti, i cui risultati determinano l'allocazione dei finanziamenti istituzionali delle università (FFO, Fondo di finanziamento ordinario);
- b) l'Abilitazione Scientifica Nazionale (ASN, istituita nel 2012), esame che consente agli accademici di candidarsi per posizioni permanenti di professore associato e professore ordinario nelle università.

Entrambi gli strumenti sono regolamentati da atti del Ministero dell'Università e della Ricerca - MUR e implementati dall'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca - ANVUR.

La VQR è gestita in maniera centralizzata dall'ANVUR ed è eseguita periodicamente utilizzando processi standardizzati. Si suppone che abbia effetti indiretti sull'attività di ricerca dei professori e dei giovani ricercatori, dovuti ai cambiamenti apportati all'interno delle università per adempiere alle regole e agli standard di valutazione, o dagli usi dei risultati che vanno oltre la loro destinazione originale: per esempio, l'uso del *rating* VQR delle pubblicazioni per le decisioni relative alla partecipazione degli accademici come coordinatori o insegnanti ai corsi di dottorato e l'individuazione dei Dipartimenti di Eccellenza.

L'ASN è anch'essa gestita in maniera centralizzata ed eseguita periodicamente dal MUR per consentire agli accademici di richiedere una attestazione di idoneità alle posizioni di professore associato o ordinario. Infatti, il superamento dell'esame ASN è un prerequisito essenziale per essere reclutati come professore associato o ordinario. Il giudizio sull'abilitazione viene formulato da accademici e si basa principalmente sulla valutazione delle pubblicazioni scientifiche presentate; tuttavia, la presenza di soglie minime di prodotti pubblicati in riviste selezionate (oltre a libri e capitoli di libri soprattutto per settori come le scienze sociali e umane) influisce sensibilmente sulla possibilità di superare l'esame.

Nonostante le differenze tra i due strumenti, sia VQR che ASN prevedono che la qualità dei risultati della ricerca degli accademici sia effettuata sulla base di una valutazione dei pari (*peer review* informata nella VQR e giudizio dei commissari nell'ASN) e di un utilizzo esteso di indicatori bibliometrici. Nei settori disciplinari (cd. settori non-bibliometrici), dove l'uso della bibliometria non è rappresentativo delle pubblicazioni del settore, è stata costruita una classificazione delle riviste nazionali e internazionali (la cosiddetta lista delle riviste di classe "A") basata sull'analisi delle riviste effettuata da una specifica commissione e periodicamente aggiornata. Pertanto, entrambi gli strumenti indicano agli accademici quali sono le caratteristiche di qualità di un risultato scientifico per ottenere il riconoscimento nazionale istituzionale e per accedere a una posizione permanente in accademia, generando in linea di principio una spinta verso l'adeguamento alle regole e ai criteri selezionati.

Il lavoro svolto nel progetto ha esaminato una serie di effetti concreti, quelli che la letteratura indica come positivi o problematici per la produzione di risultati, ovvero: i) le decisioni che riguardano il contenuto delle attività di ricerca (identificazione di problemi scientifici rilevanti, metodi usati, prospettiva a breve o a lungo termine, interdisciplinarietà, copertura, ecc.); ii) la modalità di codifica delle conoscenze prodotte facendo attività di ricerca, prestando attenzione alla variazione delle strategie e delle pratiche di pubblicazione (tipo di pubblicazione, lingua, produttività, tempestività); iii) i rapporti tra ricercatori all'inizio della carriera e *senior* che riguardano l'agenda di ricerca e le modalità di pubblicazione (eliminando o attenuando la condizione di dipendenza scientifica nei rapporti di autorità, Louvel, 2010); iv) la costruzione delle reti internazionali di collaborazione, mobilità e ricerca (Hessels e van Lente, 2008).

Ci aspettiamo che la valutazione della ricerca abbia effetti diversi sul lavoro accademico e sui modi di produzione della conoscenza, che dipendono in gran parte dal contesto istituzionale (tipo di università nella quale l'individuo lavora) e geografico (localizzazione dell'ateneo in cui operano

gli accademici), dai criteri elaborati all'interno delle comunità epistemiche di appartenenza per la valutazione della qualità del lavoro scientifico, dalla fase di carriera (ricercatori *junior* o *senior*), dal genere e dalla tipologia del nucleo familiare di riferimento. La domanda aperta è quale sia un buon bilanciamento di vantaggi e svantaggi per gli accademici; quindi, in quali condizioni la valutazione può essere un fattore di valorizzazione o marginalizzazione, migliorando o ostacolando la qualità della ricerca e generando o minimizzando le disuguaglianze, e come gli effetti possono essere misurati al fine di evitare conseguenze indesiderate.

2. Quadro teorico e dimensioni di studio

Lo sviluppo della valutazione della ricerca nei paesi europei è stato spesso guidato da riforme ispirate ai principi neoliberali, che miravano a incrementare contemporaneamente le capacità di governo dei responsabili politici e l'autonomia decisionale delle università, la ridistribuzione dei poteri e la crescente capacità delle organizzazioni di pianificare, controllare e misurare i risultati (Neave, 1991; Whitley e Gläser, 2007; Reale e Marini, 2017; Whitley, Gläser e Engwall, 2010). In questo contesto, l'attuazione della valutazione nei paesi europei è stata principalmente correlata al livello di autonomia concesso alle istituzioni e al tipo di sistema di finanziamento competitivo nei diversi paesi (Whitley e Gläser, 2007). In Italia, la debolezza nella progettazione delle politiche e la scarsa attuazione hanno portato a conseguenze indesiderate del processo di riforma e a conflitti legati all'introduzione della valutazione della ricerca (Capano *et al.*, 2016; Rebora e Turri, 2013; Reale e Seeber, 2013; Minelli *et al.*, 2015).

Diversi articoli in letteratura hanno affrontato i problemi dei cambiamenti della professione accademica a causa della trasformazione dei contesti esterni ed istituzionali, dell'emergere di nuovi incarichi e nuove sfide, di modelli di ricerca internazionalizzati e globalizzati, che spingono verso approcci interdisciplinari e aumentano sia le collaborazioni che la concorrenza (Altbach e Lewis, 1996; Moscati, 2001; Rostan, 2011). Tuttavia, poca attenzione è stata dedicata agli effetti della valutazione. Negli ultimi anni sono state raccolte evidenze empiriche sulle profonde trasformazioni della professione accademica tramite un'indagine dedicata a livello europeo (Teichler e Hohle, 2013, Machado-Taylor *et al.*, 2017), che ha affrontato solo in parte la valutazione dell'insegnamento e della ricerca. I risultati hanno evidenziato che gli accademici hanno diverse percezioni su come la valutazione della ricerca influisce sul loro lavoro, generalmente influenzate dalla *governance* nazionale e dal contesto di valutazione. In questa ricerca, l'Ita-

lia è tra i Paesi che mostra la maggiore preoccupazione sulla percezione di uno stile di gestione *top-down* che riguarda la *governance* e la valutazione (Campbell, 2013).

Altri studi si concentrano su come la valutazione esterna influisca sulla ricerca accademica. Alcuni suggeriscono che la valutazione non possa avere grandi effetti in termini di legittimazione e creazione di prestigio tra gli attori a causa delle complessità della professione accademica e delle limitate possibilità di influenzare i programmi di ricerca dei singoli (Whitley e Gläser, 2007); altri hanno mostrato evidenze positive e negative sull'impatto della valutazione sui programmi di ricerca e sui modelli di pubblicazione in diversi campi (Shalley, 1995; Osuna *et al.*, 2011; Abramo e D'Angelo, 2015a; Bloch *et al.*, 2015; Karlsson, 2017; Santos e Horta, 2018; Fontana e Valentini, 2020). allo stesso modo, altri studiosi hanno dimostrato che gli accademici tendono a criticare le regole del finanziamento competitivo e della valutazione, ma contemporaneamente sono attenti a tenerle in considerazione, reagendo secondo le esigenze del processo per ottimizzare il risultato (Lucas, 2006; Musselin, 2013).

Giova ricordare, almeno in parte, la letteratura che illustra gli effetti prodotti dall'allocazione di finanziamenti istituzionali per la ricerca e lo sviluppo (R&S) basata sulla performance, quindi finanziamenti destinati a sostenere le attività principali delle università. Le iniziative nazionali di valutazione della ricerca contribuiscono a questa allocazione attraverso diverse combinazioni di indicatori e *peer review* per migliorare le prestazioni del sistema di R&S e della concorrenza (Aagaard *et al.*, 2015; Bloch *et al.*, 2016; Kim e Bak, 2016). Altri studiosi suggeriscono invece la presenza di effetti conservativi dell'allocazione basata sulla performance e una selezione avversa causata da pregiudizi che influiscono sulla revisione e sugli indicatori tra pari, che favorirebbero le attività di ricerca in linea con gli approcci tradizionali (Zacharevicz *et al.*, 2019). Altri hanno sottolineato la presenza di attriti tra finanziamenti competitivi e pratiche di ricerca delle comunità di studiosi (Gläser e Laudel, 2016), ulteriormente esacerbati dalla riduzione del livello delle risorse pubbliche stanziate per R&S, che portano a conseguenze non intenzionali e problemi di equità (Hicks, 2012; Schneider *et al.*, 2016). I problemi legati agli effetti derivanti da finanziamenti basati sui risultati influenzano anche l'Italia, a causa della crescente importanza dei risultati VQR nell'assegnazione di finanziamenti istituzionali alle università (Geuna e Piolatto, 2016; Reale, 2017; Zinilli *et al.*, 2020). Gli argomenti portati a sostegno della bontà o della pericolosità degli effetti prodotti sono molti e utilizzano dati diversi spesso di tipo bibliometrico. Si segnalano, solo a titolo di esempio di visioni diverse, le analisi che individuano come effetto della VQR l'aumento della qualità media dell'*output* delle università che non si accom-

pagna però a un innalzamento della qualità medesima (documentato da un aumento dell'*output* su riviste di altissimo prestigio – Checchi *et al.*, 2019; Bratti *et al.*, 2021), e le analisi che segnalano la presenza di comportamenti opportunistici degli accademici per migliorare le proprie performance nelle pubblicazioni indicizzate (Baccini *et al.*, 2019).

Riassumendo, la valutazione della ricerca può avere conseguenze positive e negative a livello individuale. Per quanto riguarda le prime, la letteratura enfatizza una crescente attenzione per la qualità e la produttività, una preferenza per le riviste con migliore classificazione e nuove possibilità per i ricercatori di impegnarsi in collaborazioni internazionali. Gli aspetti negativi riguardano principalmente la preferenza per i temi di ricerca a breve termine, l'attenzione al *mainstream* per avere maggiori possibilità di pubblicazione, il basso livello di novità nel contenuto delle pubblicazioni, l'aumento di comportamenti opportunistici per aumentare il numero di output (p.e. *salami publications*)².

3. Aspetti metodologici generali

Il volume presenta studi quasi sempre caratterizzati da un approccio interdisciplinare che unisce diverse competenze nelle scienze sociali (sociologia, economia e studi organizzativi), nonché competenze in ingegneria gestionale. Inoltre, sono utilizzati diversi strumenti metodologici qualitativi e quantitativi:

- *interviste* a un campione di studiosi appartenenti a diverse università, aree di studio, genere, età e regioni geografiche. Sono state svolte interviste semi-strutturate “scientificamente informate” (Laudel e Gläser, 2007), per approfondire gli effetti derivanti dalla valutazione della ricerca nelle principali esperienze. Sono inoltre state raccolte *storie di vita quotidiana* dei ricercatori a inizio carriera (ricercatori *junior* e post-doc) per approfondire il processo attraverso il quale le informazioni provenienti dalla valutazione vengono introdotte nella loro vita quotidiana, influenzando la produzione della conoscenza;
- *survey*, di cui una a livello nazionale per raccogliere dati sugli effetti della valutazione della ricerca, caratteristiche dei rispondenti, contesti

2. La pratica che va sotto il nome di ‘salami publication’ è caratterizzata dalla distribuzione dei risultati di uno studio su più articoli di quanto sia necessario utilizzando il criterio dell’unità minima pubblicabile. Il risultato è la moltiplicazione di pubblicazioni scientifiche per ciascun risultato, che condividono la stessa popolazione di riferimento, metodi molto simili e domande di ricerca completamente correlate.

geografici e istituzionali. L'indagine nazionale consente una generalizzazione dei risultati; è stata effettuata attraverso un questionario on-line diretto a un campione casuale stratificato di accademici, rappresentante della popolazione nelle università italiane;

- *analisi bibliometriche* sulla produzione scientifica di accademici italiani utilizzando i dati *Web of Science – Core Collection of Clarivate Analytics*. L'analisi ha preso in considerazione tre periodi di cinque anni ciascuno per cogliere gli effetti sia della prima VQR (2004-2010) sia dell'ASN (il cui primo ciclo è stato lanciato nel 2012). Il primo periodo è il punto di partenza nell'asse temporale, mentre il secondo rappresenta il campo di osservazione degli effetti dei processi di valutazione progettati e implementati principalmente nel secondo quinquennio

La ricerca sul tema degli effetti prodotti dalla valutazione ha tre principali limiti metodologici che sono stati affrontati e in parte risolti.

Innanzitutto, il noto *problema dell'attribuzione degli effetti osservati*, che possono essere in linea di principio riferiti ad altre cause oltre che alla valutazione. Questo aspetto è cruciale per la validità delle osservazioni ed è particolarmente importante per quanto riguarda l'analisi dei dati raccolti attraverso la *survey* e l'analisi bibliometrica.

Per risolvere operativamente il problema, è stato adottato un approccio “contributivo” nel quale si assume che l'effetto osservato derivi dalla combinazione di più eventi causali, e quindi il compito del ricercatore è quello di stabilire l'importanza (in termini di significatività e robustezza) della valutazione rispetto a una serie di concuse individuate come rilevanti e potenzialmente suscettibili di generare il risultato osservato. Pertanto, nel lavoro di campo, una volta individuati gli effetti per i quali viene segnalata un'influenza della valutazione, si testa questa influenza, attraverso metodi quantitativi o qualitativi, rispetto ad altre spiegazioni possibili per la produzione di quel determinato effetto. Se non è possibile individuare concuse che abbiano contribuito a un determinato effetto (per esempio nel caso in cui venga riportata una percezione sulla variazione della qualità della ricerca prodotta dalla valutazione), si caratterizza il soggetto che ha espresso il giudizio sulla base di caratteristiche personali, disciplinari e di contesto che possono aver contribuito a strutturare il giudizio rilevato.

Inoltre, nel caso dell'analisi bibliometrica i cambiamenti nelle pubblicazioni e nei risultati possono essere prodotti da diversi eventi concomitanti (è noto che le analisi longitudinali bibliometriche possono essere influenzate, tra l'altro, da una distorsione sistematica correlata all'espansione della copertura delle banche dati bibliometriche). Tuttavia, da un lato, la combinazione tra diversi strumenti di indagine consente di mitigare questo inconveniente dal punto di vista delle evidenze di effetto prodotte; dall'al-

tro, assumiamo che gli effetti possono derivare da cause diverse, dove la valutazione può essere quella che gioca un ruolo preminente.

Il secondo limite è il *ritardo temporale*. Osservare gli effetti in un determinato momento non permette di catturare il complesso dei cambiamenti prodotti, perché alcuni potrebbero emergere dopo un lungo lasso di tempo e non essere visibili al momento dell'osservazione. Consideriamo questo un vincolo minore; l'analisi è stata sviluppata in una fase in cui la valutazione esterna della ricerca in Italia può considerarsi ormai istituzionalizzata; alcuni effetti sul lavoro accademico e sui modi di produzione della ricerca sono pertanto già visibili e hanno già prodotto cambiamenti consolidati, altri mostrano segnali di sviluppo futuro che permettono in ogni caso di essere individuati a fini analitici.

Un'altra fonte di distorsione può derivare dal *basso tasso di risposta* e dalle differenze di composizione dei gruppi dei rispondenti e dei non rispondenti nelle *survey*. Questo problema è stato affrontato attraverso la gestione di diverse *wave* di indagine su un campione costantemente ribilanciato in relazione alle cadute di risposte osservate (Spinello *et al.*, 2021).

Infine, siamo consapevoli che il lavoro di ricerca accademico e la produzione di conoscenza sono influenzati anche dalla valutazione dell'insegnamento, data la stretta connessione esistente tra i due ambiti. Sebbene il progetto sia incentrato sulla valutazione della ricerca, gli effetti della valutazione dell'insegnamento sono stati esplorati durante le interviste e le storie di vita quotidiana nel caso in cui fossero emersi spontaneamente dagli intervistati.

4. I contributi

Il volume può essere idealmente diviso in tre parti. La prima parte (capitoli 1, 2, 3 e 4) presenta risultati relativi agli aspetti più strettamente collegati a come la valutazione interferisca, influenzandola, sulla nozione della qualità della ricerca prodotta e come detta influenza sia differente in base al settore disciplinare di afferenza o alla posizione accademica.

Il capitolo 1 considera gli effetti prodotti dalla VQR sulla percezione degli accademici rispetto all'effetto della valutazione sulla qualità dei risultati della ricerca prodotti e utilizza i risultati emersi dalla *survey* nazionale.

Il capitolo 2 approfondisce, attraverso le storie di vita, se e come la presenza della valutazione influisca sul lavoro dei giovani ricercatori e quali reazioni essa susciti. I giovani ricercatori sono un segmento importante della popolazione accademica in quanto soggetti caratterizzati dall'aver cominciato il percorso di carriera nella ricerca quando la valutazione esterna

era già uno strumento ampiamente utilizzato dal governo e aveva già prodotto i cambiamenti organizzativi e di *governance* all'interno degli atenei per gli adempimenti previsti. A questo si aggiunga il fatto che non avendo ancora un'identità reputazionale e di ricerca consolidata, sono più soggetti alle spinte che provengono da incentivi esterni.

Nel capitolo 3 viene sviluppata invece un'analisi specifica relativa alle scienze sociali per approfondire gli effetti della valutazione sulla selezione da parte degli accademici dei temi di ricerca e, più in generale, sulla formazione dell'agenda di ricerca. Questo lavoro utilizza un metodo misto, combinando una *survey* dedicata con interviste per approfondimenti di punti difficili da esplorare attraverso un questionario.

Chiude la prima parte il capitolo 4, dove vengono riportati i risultati della *survey* nazionale in merito agli effetti della valutazione esterna sulle collaborazioni e sulla mobilità internazionale degli accademici per comprendere se e in che misura la valutazione contribuisca a promuovere o ostacolare questi aspetti legati all'organizzazione del lavoro scientifico, confrontando l'importanza della valutazione rispetto ad altri fattori influenti.

La seconda parte del volume comprende i capitoli 5 e 6 che affrontano problemi di tipo epistemologico collegati all'avvento della valutazione della ricerca. Nel primo dei due capitoli si approfondisce come sia cambiata la percezione della qualità della ricerca nel corso degli anni in due settori disciplinari, quelli delle scienze sociali e del management attraverso una revisione sistematica della letteratura che utilizza l'approccio della *Systematic Literature Network Analysis*. Nello stesso scritto si controllano gli effetti della pratica della valutazione sullo sviluppo del settore delle scienze economiche.

Nel capitolo 6 invece l'interesse è sull'individuazione delle logiche istituzionali che presiedono ai processi di pubblicazione e ai processi di allocazione delle risorse. Le pubblicazioni sono il meccanismo fondamentale di valutazione della qualità di un prodotto scientifico e si vuole indagare il modo di selezione attraverso la *peer review* per comprendere le preferenze espresse dai revisori, nonché il ruolo svolto dagli editori e la ricaduta sull'avanzamento di carriera. La distribuzione delle risorse rappresenta l'effetto prodotto dalla valutazione della ricerca scientifica. Nel capitolo si approfondiscono gli effetti dell'allocazione basata sulla performance in una particolare organizzazione, gli IRCCS (Istituti di ricovero e cura a carattere scientifico) che rappresentano una realtà contigua al mondo accademico vero e proprio.

La terza parte comprende le analisi condotte attraverso l'uso dei dati bibliometrici, confrontando i cambiamenti intervenuti tra il primo periodo 2007-2011 e il secondo periodo 2013-2017. In particolare, l'analisi bibliometrica mira a verificare l'esistenza di un cambiamento significativo nella produzione scientifica lungo le dimensioni maggiormente interessate

dagli esercizi di valutazione sopra citati: pubblicazioni e citazioni totali, citazione media per prodotto, pagine medie per paper, quota di citazioni incrociate e autocitazioni; e nelle strategie di collaborazione di ricercatori e professori (coautori medi per articolo, quota di co-autorialità nazionali-internazionali-intramurali-extramurali).

Il capitolo 7 si concentra sul confronto della performance relativa di ricerca delle università, a livello complessivo, di disciplina e di settore scientifico, prima e dopo l'introduzione della VQR, in quanto sistema di allocazione basato sulle prestazioni, contribuendo, quindi, al controllo dei risultati prodotti in letteratura da numerosi studiosi che hanno affrontato il tema.

Il capitolo 8 esplora gli effetti dell'ASN sul miglioramento della capacità di reclutamento dei docenti da parte delle università italiane. Vengono presentate due analisi. La prima è basata sulle valutazioni delle performance scientifiche dei nuovi assunti rispetto ai colleghi italiani dello stesso settore e ruolo accademico in due periodi successivi di cinque anni, prima e dopo l'avvio della VQR. La seconda analisi si concentra sull'efficacia del reclutamento da parte di tutte le università italiane, nei due quinquenni successivi, per verificare eventuali miglioramenti.

Infine, il capitolo 9 affronta un problema molto dibattuto relativo al presunto effetto che la VQR potrebbe avere sull'aumento della propensione di docenti e ricercatori ad autocitarsi all'interno dei propri paper, sviluppando altresì considerazioni sull'opportunità di migliorare in Italia l'uso degli indicatori bibliometrici per finalità di valutazione.

1. La Valutazione della Qualità della Ricerca e i suoi effetti sul lavoro accademico: differenze disciplinari e di contesto

di *Antonio Zinilli, Andrea Orazio Spinello, Emanuela Reale*

1. Introduzione

A partire dagli anni '80, la ridotta disponibilità di risorse economiche pubbliche da destinare a ricerca e innovazione, unita all'incremento dei costi associati alla ricerca, ha portato a una crescente richiesta da parte della società di maggiore trasparenza e responsabilità delle università europee rispetto all'utilizzo dei fondi pubblici (Reale e Marini, 2017). L'idea, strettamente legata al paradigma neoliberale che si andava consolidando e diffondendo negli stessi anni in Europa occidentale, era quella di aumentare la performance scientifica e didattica delle università per ottenere il maggior ritorno possibile dall'investimento pubblico (*value for money*).

In risposta a ciò, molti governi in Europa hanno cominciato a incentivare l'innovazione e l'eccellenza nella ricerca, instaurando, tra l'altro, sistemi nazionali massivi di valutazione al fine di distribuire le risorse finanziarie di base (il c.d. *core funding*) sulla base della qualità dei lavori prodotti misurata attraverso specifiche tecniche e indicatori (Hicks *et al.*, 2015).

La prima iniziativa in questa direzione fu lanciata nel Regno Unito nel 1986, con l'obiettivo di aumentare la selettività nell'assegnazione dei fondi. Da allora, il Regno Unito ha affinato e perfezionato il suo sistema, conducendo ulteriori revisioni e aggiustamenti. Le valutazioni della ricerca promosse dal governo e organizzate centralmente, spesso attraverso l'ausilio di apposite agenzie di valutazione, sono state in seguito adottate in numerosi altri paesi per misurare la qualità della ricerca nelle Università e, di conseguenza, come strumento per l'assegnazione dei finanziamenti per ricerca sulla base della performance delle università e della loro collocazione in *rating* formulati sulla base di metriche riferite alla qualità scientifica. Si è posta grande enfasi nel verificare se tali iniziative

abbiano portato a un miglioramento sia nella qualità sia nella quantità delle ricerche condotte, ma i risultati non sono ancora sufficienti per una risposta conclusiva e condivisa.

Il Capitolo intende contribuire al tema analizzando come la Valutazione della Qualità della Ricerca (VQR) italiana abbia influenzato la qualità della ricerca accademica. La parola “ricerca” abbraccia in realtà una molteplicità di attività che compongono lo sforzo per la creazione di nuova conoscenza scientifica; nel contesto di questo contributo, la “qualità della ricerca” si riferisce alla qualità delle sole pubblicazioni scientifiche. In questo lavoro per pubblicazione scientifica intendiamo un contributo accademico formalmente pubblicato, come articoli, monografie o capitoli, che riporta i risultati di ricerche originali ed è soggetto a processi di valutazione della qualità tramite revisione paritaria.

Lo scopo che ci proponiamo è esplorare la percezione degli accademici italiani in merito all’incidenza della VQR sulla qualità della loro produzione scientifica, al fine di fornire informazioni chiave per decisioni più consapevoli riguardanti questo tema.

Tre domande guidano il nostro lavoro: I) secondo la percezione degli accademici, quali effetti produce la VQR sulla qualità della loro ricerca? II) esistono differenze disciplinari nella percezione della qualità? III) quali sono i fattori più importanti che influenzano le percezioni di coloro che considerano la VQR come un elemento che produce un effetto sulla qualità della ricerca e coloro che non percepiscono alcun effetto?

L’ipotesi che intendiamo verificare è che la percezione degli effetti della VQR vari in base al campo scientifico di appartenenza, inducendo a cambiamenti strategici associati a differenti regimi epistemici. Poiché dunque vogliamo controllare l’esistenza di diverse nozioni di “qualità” nella ricerca scientifica che influenzano gli effetti prodotti dalla valutazione, le nostre analisi valorizzano le disaggregazioni di tipo disciplinare.

I risultati di questo Capitolo forniscono spunti di riflessione per i decisori politici, permettendo di comprendere il punto di vista degli accademici circa gli effetti prodotti sul lavoro di ricerca. Inoltre, i risultati possono offrire approfondimenti sulle possibili aree di miglioramento, sulle eventuali criticità e sulle specifiche esigenze legate alle diverse discipline o contesti universitari.

Il Capitolo è strutturato come segue: nel paragrafo 2, si presenta il quadro teorico di riferimento; nel paragrafo 3, si discutono dati e metodi; nel paragrafo 4, vengono presentati i risultati; e nel paragrafo 5, le conclusioni.

2. Letteratura di riferimento e ipotesi di lavoro

L'evoluzione della valutazione della ricerca nei paesi europei è stata influenzata da riforme che hanno mirato a intensificare la direzione strategica da parte dei decisori politici e a rafforzare l'autonomia delle università (Whitley, 2007). In molti paesi europei, Italia inclusa, la valutazione della ricerca si è consolidata come pilastro centrale nella *governance* della ricerca. Questo meccanismo non solo determina l'assegnazione dei finanziamenti statali e stabilisce vari sistemi di incentivi, ma ha anche un impatto significativo sull'organizzazione interna delle università e sulla distribuzione di ruoli, funzioni e responsabilità (Reale e Marini, 2017). La valutazione rappresenta, nelle intenzioni dei governi che la promuovono, un incentivo per le varie organizzazioni di ricerca verso l'adozione di strategie orientate al raggiungimento di valutazioni più elevate nella competizione per le risorse pubbliche (Zacharewicz *et al.*, 2019), ed è suscettibile dunque di produrre effetti a cascata sia a livello organizzativo sia a livello di singolo ricercatore (Reale e Zinilli, 2017).

La VQR, introdotta in Italia nel 2004 con l'obiettivo di migliorare la qualità della ricerca scientifica, ha indubbiamente innescato un vivo dibattito nel panorama accademico e, in certa misura, in quello politico. Gli esercizi valutativi, nel tentativo di identificare le istituzioni di prestigio e incentivare l'eccellenza, hanno introdotto nuove metriche e criteri. Dalle nuove analisi sono emersi giudizi positivi o negativi che spesso facevano riferimento all'uso proprio o improprio delle metriche stesse piuttosto che al valore aggiunto che l'intero processo di valutazione poteva aver avuto per l'università (Palumbo e Pennisi, 2015; Facchini e Pennisi, 2020; Fasanella e Martire, 2020).

L'esercizio VQR 2011-2014 è stato particolarmente sotto osservazione sul piano scientifico e politico. Una nota preoccupazione riguardava l'uso delle metriche associate alle riviste scientifiche come principale strumento di valutazione. In particolare, è stato evidenziato come, a causa della variabilità delle citazioni per gli articoli della stessa rivista, questo sistema potesse risultare problematico. Questa critica è stata particolarmente sottolineata da Franceschini e Maisano (2017), che hanno argomentato come una simile metodologia rendesse arduo identificare istituzioni di prestigio e quindi la qualità scientifica dei ricercatori stessi. Il dibattito si è intensificato quando si è discusso della combinazione di metriche basate sulle citazioni e sull'Impact Factor (IF) delle riviste, due *proxy* particolarmente utilizzate per valutare la qualità scientifica delle pubblicazioni. Diversi autori, tra cui Abramo e D'Angelo (2015c), hanno manifestato preoccupazioni riguardo all'approccio che combina diverse metriche. Hanno sottolineato come tale metodologia non trovi un ampio sostegno nella letteratura scien-

tifica consolidata e, di conseguenza, non goda di un pieno riconoscimento all'interno della comunità accademica.

Il disaccordo citato ha caratterizzato anche le analisi sui risultati della VQR 2004-2010. Mentre alcuni studi, come quello di Ancaiani *et al.* (2015) proprio sull'esercizio 2004-2010, hanno riconosciuto una coerenza tra revisione tra pari e analisi bibliometrica, altri, come Baccini e De Nicolao (2016), hanno contestato tali risultati. La questione relativa all'uso dell'IF come principale indicatore di qualità ha diviso la comunità accademica. Se da un lato alcune ricerche, come quella di Hicks *et al.* (2015), lo criticavano (principalmente perché tende a creare disparità tra aree scientifiche), altri studi, come quello di Hunt *et al.* (2010), ne sottolineavano la coerenza e la validità. Tuttavia, al di là di questi aspetti tecnici e metodologici, un'altra critica radicata era la mancanza di trasparenza nel processo VQR, soprattutto perché i dati specifici non erano pubblicamente accessibili. Questa opacità, combinata con la variabilità delle discipline all'interno delle università e dei dipartimenti, ha complicato ulteriormente le comparazioni tra istituzioni e tra ricercatori provenienti da campi scientifici differenti.

Sebbene l'impatto della VQR sul contesto scientifico italiano sia un argomento ampiamente trattato, la letteratura raramente fornisce analisi empiriche approfondite riguardo alla percezione degli accademici circa l'effetto della VQR sulla qualità del loro lavoro scientifico. Più in generale, è estremamente complesso individuare un nesso causale tra valutazione e suoi effetti, anche quando si restringe il campo di osservazione alle sole pubblicazioni scientifiche.

La definizione di “qualità” nel contesto della ricerca accademica è materia di intenso dibattito che si traduce in posizioni contrastanti, come sottolineato da diversi studi (Langfeldt *et al.*, 2020; Aksnes *et al.*, 2019; Abramo, D'Angelo e Reale, 2019). Il concetto di “qualità” è piuttosto fluido e sottoposto a continue trasformazioni; è influenzato non solo dalle mutevoli dinamiche della ricerca stessa, ma anche da forze esterne, come le metriche di valutazione. Questa fluidità del concetto di qualità si riflette chiaramente nella percezione degli accademici, che si trova al crocevia tra l'adesione a standard scientifici consolidati e la risposta alle pressioni esterne, come quelle indotte dalla VQR. Secondo Langfeldt *et al.* (2020), esistono due principali definizioni di qualità che spesso sono in contrasto tra loro. La prima definizione è legata alle specifiche comunità scientifiche, che si basano su regole e metodi consolidati per produrre conoscenza. La seconda definizione, invece, è legata alla politica e si riflette nelle comunità esterne al mondo della ricerca. Queste ultime tendono a valutare la qualità basandosi su indicatori specifici, dando origine a differenti criteri di valutazione. Pertanto, le valutazioni esterne, come la VQR, possono influenzare signifi-

cativamente le percezioni della qualità della ricerca, spingendo ricercatori e istituzioni verso particolari standard nelle loro pubblicazioni e contenuti. Questa percezione di “qualità” può essere ulteriormente modulata dalla comunità epistemica di riferimento e dal contesto istituzionale in cui la ricerca viene condotta. Gli studiosi hanno fornito sia prove a favore che contro riguardo all’effetto della valutazione sull’agenda di ricerca e sulla qualità in vari ambiti scientifici (Osuna *et al.*, 2011; Abramo e D’Angelo, 2015c). Il dibattito citato mostra la complessità e la sfaccettatura del rapporto tra valutazione e produzione scientifica in diverse aree disciplinari. La questione rimane aperta e sottolinea l’importanza di una riflessione continua sulla metodologia e sui criteri di valutazione adottati nel panorama accademico.

In questo contesto, il presente Capitolo intende approfondire specificamente la percezione degli accademici italiani. L’obiettivo centrale è comprendere come gli accademici percepiscono gli effetti della VQR sul loro lavoro scientifico. La domanda fondamentale a cui cerchiamo di rispondere è:

D: il campo disciplinare e il contesto istituzionale in cui un accademico opera influenzano la sua percezione degli effetti della VQR sulla qualità della sua ricerca?¹

Rispetto alla domanda appena evidenziata, l’ipotesi che intendiamo controllare, seguendo la letteratura appena citata, è:

Hp: la percezione degli accademici non è omogenea, ma è piuttosto influenzata da una complessa interazione di fattori disciplinari e istituzionali.

L’ipotesi considera dunque che le diverse comunità scientifiche, con le loro specifiche regole e i loro metodi per produrre conoscenza (Langfeldt *et al.*, 2020), siano influenzate in modo differente dalla valutazione della ricerca. Allo stesso modo, i contesti istituzionali e geografici, con le loro specifiche dinamiche, potrebbero orientare il modo in cui gli accademici vedono e rispondono ai criteri di valutazione esterna. Questo si deve al fatto che la VQR – come altri esercizi sviluppati in ambito europeo primo fra tutti il REF nel Regno Unito (Whitley, 2007; Zacharewicz *et al.*, 2019), riveste un ruolo fondamentale nella *governance* della ricerca in Italia. Infatti, oltre a guidare l’allocazione delle risorse, essa modella profondamente

1. Nel contesto di questo Capitolo, il concetto di qualità della ricerca si intende come l’insieme di attributi che caratterizzano la ricerca accademica, e non si basa su una definizione preconstituita. Il concetto è operativizzato da domande specifiche nel questionario utilizzato per valutare gli effetti della VQR sulla percezione degli accademici.

la struttura delle istituzioni accademiche, definendone le priorità, le regole operative e, in ultima analisi, l'orientamento e la direzione delle attività di ricerca (Reale e Pennisi, 2013; Facchini e Pennisi, 2020).

Gli effetti della VQR e la percezione che ne hanno gli accademici cambiano quindi sulla base del contesto disciplinare e istituzionale.

3. Dati e metodi

I dati utilizzati provengono da un'indagine nazionale svolta in Italia nel 2020-2021. L'indagine, basata su un questionario strutturato, ha coinvolto un campione probabilistico della popolazione di tutti gli accademici delle università italiane, raccogliendo risposte da 1.365 unità (Spinello *et al.*, 2021).

Nell'ambito della rilevazione sulla percezione degli effetti della VQR da parte degli accademici, l'indagine ha raccolto dati su caratteristiche individuali come genere, età, ruolo accademico e area disciplinare, e contesti geografici e istituzionali, seguendo la categorizzazione ISTAT in 5 macroaree geografiche (Italia nord-occidentale, Italia nord-orientale, Italia centrale, Sud Italia e Isole).

Riguardo al ruolo accademico, sono state considerate le cinque posizioni introdotte dalla Legge Gelmini (L. 240/2010): ricercatore ad esaurimento (pre-riforma Gelmini), RTD-A (ricercatore a tempo determinato di tipo A), RTD-B (ricercatore a tempo determinato di tipo B), professore associato e professore ordinario. Prima della riforma Gelmini, era presente la sola figura del ricercatore, che rappresentava una posizione a tempo indeterminato nell'ambito accademico. Con l'introduzione della riforma Gelmini, sono state introdotte due nuove figure: RTD-A e RTD-B. La posizione RTD-A rappresenta una posizione iniziale all'università, generalmente destinata a ricercatori nella fase iniziale della loro carriera accademica. Questo tipo di contratto ha una durata limitata ed è spesso l'anticamera per posizioni più stabili all'interno dell'università. La posizione RTD-B, d'altro canto, è una posizione più avanzata, spesso coperta da ricercatori che hanno già accumulato esperienza e hanno dimostrato le loro capacità in termini di produzione scientifica. La posizione di RTD-B, dopo un periodo di tre anni e a seguito di una valutazione interna condotta dal dipartimento, viene convertita nella posizione di professore associato. La figura del ricercatore (ad esaurimento) a tempo indeterminato è stata sostituita da RTD-A e RTD-B. Nelle analisi successive abbiamo aggregato le tre figure – ricercatore (ad esaurimento), RTD-A e RTD-B – in un'unica categoria generica “ricercatore”.

La categorizzazione dell'area disciplinare degli accademici ha seguito quella del Ministero dell'Università e della Ricerca, considerando le 14 aree CUN:

1. Scienze matematiche e informatiche;
2. Scienze fisiche;
3. Scienze chimiche;
4. Scienze della Terra;
5. Scienze biologiche;
6. Scienze mediche;
7. Scienze agrarie e veterinarie;
8. Ingegneria civile ed architettura;
9. Ingegneria industriale e dell'informazione;
10. Scienze dell'antichità, filologico-letterarie e storico-artistiche;
11. Scienze storiche, filosofiche, pedagogiche e psicologiche;
12. Scienze giuridiche;
13. Scienze economiche e statistiche;
14. Scienze politiche e sociali.

Le principali caratteristiche della *survey* sono riportate nel prossimo paragrafo. Per un maggiore dettaglio e per consultare il questionario completo, si rimanda a Spinello *et al.* (2021).

Per comprendere le percezioni degli accademici riguardo gli effetti della VQR sulla qualità della loro ricerca è stato utilizzato un modello LOGIT. Il modello LOGIT permette di osservare le relazioni tra una variabile dipendente binaria e una serie di variabili indipendenti. La variabile dipendente rappresenta la percezione degli accademici sulla qualità della loro ricerca in relazione alla VQR. Questa variabile è stata formulata a partire dalla seguente domanda del questionario: *Pensando al suo lavoro di ricerca negli ultimi dieci anni, la necessità di produrre pubblicazioni in linea con le indicazioni della VQR ha influito sulla qualità della sua ricerca?*

Per garantire una coerenza con l'oggetto dello studio, solo gli accademici che avevano effettivamente partecipato ad almeno una edizione della VQR sono stati coinvolti e invitati a rispondere a questa domanda. Ciò ha assicurato che le risposte ottenute fossero basate sull'esperienza diretta e personale degli accademici, e non su opinioni generiche. Le variabili indipendenti, invece, sono state selezionate per analizzare quali fattori potrebbero influenzare la percezione degli accademici e verificare la nostra ipotesi iniziale.

Con il modello LOGIT possiamo dunque identificare quali fattori hanno un impatto significativo sulla probabilità che un accademico risponda “La VQR ha influito sulla qualità” (valore della dummy = 1) oppure “La VQR non ha influito sulla qualità” (valore della dummy = 0). In questo tipo di analisi, è essenziale considerare un punto di riferimento o

benchmark. Nel contesto del modello LOGIT, il *benchmark* rappresenta la categoria di base di una variabile discreta. Il *benchmark* è cruciale poiché fornisce un quadro di confronto per le altre categorie della variabile, facilitando una comprensione chiara delle differenze relative e dei fattori che hanno un impatto significativo sulla probabilità in questione. Di seguito, vengono indicate la variabile dipendente e le variabili indipendenti.

Tab. 1 - Variabili utilizzate nel modello LOGIT

Variabili	Descrizione
Variabile dipendente	
La VQR influisce sulla qualità della ricerca?	Variabile dicotomica: 0 La VQR non ha influito sulla qualità 1 La VQR ha influito sulla qualità
Variabili indipendenti	
Genere	Genere dei rispondenti: 0 Femmina 1 Maschio
Bambini	Presenza dei bambini: 0 no 1 Sì, uno 2 Sì, più di uno
Area geografica	Distribuzione geografica dell'università di appartenenza: 1 Italia Nord-Ovest 2 Italia Nord-Est 3 Italia Centrale 4 Sud Italia 5 Isole
Dimensione Università	Dimensione universitaria di appartenenza dell'intervistato in termini di studenti iscritti e personale. La dimensione delle università considerate è calcolata tenendo in considerazione l'area CUN: 1 Grande università 2 Media università 3 Piccola università
Ruolo accademico	Ruolo accademico degli intervistati: 1 Professore Ordinario 2 Professore Associato 3 Ricercatore (prima Legge Gelmini), ricercatore a tempo determinato di tipo A, ricercatore a tempo determinato di tipo B

Tab. 1 - segue

Variabili	Descrizione
Aree CUN	Elenco delle macro-discipline scientifiche di appartenenza (aree CUN): 1. Scienze matematiche e informatiche 2. Scienze fisiche 3. Scienze chimiche 4. Scienze della Terra 5. Scienze biologiche 6. Scienze mediche 7. Scienze agrarie e veterinarie 8. Ingegneria civile ed architettura 9. Ingegneria industriale e dell'informazione 10. Scienze dell'antichità, filologico-letterarie e storico-artistiche 11. Scienze storiche, filosofiche, pedagogiche e psicologiche 12. Scienze giuridiche 13. Scienze economiche e statistiche 14. Scienze politiche e sociali

4. La survey

La *survey* si è basata sull'invio di un questionario strutturato a un campione rappresentativo della popolazione accademica appartenente alle università italiane soggette alla valutazione da parte dell'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANVUR). L'elenco dei partecipanti è stato estratto alla data del 31 luglio 2019 dal database MUR-CINECA (<https://cercauniversita.cineca.it>) che comprendeva 55.535 accademici afferenti a 98 università italiane. Il processo di acquisizione dei dati è stato implementato attraverso il metodo dell'intervista via web (*Computer Assisted Web Interviewing – CAWI*), sfruttando i servizi dell'app Lime Survey, integrata all'interno della piattaforma per la realizzazione di indagini statistiche del Consiglio Nazionale delle Ricerche (<https://survey.cnr.it>). Gli indirizzi e-mail dei rispondenti sono stati estratti esclusivamente tra coloro che erano inclusi nel campione costruito per la *survey*. Ciò ha garantito che le e-mail fossero inviate solo agli accademici selezionati in modo stratificato secondo i dati del Ministero dell'Università e della Ricerca (MUR), assicurando così una rappresentanza adeguata delle diverse discipline e contesti istituzionali.

Il questionario generale è stato strutturato in cinque sezioni – “Strategia di pubblicazione”, “Produttività scientifica”, “Pratica di ricerca”, “Collaborazioni” e “Mobilità internazionale” – oltre a una sezione iniziale finalizzata

a raccogliere le informazioni di profilo del rispondente. La progressione delle domande in ciascuna sezione ha seguito una struttura “a imbuto”, con domande generali che progressivamente diventavano più specifiche, con l’obiettivo di esplorare i vari aspetti della dimensione considerata. Ogni sezione includeva anche uno spazio aperto finale per “ulteriori commenti”, che ha consentito ai partecipanti di esprimersi liberamente sull’argomento proposto o segnalare eventuali problematiche relative alla compilazione. L’intervista completa consisteva di 44 domande e richiedeva circa 13 minuti per essere completata.

Per la selezione del campione è stato utilizzato un metodo di campionamento casuale stratificato “a più stadi”. Il processo è stato diviso in due fasi: inizialmente sono state selezionate unità di campionamento più ampie, e in seguito unità di campionamento più piccole all’interno delle prime. Le unità più grandi erano costituite dagli atenei, mentre le unità più piccole erano costituite dagli accademici. In entrambe le fasi è stato effettuato un campionamento casuale stratificato delle unità. La proporzione di alcune categorie di stratificazione della popolazione ha determinato il numero di unità estratte in ciascuna fase. Le categorie considerate nella prima fase sono state la posizione geografica e le dimensioni dell’università, mentre nella seconda fase sono state il genere, l’area scientifica CUN e la posizione accademica degli accademici appartenenti a ciascuna università.

La prima *wave* della *survey* è partita nell’ottobre 2020 e si è protratta per 16 giorni. A questa è seguita una serie di altre tre *wave*, ciascuna con una durata compresa tra i 14 e i 16 giorni, tra novembre 2020 e marzo 2021. Al termine delle quattro *wave*, è stata effettuata una fase di richiamo di 6 giorni per i non rispondenti. La *survey* si è conclusa alla fine di marzo 2021. L’indagine ha ricevuto 1.365 risposte, con un tasso di risposta del 43,1%, che sale al 43,8% considerando solo le unità invitate a partecipare che hanno effettivamente ricevuto l’e-mail di invito (3.120 su 3.165). L’analisi delle risposte pervenute ha confermato una buona rappresentatività del campione, considerando le cinque categorie utilizzate per stratificare gli accademici, vale a dire posizione geografica, dimensioni dell’università, genere, area scientifica CUN e posizione accademica.

Come già detto, per studiare gli effetti della VQR sulla qualità della ricerca, è stata utilizzata una domanda specifica del questionario: *Pensando al suo lavoro di ricerca negli ultimi dieci anni, la necessità di produrre pubblicazioni in linea con le indicazioni della VQR ha influito sulla qualità della sua ricerca?* Per garantire coerenza con l’oggetto dello studio, solo gli accademici che avevano partecipato ad almeno una edizione della VQR sono stati coinvolti e invitati a rispondere a questa domanda. La qualità

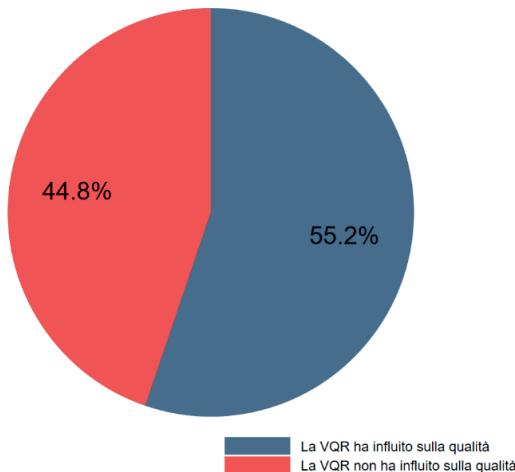
della ricerca percepita dai rispondenti si riferisce alla valutazione soggettiva dell'impatto della VQR sulla loro produzione scientifica, variabile in base al settore disciplinare di appartenenza. Questa percezione include aspetti come la rilevanza, l'innovatività e l'allineamento della ricerca agli standard del proprio campo. La variabile è stata quindi categorizzata come “1” se l'accademico percepiva che la VQR aveva influenzato la qualità della sua ricerca e “0” se non percepiva alcuna influenza.

5. Risultati: analisi descrittiva

Prima di procedere all'analisi delle relazioni tra le variabili utilizzando il modello LOGIT, in questo paragrafo presentiamo alcune analisi descrittive per comprendere la distribuzione dei partecipanti all'indagine rispetto alla domanda del questionario oggetto del nostro studio.

La Fig. 1 evidenzia la percezione predominante tra gli intervistati in merito all'influenza della VQR sulla qualità del lavoro scientifico.

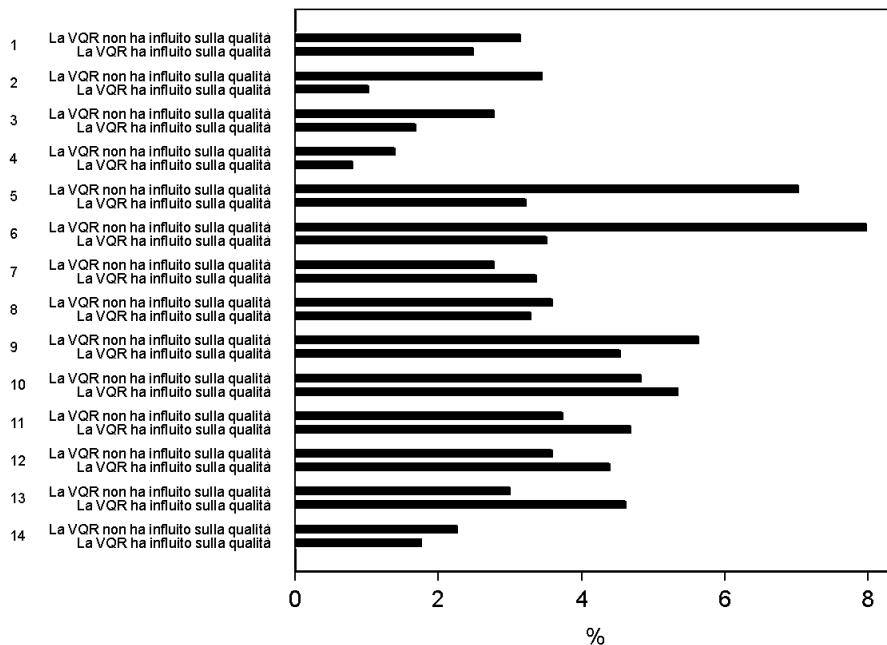
Fig. 1 - Percentuale dei rispondenti in merito all'influenza della VQR



Il 55,2% degli intervistati ritiene che la VQR abbia avuto un impatto sulla qualità della loro ricerca. Questa maggioranza relativa suggerisce che più della metà degli intervistati vede un effetto tangibile nella VQR, considerandola come un fattore che ha portato a cambiamenti. D'altro can-
to, è interessante notare che una percentuale pari al 44,8% sostiene che la

VQR non abbia avuto alcun effetto. Questa percentuale rappresenta una porzione considerevole del campione e suggerisce che esiste una quantità significativa di rispondenti che non considera la valutazione un fattore di cambiamento delle proprie pratiche o risultati. Questa differenza nelle risposte sottolinea l'importanza di esplorare l'efficacia e la percezione della VQR nel contesto accademico rispetto alle aree disciplinari. La successiva Fig. 2 mostra la distribuzione delle risposte in relazione alle diverse aree CUN. Le risposte sono suddivise per ogni area CUN in due categorie principali: "La VQR non ha influito sulla qualità" e "La VQR ha influito sulla qualità". Il grafico presenta la distribuzione percentuale delle risposte in ciascuna delle 14 aree CUN, esprimendo la percentuale di ciascuna risposta rispetto al totale delle risposte raccolte. Il test chi-quadro di Pearson (χ^2) ha mostrato una differenza nelle risposte tra le diverse aree CUN (statisticamente significativa). Questo risultato suggerisce che l'appartenenza a un'area disciplinare influisce in modo significativo sulla percezione dell'impatto della VQR sulla qualità della ricerca. Questo conferma che la disciplina è un fattore importante nella percezione degli effetti della valutazione sulla ricerca accademica.

Fig. 2 - Percentuale dei rispondenti in merito all'influenza della VQR per area CUN



La Fig. 2 evidenzia come soprattutto per le aree delle Scienze Sociali e Umane “La VQR ha influito sulla qualità” (ad esempio, le aree 10. Scienze dell’antichità, filologico-letterarie e storico-artistiche, 11. Scienze storiche, filosofiche, pedagogiche e psicologiche, 12. Scienze giuridiche, 13. Scienze economiche e statistiche), mentre le aree delle Scienze Fisiche e Ingegneristiche e Scienze della Vita optano per la risposta secondo la quale “La VQR non ha influito sulla qualità” (ad esempio, le aree CUN 1. Scienze matematiche e informatiche, 2. Scienze Fisiche, 3. Scienze chimiche e particolarmente visibile nelle aree 5. Scienze biologiche e 6. Scienze mediche).

I commenti dei rispondenti offrono una serie di spunti interessanti rispetto ai risultati dell’analisi descrittiva. Il primo elemento da segnalare sono le divergenze di percezione sull’impatto della valutazione sulla qualità della ricerca tra accademici appartenenti a discipline nell’ambito delle scienze sociali da una parte e all’ambito delle scienze sperimentali dall’altra. Sebbene siano emersi pareri sia positivi che negativi da entrambe le parti, alcuni di questi spiccano per chiarezza e nettezza delle posizioni espresse. Ne riportiamo alcuni richiamando il lettore a considerare che, trattandosi di commenti liberi, gli intervistati a volte si riferiscono all’intero sistema valutativo della ricerca, includendo nei propri giudizi sia la VQR che l’Abilitazione Scientifica Nazionale (ASN).

Per quanto riguarda le scienze sociali e umane, un intervistato di area 12 osserva che:

[...] la spinta a pubblicare esercitata da ASN e VQR sia in realtà molto deleteria per la qualità della ricerca, dal momento che induce a considerare soprattutto gli aspetti quantitativi della ricerca, a scapito della qualità di quanto realizzato. [...] Personalmente abolirei sia il sistema di ASN che di VQR così come concepiti in questo momento (Ricercatore a tempo determinato di tipo A, piccola università dell’area Sud e Isole, Area CUN 12).

Ancora più radicale nel giudizio è una collega di area 10:

Per le Scienze Umane, i criteri VQR e ASN mi paiono punitivi, nazionalistici, anacronistici e poco orientati verso la qualità della ricerca (Ricercatore a esaurimento, grande università del Nord-Ovest, Area CUN 10).

Un professore di area 13 suggerisce che l’effetto sulla qualità non è il problema principale che la valutazione produce; esso piuttosto è da ricercarsi nella spinta esercitata dalla valutazione ad aderire alla ricerca *mainstream*:

[...] c’è la questione della necessità di adattarsi a argomenti di ricerca e modalità di fare ricerca secondo i canoni mainstream e quindi della pressione ad adattarti

a condizioni che non rispecchiano interessi e esperienza. D'altra parte, se non ti adatti non esci nelle riviste di fascia A e la valutazione della tua ricerca (indipendentemente dall'effettiva qualità) scende (Professore ordinario, piccola università del Nord-Est Area CUN 13).

Secondo un'altra intervistata appartenente sempre all'area scientifica 13, il rischio, generato dalla valutazione, di impoverimento della qualità scientifica legittima un atteggiamento di resistenza e di non adesione alla spinta verso una maggiore produttività scientifica:

Gli indicatori ASN e VQR avrebbero contribuito ad una pesante distorsione della mia produttività scientifica verso la quantità, penalizzando la qualità. Con consapevolezza ma assumendomi grossi rischi in termini di possibilità di carriera entro l'università italiana, ho deciso di NON cambiare indirizzo e continuare a concentrarmi sul produrre ricerca di qualità invece che ricerca in quantità (Professore ordinario, grande università del Nord-Est, Area CUN 13).

Come già evidenziato, nel contesto delle scienze definite sperimentali, si nota un'incidenza minore di risposte che attribuiscono alla valutazione un'influenza sulla qualità della ricerca. Viene talvolta sottolineata la concordanza e l'aderenza tra modelli di ricerca intrapresi personalmente o legati al contesto disciplinare e criteri di qualità proposti dalla VQR. Una professoresca commenta:

Le mie pubblicazioni sono di elevato livello da molti anni prima dell'inizio della VQR. Quindi ho solo continuato a produrre nello stesso modo (Professore Ordinario, grande università del Sud, Area CUN 6).

Tale concordanza può far sì che l'accademico non rifletta affatto sulla valutazione prima di approcciarsi a un lavoro di ricerca, ma riceva gratificazione *ex post*, una volta ricevuti i risultati:

La mia attività di ricerca e la qualità non ha relazione con le valutazioni. Ma sono d'accordo che questo ha reso gratificante vedere che la mia attività di ricerca era in linea con le scelte della valutazione (Professore ordinario, grande università del Centro, Area CUN 4).

D'altro canto, come detto, l'opinione di mancata influenza della valutazione sulla qualità dei propri lavori, sebbene maggioritaria, non è riscontrabile in tutti i commenti dei docenti appartenenti alle scienze sperimentali. Così commenta un professore:

Le indicazioni ASN e VQR necessariamente peggiorano la qualità. Per pubblicare su certe riviste bisogna per forza aggiungere coreografie ai dati [...], ma senza

quelle non si pubblica. Senza i parametri VQR i dipartimenti e gli atenei prendono meno finanziamenti; quindi, chi non segue questi parametri è poco furbo [...] (Professore associato, grande università del Nord-Ovest, Area CUN 4).

E anche chi afferma di non farsi influenzare, vede una tendenza contraria per i giovani ricercatori, accompagnata da un giudizio netto:

Ho cercato di non farmi influenzare dalle indicazioni della ASN e della VQR, ma vedo intorno a me altri (perlopiù giovani, purtroppo) pensare solo a quello. Il danno è ormai enorme (Ricercatore a esaurimento, grande università del Nord-Est, Area CUN 1).

6. Risultati: modello LOGIT

Per testare l'ipotesi che il campo disciplinare e il contesto istituzionale influenzino la percezione degli effetti della VQR sulla qualità della ricerca degli accademici, è stato utilizzato un modello LOGIT. La regressione logistica è particolarmente adatta a questo scopo perché permette di analizzare variabili binarie o dicotomiche, come la percezione della VQR come influente o non influente sulla qualità della ricerca accademica. Nell'analisi sono state individuate inoltre tre tipologie di variabili: istituzionali, di contesto e individuali. Le variabili istituzionali includono la dimensione dell'università; le variabili di contesto comprendono la geolocalizzazione dell'università di appartenenza del rispondente; mentre le variabili individuali comprendono il genere, la presenza di figli, il ruolo accademico e l'area disciplinare. Le variabili istituzionali e individuali sono particolarmente rilevanti perché aiutano a rispondere all'ipotesi.

In Tab. 2 si analizza l'effetto delle diverse variabili indipendenti sulla percezione degli accademici riguardo all'influenza (positiva o negativa) della VQR sulla qualità del loro lavoro scientifico. Nelle analisi sul ruolo del ricercatore all'interno dell'università, il professore ordinario funge da *benchmark*. Per verificare la robustezza dei risultati, abbiamo effettuato analisi sostituendo il *benchmark* iniziale del "professore ordinario" con le altre categorie, "professore associato" e "ricercatore". Anche con questi cambiamenti, i risultati ottenuti restano invariati, indicando che le relazioni osservate tra le variabili indipendenti e la percezione dell'influenza della VQR sulla qualità della ricerca sono robuste e non dipendono dalla scelta del *benchmark*. Abbiamo scelto il "professore ordinario" perché, oltre a rappresentare la figura accademica di grado più elevato e stabile, consente un confronto efficace con i professori associati, che costituiscono la seconda figura senior nel contesto

universitario. Questo confronto è particolarmente utile per valutare come la percezione dell'influenza della VQR vari anche tra le due figure più senior, distinguendo l'effetto della valutazione rispetto a posizioni che, pur essendo stabili, si differenziano in termini di autonomia e pressione per l'avanzamento. Diverse variabili indipendenti, tra cui genere, presenza di figli, ruolo accademico, dimensione dell'università, geolocalizzazione e area scientifica sono state considerate nell'analisi. Il modello LOGIT è stato apprezzato utilizzando il metodo di classificazione per determinare quanto accuratamente sia in grado di predire le classi dei dati. Il valore del 79,8% nella “Percentuale di Classificazioni Corrette”, evidenziato alla fine della Tab. 2, rappresenta la proporzione di osservazioni totali correttamente classificate dal modello. Questo indica che su tutte le previsioni fatte dal modello, circa l’80% sono risultate corrette rispetto ai dati effettivi. Questo valore sottolinea la capacità del modello di predire con una certa precisione le classi dei dati, riflettendo l'affidabilità dell'accuratezza del modello.

Tab. 2 - Risultati del modello LOGIT

Variabile	Coefficiente (Errore standard)
Maschio	-0,0999 (0,120)
Bambino; Sì, uno	0,0311 (0,137)
Bambino; Sì, più di uno	0,0971 (0,153)
Professore associato	0,0951 (0,129)
Ricercatore	0,426 (0,189) **
Isole	0,515 (0,221) **
Italia nord-est	0,158 (0,169)
Italia nord-ovest	0,182 (0,162)
Sud Italia	0,319 (0,171) +
Media università	0,131 (0,133)
Piccola università	0,349 (0,159) **
Scienze fisiche	-0,865 (0,393) **
Scienze chimiche	-0,292 (0,358)
Scienze della Terra	-0,146 (0,454)
Scienze biologiche	-0,489 (0,300)
Scienze mediche	-0,511 (0,295) +
Scienze agrarie e veterinarie	0,499 (0,325)
Ingegneria civile ed architettura	0,140 (0,314)

Tab. 2 - segue

Variabile	Coefficiente (Errore standard)
Ingegneria industriale e dell'informazione	0,0453 (0,292)
Scienze dell'antichità, filologico-letterarie e storico-artistiche	0,395 (0,296)
Scienze storiche, filosofiche, pedagogiche e psicologiche	0,485 (0,303)
Scienze giuridiche	0,411 (0,307)
Scienze economiche e statistiche	0,741 (0,314) **
Scienze politiche e sociali	-0,0477 (0,361)
Costante	-0,670 (0,313) **
Osservazioni	1.365
Legenda: + p < 0.05, ** p < 0.01	
Percentuale di Classificazioni Corrette	79,8%

Nota: rispetto alle variabili che utilizzano l'Area CUN, l'area disciplinare Matematica e Informatica funge da benchmark.

I risultati mostrano come le diverse variabili esplicative influenzino la probabilità di risposta, ossia la percezione degli accademici rispondenti riguardo gli effetti della VQR sulla loro qualità scientifica. I risultati delineano un panorama complesso sulla percezione degli accademici in merito agli effetti della VQR sulla qualità scientifica del loro lavoro di ricerca. La variabile “Maschio” (utilizzando la categoria “Femmina” come riferimento) non presenta significatività statistica. Questo suggerisce che l’effetto della VQR sulla percezione della qualità scientifica non differisce tra i generi. Analogamente, la variabile “Bambini” non è significativa, indicando che la presenza o assenza di figli non influenza le percezioni degli accademici rispetto all’impatto della VQR sulla loro ricerca. Non sembra esserci quindi un legame diretto tra avere figli e la percezione dell’effetto della VQR sulla qualità scientifica. Questa informazione segnala che le pressioni personali o familiari potrebbero non influenzano significativamente le percezioni sugli effetti della VQR.

La variabile “Area scientifica” mostra invece che i docenti in alcune discipline percepiscono significativamente gli effetti della VQR sulla loro qualità scientifica rispetto ai docenti appartenenti ad altre aree. Le aree disciplinari di Fisica e Medicina hanno un coefficiente negativo significativo, suggerendo che gli accademici in queste aree hanno un’alta probabilità di considerare la VQR non influente rispetto alla qualità scientifica del proprio lavoro. Al contrario, gli accademici in Economia e Statistica tendono

a percepire l'effetto della VQR come significativo. Nelle altre aree disciplinari non si nota alcuna tendenza statisticamente significativa nella considerazione dell'effetto della valutazione.

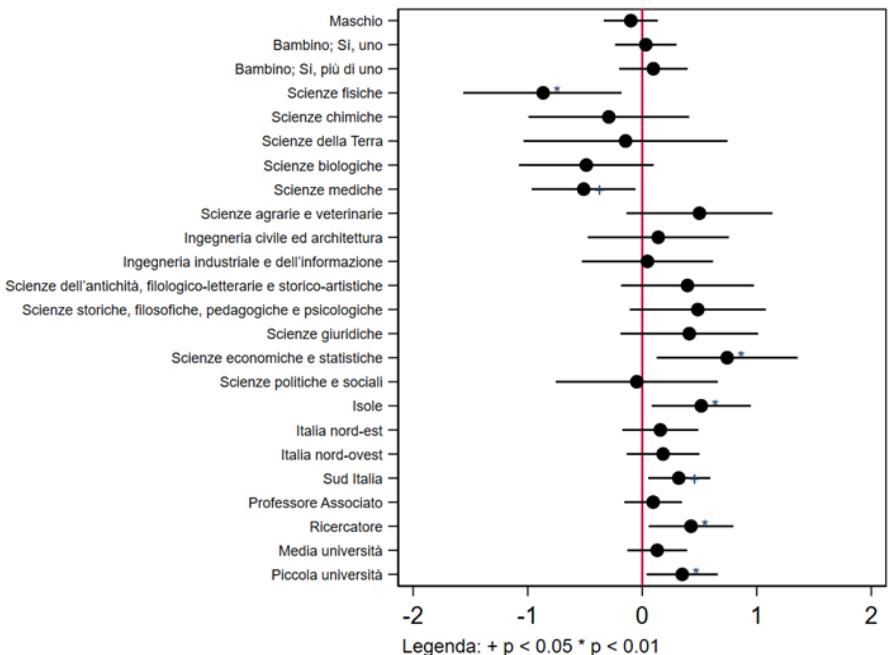
La geografia gioca un certo ruolo. Alcune aree territoriali mostrano una percezione più forte dell'impatto della VQR. Ciò potrebbe riflettere differenze nelle culture accademiche regionali, nelle risorse disponibili o nelle pressioni locali per la pubblicazione e l'acquisizione di finanziamenti. Il ruolo all'interno dell'istituzione accademica emerge come un altro fattore chiave. I ricercatori, rispetto ai professori ordinari (*benchmark*), percepiscono un maggiore impatto della VQR. Ciò potrebbe essere dovuto al fatto che i ricercatori, essendo spesso in una fase iniziale della loro carriera, potrebbero sentirsi più pressati dalle metriche di valutazione come la VQR, rispetto ai professori ordinari e associati che potrebbero avere una posizione più consolidata e sicura.

Infine, la dimensione dell'università sembra contare. Gli accademici provenienti da università più piccole sentono l'effetto della VQR in modo più acuto. Questo potrebbe suggerire che in istituzioni di dimensioni ridotte, dove le risorse sono più limitate, la competizione per queste risorse potrebbe essere più intensa e la VQR potrebbe quindi assumere un'importanza maggiore. Questa maggiore percezione dell'effetto della VQR potrebbe essere legata a questioni di controllo sociale e *governance*, che variano significativamente con la dimensione dell'organizzazione. In università più piccole, la struttura di *governance* è spesso più centralizzata e il controllo sociale più stretto, il che potrebbe portare a una maggiore pressione sugli accademici per conformarsi alle direttive della VQR. Al contrario, nelle università più grandi, una struttura di *governance* più distribuita e una maggiore diversità di risorse e opportunità potrebbero attenuare l'effetto percepito della VQR, permettendo una maggiore autonomia individuale e un ambiente meno competitivo.

Il termine costante rappresenta l'intercetta del modello, e in questo caso rappresenta il *log-odds* della variabile dipendente quando tutte le variabili indipendenti sono pari a zero.

Al fine di fornire una visualizzazione chiara e sintetica di questi risultati, la Fig. 3 presenta un grafico degli intervalli di confidenza per le *odds ratio* di ciascuna variabile. Questo grafico consente di identificare facilmente quali variabili hanno effetti statisticamente significativi. Va notato che, anche se la costante (o intercetta) del modello è stata omessa dal grafico per ragioni di chiarezza, essa è stata inclusa nei calcoli e nell'analisi complessiva (come mostrato nella Tab. 2). La Fig. 3 è da considerarsi, quindi, come uno strumento visivo complementare per comprendere meglio le dinamiche presentate nei risultati del modello LOGIT.

Fig. 3 - Rappresentazione grafica del modello LOGIT e intervallo di confidenza

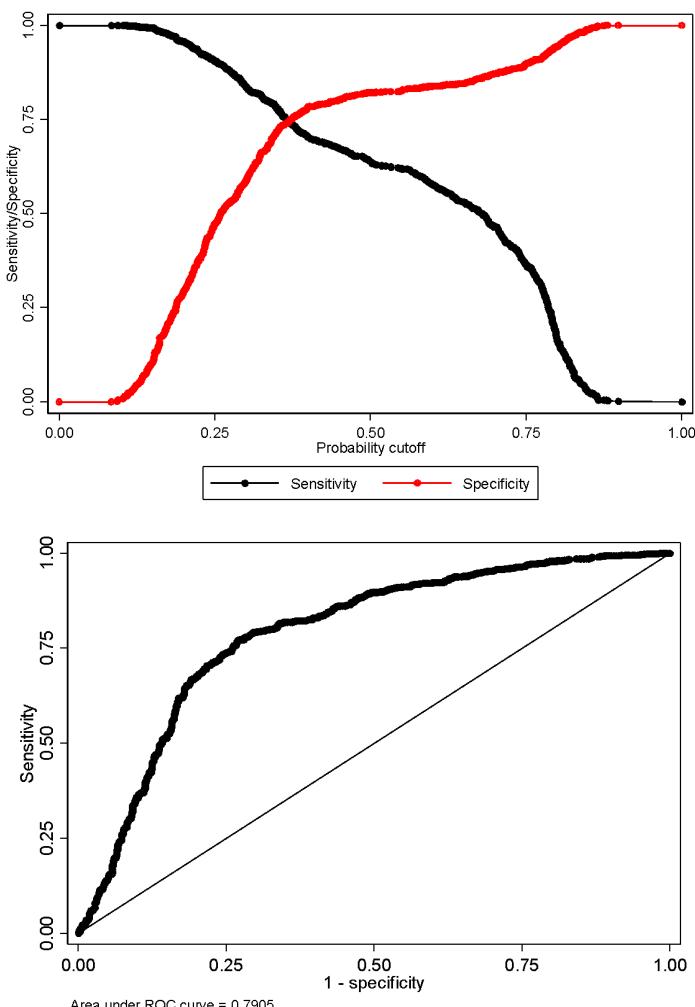


A causa delle caratteristiche proprie e dei limiti tipici degli approcci *survey*, è comune riscontrare una maggiore incertezza e, di conseguenza, una maggiore varianza nei dati raccolti. Questo può spiegare perché gli intervalli di confidenza mostrati nella Fig. 3 appaiano relativamente ampi.

L'adeguatezza del modello LOGIT adattato è stata valutata utilizzando procedure diagnostiche standard, quali i grafici di sensibilità/specificità e la curva ROC (Receiver Operating Characteristic). Un elemento fondamentale nella valutazione della performance del modello LOGIT è la comprensione della sua sensibilità e specificità a vari livelli di probabilità *cut-off*, ovvero il valore soglia utilizzato per classificare il risultato della nostra variabile dipendente “Effetto della VQR” in base a una probabilità predetta. La sensibilità misura la capacità del modello di identificare correttamente i “casi positivi” (cioè quando il modello prevede correttamente che un evento accada, nel nostro caso, che la VQR abbia influito sulla qualità della ricerca). La specificità, invece, indica la capacità del modello di identificare correttamente i “casi negativi” (cioè quando il modello prevede correttamente che un evento non accada, nel nostro caso, che la VQR non abbia influito). La curva ROC mostra la sensibilità rispetto al tasso di falsi positivi (quando il modello prevede un evento che in realtà non si verifica).

a vari livelli di soglia (“cut-off”), permettendo di valutare la precisione complessiva del modello. L’area sotto la curva ROC (AUC) quantifica la capacità discriminante del modello: un valore vicino a 1 indica un’elevata capacità di classificazione, mentre un valore vicino a 0,5 suggerisce una previsione casuale. La Fig. 4a rappresenta un grafico della sensibilità/specificità in funzione del punto di cut-off della probabilità. Le due curve si incrociano abbastanza vicino all’asse verticale. Questo suggerisce che il modello adattato è adeguatamente sensibile e specifico.

Fig. 4 - (4a) Sensibilità/specificità rispetto al punto di cut-off della probabilità e (4b) Area sotto la curva ROC



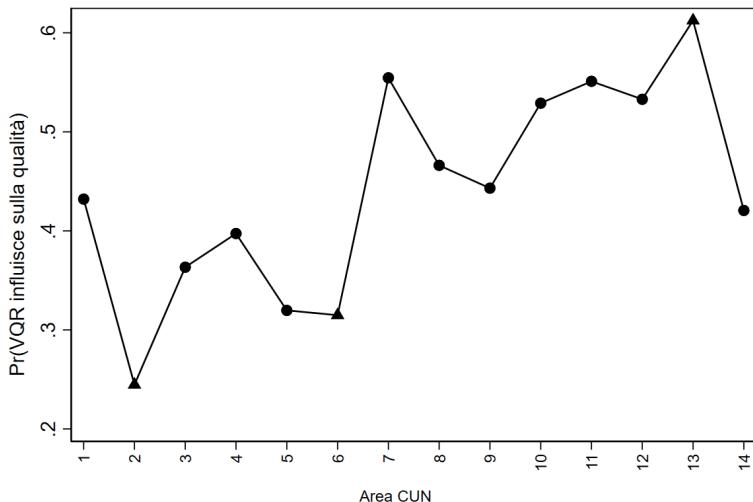
Il grafico ROC è uno strumento fondamentale per visualizzare e valutare la capacità del nostro modello di discriminare tra le 2 classi in esame. Un'area sotto la curva ROC vicina a 1 indica un modello di previsione eccellente, mentre un'area vicina a 0,5 suggerisce un modello che non ha capacità discriminatoria migliore di una scelta casuale. La Fig. 4b presenta la Curva ROC (Receiver Operating Characteristic) rispetto al modello LOGIT presentato in Tab. 2. L'ampiezza dell'area che giace sotto il grafico ROC rappresenta una misura della variazione spiegata dal modello LOGIT adattato. In questo specifico caso, l'area sotto il grafico ROC è del 79%, un valore che supera significativamente il 75% (ritenuto dalla letteratura la soglia minima di accettazione) (Hosmer e Lemeshow, 2000).

La proporzione di varianza spiegata e la poca proporzione di varianza non spiegata indicano che il modello adattato è affidabile nello spiegare la variabilità della risposta “Effetto della VQR” in funzione delle variabili esplicative. Tuttavia, è importante sottolineare che, trattandosi di dati derivanti da un'indagine, la varianza delle risposte fornite risulta elevata. Questo aspetto comporta che il grafico non sia perfettamente delineato. Nonostante ciò, il modello fornisce comunque un quadro significativo e molto informativo della situazione, pur tenendo conto delle imperfezioni dovute alla natura stessa dei dati raccolti tramite *survey*.

La Fig. 5 rappresenta le probabilità marginali per le 14 aree scientifiche CUN. Si nota una marcata variabilità tra discipline. In particolare, le aree relative a Fisica e Medicina mostrano probabilità sensibilmente inferiori rispetto alle altre, indicando che gli accademici appartenenti a queste aree hanno una minore probabilità di affermare che la VQR abbia avuto un impatto positivo o negativo sulla qualità della loro ricerca. Al contrario, le Scienze Economiche e Statistiche registrano una probabilità superiore ($>$ di 0,60), suggerendo una maggiore tendenza a ritenere che la VQR abbia influenzato negativamente la loro qualità di ricerca. Nella Fig. 5, i cerchi indicano le aree CUN in cui l'effetto della VQR sulla qualità della ricerca non è statisticamente significativo, mentre i triangoli rappresentano le aree in cui questo effetto è significativo. Le aree CUN con probabilità non statisticamente significativa (indicate nel grafico con il simbolo del cerchio) sono state escluse dall'interpretazione. Tuttavia, risulta evidente che le Scienze Sociali e Umane (dall'area CUN 10 in avanti) mostrano in media una maggiore probabilità di ritenere che la VQR abbia influenzato la qualità della loro ricerca.

La Fig. 5, come detto, evidenzia una marcata variabilità tra le diverse discipline. In particolare, si nota una differenza notevole tra le aree relative a Fisica e Medicina (rispettivamente area CUN 2 e 6) rispetto alle Scienze Economiche e Statistiche (area CUN 13). Le Scienze Fisiche e della Vita, tradizionalmente ancorate alla pratica di pubblicare *paper* scientifici su

Fig. 5 - Probabilità marginali per Area Scientifica CUN



riviste indicizzate, sembrano risentire meno dell'effetto della VQR sulla percezione della qualità della ricerca. Questa tendenza potrebbe derivare dal fatto che tali discipline si sono stabilizzate da tempo attorno a pratiche di pubblicazione non diverse da quelle premiate dalla VQR, che sono riconosciute dalla comunità epistemica come principale veicolo di circolazione dei risultati ottenuti attraverso il lavoro scientifico.

Le Scienze Economiche e Statistiche mostrano invece una probabilità decisamente superiore. Questo potrebbe suggerire che queste aree, che tradizionalmente potrebbero aver adottato diverse forme di circolazione oltre all'articolo scientifico su rivista internazionale, hanno dovuto affrontare una maggiore sfida nell'adeguarsi ai criteri della VQR. L'evoluzione e l'adattamento delle pratiche di ricerca e pubblicazione in questi campi potrebbero aver influenzato la percezione dell'impatto della VQR sulla qualità della loro ricerca.

7. Conclusioni

Alla luce dei risultati ottenuti, il nostro studio ha dimostrato una notevole variabilità nella percezione della VQR tra gli accademici italiani, influenzata da una serie di fattori disciplinari e istituzionali. Alcuni di questi risultati sono in linea con studi precedenti che hanno sottolineato l'importanza della disciplina e del contesto istituzionale nella modulazio-

ne della percezione della qualità da parte degli accademici (Bratti *et al.*, 2021; Demetrescu *et al.*, 2020b). Tuttavia, al di là dei risultati che confermano le evidenze della letteratura esistente, il nostro studio apporta nuovi elementi di discussione.

In primo luogo, le variabili demografiche come il genere e la presenza di figli non sembrano influenzare significativamente la percezione della VQR sulla qualità scientifica. Questo potrebbe indicare che le pressioni o i vantaggi percepiti dalla VQR sono simili per uomini e donne nel mondo accademico, suggerendo un terreno comune in termini di valutazione della ricerca. La variabile relativa alla presenza di figli non ha evidenziato correlazioni, suggerendo che gli impegni familiari o le responsabilità legate all'avere figli non alterano sostanzialmente la percezione della VQR tra gli accademici.

Invece, l'area disciplinare, la posizione geografica, il ruolo accademico e la dimensione dell'istituzione hanno mostrato variazioni nella percezione. Ad esempio, mentre discipline come la Fisica tendono a percepire meno l'influenza della VQR, aree disciplinari come Economia e Statistica ne riconoscono un impatto più pronunciato. Queste differenze sottolineano l'importanza di considerare le specificità di ciascuna comunità accademica nella valutazione degli effetti della VQR. Anche gli accademici in Medicina tendono a percepire un impatto minore della VQR sulla loro ricerca, potenzialmente a causa delle caratteristiche intrinseche di tale disciplina.

La Fisica e la Medicina vantano una lunga tradizione di pubblicazioni su riviste scientifiche indicizzate; pertanto, la VQR che privilegia metriche associate a indici citazionali anche come integrazione alla *peer review*, potrebbe guidare pratiche non dissimili da quelle già presenti. Invece, aree disciplinari come l'Economia e la Statistica sembrano più influenzati dalla VQR.

Ciò potrebbe essere dovuto a un maggiore sforzo di adattamento ai criteri valutativi della VQR, che avrebbe portato gli studiosi a rivedere, in alcuni casi, le loro metodologie e approcci di ricerca. La triangolazione tra le analisi quantitative e l'analisi dei commenti liberi forniti dagli accademici durante la compilazione del questionario ha ulteriormente evidenziato l'importanza del contesto disciplinare nel modellare percezioni differenziate rispetto agli effetti della VQR sulla qualità della ricerca. Coloro che hanno scelto di integrare le risposte chiuse con commenti hanno in genere sollevato perplessità riguardo a eventuali esiti positivi scaturiti dagli esercizi valutativi. Maggiori critiche sono state espresse da rappresentanti delle Scienze Umane e Sociali, mentre quelli delle scienze sperimentali hanno mantenuto un atteggiamento più neutrale, limitando i toni polemici.

Anche la posizione geografica dell'affiliazione accademica ha mostrato variazioni nella percezione dell'influenza della VQR. Ciò potrebbe riflette-

re differenze culturali, disponibilità di risorse o priorità di ricerca tra diverse regioni. Il ruolo accademico si è rivelato un altro fattore discriminante. I ricercatori, che sono all'inizio della loro carriera o in posizioni meno stabili, sembrano percepire la VQR come un fattore più influente rispetto ai professori ordinari. Questo potrebbe indicare che la VQR ha un impatto maggiore su coloro che sono ancora in fase di consolidamento della loro posizione accademica.

Infine, la dimensione dell'università ha mostrato che gli accademici di atenei più piccoli percepiscono la VQR come un fattore più determinante nella valutazione della qualità della loro ricerca rispetto a quelli di istituti più grandi. Ciò potrebbe essere dovuto al fatto che in università più piccole, le metriche di valutazione esterne come la VQR possono avere un peso maggiore nel determinare finanziamenti, riconoscimenti e opportunità.

I risultati del presente Capitolo mostrano differenze, soprattutto rispetto alla comunità scientifica di appartenenza e al contesto istituzionale e geografico in cui si lavora. Comprendere i diversi ambiti in cui la VQR viene esercitata porterebbe inevitabilmente a un sistema di valutazione più equilibrato, in grado di rispecchiare accuratamente la realtà del contesto accademico. Questo è essenziale per garantire che la valutazione sia non solo precisa, ma anche equa. Inoltre, comprendere la percezione degli accademici riguardo agli effetti della VQR sulla loro qualità scientifica può guidare i *policy maker* nel rafforzare l'accettazione e la fiducia nel sistema di valutazione. Una visione chiara e accurata della percezione della VQR tra gli accademici non solo offre informazioni preziose ai *policy maker*, ma fornisce anche gli strumenti necessari per apportare modifiche mirate nei futuri esercizi di valutazione. Queste modifiche, basate su *feedback* reali, possono rendere il processo di valutazione molto più trasparente ed equo, assicurando che sia in sintonia con gli obiettivi a lungo termine della ricerca scientifica in Italia e delle comunità accademiche.

La VQR, come ogni altro strumento di valutazione, è in evoluzione continua. È essenziale che essa venga aggiornata regolarmente per riflettere le dinamiche mutevoli del panorama accademico e delle esigenze di ricerca. Per il prossimo ciclo di valutazione (2020-2024), emergono indicazioni di potenziali cambiamenti al sistema, con un focus crescente sul trasferimento tecnologico, la capacità di attrarre finanziamenti internazionali, l'instaurazione di rapporti di collaborazione e lo scambio con enti di ricerca, sia pubblici che privati².

2. www.mur.gov.it/sites/default/files/2023-08/Decreto%20Ministeriale%20n.%20998%20del%2001-08-2023.pdf.

Dato che la VQR influisce direttamente sull’assegnazione della quota premiale del Fondo di Finanziamento Ordinario (FFO), è necessario che le università e gli istituti di ricerca siano proattivamente informati e pronti ad adeguarsi a questi cambiamenti. Devono garantire che le loro iniziative e strategie siano allineate con questi nuovi parametri valutativi. Un’insufficiente preparazione o adesione a tali criteri potrebbe infatti tradursi in ulteriori squilibri a livello istituzionale e geografico, provocando effetti di accumulazione o disparità nelle risorse e nelle opportunità di ricerca, con potenziali ripercussioni sulla qualità e l’efficacia della ricerca stessa in Italia.

2. Le scelte collegate alla carriera dei giovani ricercatori: storie di vita e narrazioni sugli effetti della valutazione

di *Serena Fabrizio, Ugo Finardi, Lucio Morettini,
Emanuela Reale*

1. Introduzione

Negli ultimi decenni del ventesimo secolo l'attenzione verso la valutazione delle politiche relative all'università e alla ricerca scientifica e tecnologica è cresciuta nel dibattito pubblico tra decisori politici, stakeholder e opinione pubblica in generale (Shriberg, 2002; Maassen e Stensaker, 2011; Reale e Seeber, 2013). Questa attenzione è collegata ad un rinnovato interesse, in diversi Paesi europei, nei confronti di politiche orientate a promuovere l'efficacia dell'organizzazione delle attività universitarie e il contributo che le università possono portare al benessere economico e sociale (Martin, 2000; Gläser e Whitley, 2007; Martin, 2016).

La promozione di politiche universitarie efficaci ha comportato, di pari passo, la creazione di strumenti utili ai decisori politici per valutare sia la qualità scientifica del lavoro svolto dagli accademici sia il suo impatto sociale, per poter quindi allocare le risorse in modo da migliorare il sistema universitario nel suo complesso (Aagaard *et al.*, 2015; van den Besselaar, *et al.*, 2017). Peraltro, la presenza di questi strumenti non è senza effetto: diversi studi infatti hanno rilevato che la valutazione ha un'influenza diretta sull'attività dei ricercatori, con risultati non univoci sulla qualità della ricerca (Gläser e Laudel, 2016; Kivistö e Kohtamäki, 2016; Karlsson, 2017; Kivistö e Mathies, 2023).

In questo capitolo intendiamo focalizzare l'attenzione sui giovani accademici italiani soggetti in particolare all'Abilitazione Scientifica Nazionale (ASN), il processo di valutazione della produzione scientifica dei ricercatori, propedeutico alla progressione di carriera, basato sul confronto della situazione individuale rispetto al quadro nazionale del settore scientifico di riferimento attraverso alcuni indicatori quali il numero di articoli pubblica-

ti, il numero di citazioni (settori scientifici c.d. bibliometrici) e il numero di articoli pubblicati su riviste selezionate (settori c.d. non bibliometrici).

Il capitolo presenta i principali risultati emersi da un'indagine, rivolta ai ricercatori a tempo determinato delle università italiane, che ha adottato un approccio narrativo per esplorarne le aspettative e le dinamiche di carriera. Le domande di ricerca cui si cerca di dare risposta sono: come la valutazione esterna influenza il vissuto accademico e professionale dei giovani ricercatori? Quale è l'importanza della valutazione rispetto ad altri fattori che influenzano le scelte professionali e di carriera dei giovani ricercatori?

Le evidenze empiriche sono state raccolte attraverso delle “storie di vita” per far emergere le esperienze personali dei giovani ricercatori e ricercatrici, offrendo una prospettiva dettagliata sulla capacità della valutazione esterna di influenzare le loro scelte di carriera.

La metodologia utilizzata include la definizione di un profilo dettagliato degli intervistati, considerando indicatori come “genere”, “posizione accademica”, “dimensioni dell'università” e “area geografica”, per garantire una distribuzione equilibrata. L'analisi documentale del percorso accademico degli intervistati precede l'approfondimento attraverso le Storie di Vita con approccio *informato* (Bichi, 2000; Laudel e Gläser, 2007) per facilitare il coinvolgimento emotivo dell'intervistatore con l'intervistato.

Sono state condotte 18 Storie di Vita rivolte ai ricercatori a tempo determinato degli atenei italiani. I soggetti intervistati sono ricercatori che hanno intrapreso la carriera accademica e che hanno sviluppato già alcuni anni di esperienza di ricerca durante il dottorato ed eventualmente in posizione post-doc, ma che rivestono ancora la posizione di Ricercatori a Tempo Determinato (RTD)¹.

2. Contesto di riferimento dello studio e obiettivi conoscitivi

L'impatto della valutazione della ricerca è un tema che ha acquisito una crescente centralità in letteratura nel corso del tempo, data la sua natura poliedrica e complessa. Da un lato gli strumenti di valutazione ester-

1. Gli RTD sono figure professionali a tempo determinato presenti nell'ordinamento accademico italiano; il contratto prevede una durata triennale dell'incarico al termine del quale è possibile ottenere un altro incarico a tempo determinato (profilo a, RTDa) o essere inquadrati a tempo indeterminato come professori associati qualora il profilo del ricercatore soddisfi i parametri dell'Abilitazione Scientifica Nazionale (ASN) (profilo b, RTDb). Anche se le posizioni lavorative degli RTDa e RTDb non sono definitive, sono anch'essi soggetti alla Valutazione della Qualità della Ricerca (VQR).

ni sono un elemento che concretizza e struttura le politiche per la ricerca, dato il loro valore direzionale rispetto ai temi di ricerca (Gläser e Laudel, 2016), l'impatto che hanno sulle carriere dei ricercatori (Altbach e Lewis, 1996; Moscati, 2001; Rostan, 2011) e la possibilità di alterare sensibilmente l'identità delle università e degli enti di ricerca coinvolti (Kivistö e Kohtamäki, 2016; Karlsson, 2017). Parallelamente a questi effetti di contesto, diversi lavori hanno approfondito l'impatto della valutazione su produttività scientifica e modalità di pubblicazione, con risultati che presentano elementi controversi.

Diversi contributi sottolineano come la valutazione abbia portato ad un diffuso aumento quantitativo di lavori pubblicati (Butler, 2003b; van den Besselaar *et al.*, 2017; Gläser e Laudel, 2016) e ad un maggiore interesse verso pubblicazioni di tipo internazionale che hanno ampliato gli orizzonti dei ricercatori (Hammarfelt e de Rijcke, 2015). All'interno di questo quadro generale emergono sfumature legate ai diversi effetti sulla qualità della ricerca. Da un lato, infatti, esistono analisi che mostrano come ad un aumento del numero di pubblicazioni si associa un generale aumento delle citazioni, utilizzate come approssimazione della qualità del lavoro (Larivière e Costas, 2016; van den Besselaar *et al.*, 2017); dall'altro, viene sottolineata una duplice generale tendenza: quella a produrre lavori con incrementi di conoscenza minimali su temi già consolidati (Coulthard e Keller, 2016) e quella relativa a potenziali effetti di indebolimento delle comunità nazionali di ricerca causati da adesioni repentine e radicali a nuovi standard internazionali e all'abbandono di temi e modalità di pubblicazione proprie dei sistemi nazionali, decretandone di fatto la scomparsa (Mathies *et al.*, 2020).

Nell'ambito di questo filone di letteratura negli ultimi anni hanno guadagnato spazio le analisi degli effetti della valutazione sui ricercatori ad inizio carriera, una classe di ricercatori che presenta di per sé diverse fragilità connesse alla fase iniziale della carriera: fra tutte, la necessità di guadagnare rapidamente reputazione e visibilità all'interno della comunità scientifica di riferimento, la necessità di accumulare un buon numero di pubblicazioni riconosciute, di allargare il proprio network e di sviluppare una capacità autonoma di attrarre fondi per ricerca. D'altra parte, se l'adesione ai criteri e agli indicatori usati dalla valutazione esterna implementata dal governo può risultare più facile per questi soggetti, dato che ancora non hanno sviluppato un *modus operandi* consolidato, può allo stesso tempo influenzare in maniera netta lo sviluppo della loro attività di ricerca.

Infatti, alcuni autori, tra quelli che analizzano l'impatto della valutazione sulle pratiche di pubblicazione dei ricercatori, mettono in luce come esista una discrepanza tra risultati relativi a giovani ricercatori e quelli re-

lativi agli accademici più esperti, con i primi che hanno incrementi meno consistenti nel numero di pubblicazioni, ma in proporzione registrano una maggiore propensione ad ampliare il numero di citazioni (Aagaard *et al.*, 2015; Larivière e Costas, 2016). Su impulso delle pratiche di valutazione esterne, i giovani ricercatori tendono inoltre ad adottare pratiche di pubblicazione strumentali, concentrandosi cioè sullo sviluppo di specifiche competenze di ricerca utilizzabili su più lavori contemporaneamente (Johann *et al.*, 2024), favorendo lo sviluppo di metodologie di analisi quantitative a scapito di metodologie qualitative (Coulthard e Keller, 2016).

Gli effetti della valutazione sulla carriera dei giovani ricercatori appaiono comunque più complessi di quanto possa essere rilevato dagli indicatori bibliometrici. La spinta a cercare di realizzare pubblicazioni con maggiore impatto spinge i ricercatori a concentrare i loro sforzi in settori consolidati, abbandonando lo sviluppo di nuove aree di ricerca (Johann *et al.*, 2024). Per motivi analoghi, i ricercatori ad inizio carriera sono scoraggiati ad intraprendere percorsi di ricerca interdisciplinari che difficilmente aderiscono ai parametri di valutazione *mainstream* (Sobey *et al.*, 2013; Mäkinen *et al.*, 2024) così come sono scoraggiati ad impostare le proprie attività di ricerca in modo da ottenere risultati trasferibili verso stakeholders, ma scarsamente considerati dall'ambiente accademico (Wróblewska *et al.*, 2023).

Dalla letteratura emerge che i giovani ricercatori subiscono in maniera più accentuata gli effetti della valutazione, al punto che questi diventano un elemento essenziale nella formazione dell'immagine professionale che i ricercatori danno di sé, portandoli a frustrare le proprie aspirazioni e il proprio talento per cercare di avviare e consolidare la loro carriera (Lindahl *et al.*, 2020; Mula-Falcón e Caballero, 2023) e a sviluppare un rapporto simbiotico con il proprio mentore (Van Balen *et al.*, 2012). Ciò accade senza che entrambi questi elementi siano effettivamente risolutivi per risolvere il rapporto tra valutazione e percorso di ricerca. Sebbene sia largamente confermato che la valutazione esterna abbia un effetto diretto sul modo di fare ricerca, tanto più per i ricercatori ad inizio carriera, i meccanismi che attivano questi effetti non sono ancora del tutto chiariti, poiché connessi al campo di studi, alle consuetudini nelle modalità di produzione della conoscenza (Gläser e Laudel, 2017) e ai contesti accademici in cui ricercatori si trovano a lavorare (Thomas e Nedeva, 2012).

Il presente capitolo si inserisce in questo quadro di riferimento, focalizzandosi sugli effetti prodotti dagli strumenti di valutazione previsti nel sistema della ricerca italiano sui ricercatori ad inizio carriera.

3. Dimensione narrativa degli effetti della valutazione: Storie di Vita “informate”

Le interviste, nella forma più o meno strutturata, rappresentano uno degli strumenti principali utilizzati in studi qualitativi. Questi studi possono riguardare, ad esempio, la comprensione delle relazioni tra l'effettiva implementazione di una politica o di un regolamento e i conseguenti effetti sul vissuto e sulle decisioni individuali.

Nardon *et al.* (2021) hanno sperimentato l'utilizzo delle interviste come strumento di riflessione per cogliere l'impatto sociale delle politiche, proponendo tecniche di intervista per facilitare la riflessione dei partecipanti su questo tema specifico. In particolare, il loro studio evidenzia come le ‘interviste riflessive’² possano aiutare i ricercatori a esplorare e comprendere in modo più approfondito le conseguenze delle politiche di valutazione sulla propria attività e carriera accademica, promuovendo un dialogo che può portare a un miglioramento delle pratiche e a una maggiore consapevolezza degli effetti sociali e istituzionali delle politiche stesse. La necessità di utilizzare questo tipo di tecniche è più evidente nei casi in cui le finalità di studio siano legate a fenomeni difficilmente rilevabili con metodi quantitativi, come nel caso del vissuto o delle dinamiche all'interno di un sistema o di una comunità rispetto a un dato evento, o nelle quali il confronto tra aree vicine ma comunque diverse tra loro non può essere riconducibile a indicatori univoci, come potrebbe essere ad esempio il caso di aree di ricerca diverse all'interno dello stesso sistema accademico (Gibson e Hazelkorn, 2017).

L'approccio biografico-narrativo utilizzato in questo studio, fa riferimento sia alla metodologia delle interviste biografiche (Bertaux, 1981, Bichi, 2000), che si focalizza sulla raccolta di informazioni sulla vita di un individuo attraverso interviste mirate, sia alla *Critical Communicative Methodology* (CCM), utilizzata nelle scienze sociali, in particolare nel campo della ricerca qualitativa (Flecha *et al.*, 2011, Gómez, 2017).

L'intervista biografica si concentra sulla Storia di Vita dell'individuo, esplorando eventi passati, decisioni cruciali e cambiamenti significativi. Si

2. Secondo Nardon *et al.* (2021) l'intervista riflessiva si distingue dalle tecniche tradizionali poiché incoraggia i partecipanti a riflettere sulle proprie esperienze e a considerare nuove prospettive. Questo approccio non solo raccoglie dati, ma promuove anche una maggiore consapevolezza e comprensione tra i partecipanti, potenziando l'impatto sociale della ricerca. Per implementare efficacemente l'intervista riflessiva, i ricercatori devono creare un ambiente sicuro e aperto che favorisca la riflessione e il dialogo. È fondamentale che l'intervistatore sia empatico e flessibile, adattando le domande in base alle risposte dei partecipanti per approfondire le loro riflessioni.

distingue per l'orientamento alla vita personale, al contesto e al significato personale attribuito alle esperienze (Bertaux, 1981; Bichi, 2002). Segue un approccio narrativo, riflettendo l'importanza di una ricostruzione retrospettiva degli eventi, e incoraggia la riflessività e l'auto interpretazione da parte dell'intervistato. L'approccio è flessibile e mira a individuare *pattern* e connessioni nella vita dell'intervistato.

La CCM si concentra sull'importanza della comunicazione e del linguaggio nelle dinamiche sociali e cerca di analizzare in modo critico le interazioni linguistiche. Questo approccio riconosce il potere del linguaggio nel plasmare le prospettive e le esperienze delle persone. La metodologia mira a esplorare come il linguaggio contribuisca alla costruzione del significato sociale e alle dinamiche di potere all'interno di una comunità o di un contesto specifico.

La CCM incoraggia inoltre una partecipazione attiva e consapevole dei partecipanti alla ricerca, cercando di incorporare le loro voci e le loro prospettive nel processo di analisi. In questo modo, si cerca di andare oltre la mera raccolta di dati per comprendere in modo più approfondito le relazioni sociali e le sfide che emergono attraverso la comunicazione dialogica.

La dimensione biografica, unita a quella dialogica propria della CCM nella conduzione delle Storie di Vita *informate*, prevede di invitare i partecipanti a raccontare eventi passati e presenti relativi all'argomento oggetto della ricerca, senza esplicitarlo, puntualizzando con gli intervistati solo “temi di conversazione”, evidenziando elementi significativi e formulando ipotesi su aspettative future. La componente di analisi temporale è importante in quanto ai partecipanti viene sia chiesto di identificare l'evoluzione delle loro esperienze passate sia di formulare ipotesi legate alle prospettive per il futuro (Sanders e Munford, 2005; Gómez *et al.*, 2011; Brown *et al.*, 2013; Puigvert, 2012).

In questo modo è possibile aiutare chi svolge la ricerca a ricostruire un modello di riferimento, allo scopo di isolare gli elementi che influenzano direttamente le esperienze personali o che, a loro volta, contribuiscono a modificare la struttura del modello (Tellado *et al.*, 2020; Ruiz-Eugenio *et al.*, 2020; Prieto-Flores *et al.*, 2016).

Raccogliendo le testimonianze dirette degli accademici direttamente coinvolti nell'implementazione delle decisioni politiche, le Storie di Vita (SDV) si rivelano cruciali nello svelare i meccanismi interni di trasmissione delle informazioni; questo processo facilita l'identificazione dei punti critici degli strumenti di intervento esterni o le politiche (Lo *et al.*, 2019).

Lo strumento delle SDV si dimostra particolarmente utile per esplorare il possibile legame tra l'applicazione di una politica di valutazione e gli effetti che questa può avere sull'esperienza personale dei ricercatori che ne sono soggetti. Il prossimo paragrafo presenta le principali fasi di costruzione del protocollo SDV.

4. Dalla definizione delle dimensioni analitiche ai “temi di conversazione”: il protocollo delle Storie di Vita

La valutazione della ricerca può avere effetti diversi sul lavoro accademico: essi dipendono dal contesto istituzionale e geografico in cui operano gli accademici, dalle aree a cui appartengono, dalla fase della loro carriera (junior o senior) e dalle tematiche di genere. Le Storie di Vita (SDV) sono quindi state disegnate per raccogliere elementi utili a osservare e comprendere se e come gli elementi di valutazione si inseriscono nella vita lavorativa dei ricercatori e se/quanto ne hanno influenzato la strutturazione, lo sviluppo e le prospettive di medio e lungo termine.

Sono stati indagati quattro possibili effetti prodotti sulle dimensioni analitiche, a partire dalle domande di ricerca generali del progetto (in corsivo) e dall’analisi della letteratura di riferimento (cfr. anche Capitolo 1):

- *Decisioni che influenzano il contenuto delle attività di ricerca.* Ad esempio, l’identificazione di problemi scientifici rilevanti, la metodologia utilizzata, la prospettiva di lungo periodo o di breve periodo, l’interdisciplinarità, la copertura, ecc. (Santos e Horta, 2018). Questo aspetto è stato analizzato anche da Karlsson (2017) e Hessels e van Lente (2008), i quali esplorano come le scelte di ricerca possono orientarsi verso tematiche ritenute più rilevanti o finanziabili.
- *Modalità di codificazione delle conoscenze prodotte attraverso l’attività di ricerca.* Osservando i cambiamenti nelle strategie e nelle pratiche di pubblicazione, come il tipo di pubblicazione, la lingua, la produttività, e la tempistica, Hessels e van Lente riflettono sull’impatto della valutazione sulle modalità di pubblicazione, notando come le pressioni per pubblicare frequentemente possano influenzare la scelta delle riviste e la qualità delle pubblicazioni (2008).
- *Cambiamenti nelle relazioni tra ricercatori a inizio carriera e senior.* Eliminando o attenuando la condizione di dipendenza scientifica nei rapporti di autorità, ed effetti sui temi di ricerca e sulle modalità di pubblicazione. Santos e Horta (2018) esaminano come la valutazione possa ridurre la dipendenza dei giovani ricercatori dai senior, promuovendo una maggiore autonomia nelle scelte di ricerca.
- *Collaborazioni internazionali, mobilità e reti di ricerca.* La mobilità accademica e le collaborazioni sono elementi centrali nello sviluppo delle carriere accademiche e sono influenzate dalle politiche di valutazione, come discusso in uno studio di Minelli *et al.* (2006) che mette in luce come la valutazione può incentivare la mobilità internazionale e la partecipazione a reti di ricerca globali.

È stato quindi costruito un protocollo delle SDV informate basato su uno schema concettuale nel quale sono state sintetizzate le dimensioni di analisi individuate a partire dalle riflessioni di alcuni autori in particolare (Santos e Horta, 2018; Hessels e van Lente, 2008; Karlsson, 2017; Minelli *et al.*, 2006). Le dimensioni fanno riferimento a due livelli a cui è possibile ricondurre gli effetti della valutazione sulle pratiche accademiche:

- *Individuale*. Riguarda la produzione di conoscenza, le pratiche di pubblicazione, e l'organizzazione delle attività di ricerca. Hessels e van Lente (2008) esplorano come la valutazione possa orientare le scelte di ricerca individuali, influenzando le priorità e i metodi utilizzati. Santos e Horta (2018) sottolineano come le valutazioni possano favorire l'autonomia dei giovani ricercatori nelle loro scelte di ricerca, riducendo la dipendenza dai ricercatori senior.
- *Organizzativo*. Include il contesto accademico di riferimento, il management della ricerca, la gestione delle risorse finanziarie, e il sistema di potere (Tirado *et al.*, 2022). Minelli *et al.* (2006) discutono l'impatto delle valutazioni a livello organizzativo, evidenziando come queste possano influenzare la gestione delle risorse e le politiche interne delle istituzioni accademiche.

In fase di selezione degli intervistati, le dimensioni di analisi sono state utilizzate per costruire la scheda di profilo propedeutica allo svolgimento delle SDV.

Tab. 1 - Dimensioni analitiche delle Storie di Vita

Ambiti di analisi	Elementi di analisi
Percorso Accademico	<i>Influenza del mentore</i> (Laudel e Gläser, 2007; Santos e Horta, 2018): individuare ruolo di mentore e influenza di tale figura nell'attività di ricerca. <i>Determinanti della strategia di pubblicazione</i> (Hendrik e van Dalen, 2012; Karlsson, 2017): motivazioni che guidano le scelte di pubblicazione rispetto alla volontà di presentare un percorso di crescita scientifica individuale, alle esigenze dettate dal gruppo di ricerca di cui si è parte o alla preferenza, innata o indotta, per la costanza del numero di pubblicazioni indipendentemente dall'effettivo sviluppo dell'attività di ricerca (<i>publish or perish</i>).
Attività/interessi di ricerca	<i>Multidisciplinarietà</i> (Maassen e Stensaker, 2011; Hessels e van Lente, 2008), preferenza verso argomenti trasversali determinata da un approccio di studio multidisciplinare. <i>Approccio conservativo vs approccio volto alla scoperta</i> (Martin, 2016; Hessels e van Lente, 2008): confronto tra gli approcci di ricerca volti ad individuare la preferenza per argomenti più sicuri e con una più solida tradizione rispetto alla preferenza per i campi emergenti, associati a scoperte la cui risonanza scientifica è potenzialmente maggiore al pari dei rischi per risultati inconcludenti

Tab. 1 - segue

Ambiti di analisi	Elementi di analisi
Pratiche di pubblicazione	<p><i>Tipo di pubblicazione</i> (Gunnulfsen., 2021; Karlsson, 2017): tipologie di pubblicazioni maggiormente curate e livello di coerenza rispetto all'area disciplinare di riferimento.</p> <p><i>Tipo di rivista</i> (Altman et al., 2020): tipo di rivista a cui si sottopongono i lavori di ricerca e motivazioni delle scelte (es. affinità della rivista rispetto al tema trattato o visibilità e prestigio della rivista scelta).</p> <p><i>Approccio conservativo vs approccio volto alla scoperta</i>: approccio che mira a inserirsi nelle pratiche di pubblicazioni in un solco già esistente come tipologia e modalità di esposizione dei temi di ricerca <i>vs</i> nuovi modi di presentazione del lavoro di ricerca, potenzialmente più innovativi e ambiziosi.</p>
Collaborazioni accademiche	<p><i>Volontà di collaborare vs invito a collaborare</i> (Groves et al., 2018; Santos e Horta, 2018): sviluppo delle collaborazioni di ricerca sulla base di una precisa volontà di iniziativa nella ricerca di partner contrapposta allo sviluppo di dinamiche in cui la collaborazione non dipende dalla sola volontà del ricercatore (inclusione in progetti di ricerca con alto numero di ricercatori, spinta e/o imposizione di altri soggetti, ecc.). Le due situazioni possono coesistere, in fasi diverse, della carriera di un ricercatore.</p> <p><i>Influenza del mentore</i>: rispetto a questa area di analisi va individuata l'eventuale influenza della figura riconosciuta come mentore nella determinazione delle collaborazioni accademiche.</p>
Mobilità	<p><i>Volontà di collaborare (passivo) vs invito a collaborare (attivo)</i>: rispetto a questa area di analisi vanno indagati gli elementi relativi allo svilupparsi di collaborazioni di ricerca che hanno un effetto diretto sulla mobilità del ricercatore, determinando se la mobilità sia una conseguenza di scelte del ricercatore o di dinamiche esterne in tutto o in parte alla sua volontà.</p> <p><i>Omogeneità vs Eterogeneità</i> (Brown et al., 2013; Minelli et al., 2006) relativa alla gamma e alla tipologia di luoghi per la generazione di conoscenza (nuove reti di collaborazione).</p>
Aspettative di carriera	<p><i>Prestigio</i> (Dahler-Larsen, 2012; Santos e Horta, 2018): attività di ricerca volta al riconoscimento da parte dei propri pari anche finalizzato allo sviluppo futuro della carriera.</p> <p><i>Scoperta</i> (Hessels e van Lente, 2008): preferenza per l'esplorazione di campi emergenti e ancora poco approfonditi, e in cui esista un basso numero di pubblicazione, rispetto ad altri campi maggiormente esplorati.</p> <p><i>Conservativo</i> (<i>Ibidem</i>): preferenza per argomenti di ricerca consolidati, e tendenzialmente <i>mainstream</i> rispetto al proprio campo di ricerca specifico.</p>
Contesto Accademico Effetti della valutazione	<p><i>Apprendimento organizzativo</i> (Minelli et al., 2006): effetti della valutazione in termini di aumento di conoscenza, cambiamento nelle pratiche di lavoro (team e pubblicazioni), accettazione della competizione.</p> <p><i>Sviluppo di risorse</i> (Mathies et al. 2020): effetti della valutazione sulle dinamiche interne alle università in termini di allocazione di risorse, progressione di carriera, trasformazione della struttura e miglioramento e della gestione delle competenze del capitale umano.</p>

Tab. 1 - segue

Ambiti di analisi	Elementi di analisi
	<i>Sistema di potere (Minelli et al., 2006; Tirado et. al 2022): effetti della valutazione sul consolidamento sia degli Enti incaricati della valutazione che degli Enti valutati, così come quelli dovuti al miglioramento delle capacità decisionali degli enti di governo delle Università e dei dipartimenti.</i>

I temi di conversazione sono presentati in modo ampio e generico per permettere all'intervistato di presentare gli argomenti che reputa rilevanti in base alla sua gerarchia di priorità, e per costruire in maniera autonoma, spontanea e personale la gamma di percezioni ed esperienze ritenute significative relativamente ad ogni punto proposto dall'intervistatore (Bichi, 2000, 2002).

Le dimensioni analitiche non sono mai state citate esplicitamente dall'intervistatore, per limitare l'influenza o la direzionalità da parte dell'intervistatore su giudizi, percezioni e priorità individuate dagli intervistati sugli argomenti trattati. L'intervistatore ha avuto cura di annotare gli elementi direttamente afferenti alle dimensioni di indagine emersi spontaneamente dalle risposte fornite.

L'intervistato è stato lasciato molto libero nel racconto, raccogliendo e rilanciando anche rispetto ai suoi temi ed alle sue preferenze di narrazione. In questo caso, se gli effetti della valutazione sulle pratiche e sul vissuto non fossero emersi spontaneamente nel discorso, non sarebbero stati rilanciati dall'intervistatore in alcun modo.

Temi di conversazione

Può raccontare la sua giornata (o settimana) tipo?

Può raccontare il suo percorso accademico dal dottorato all'attuale posizione: vissuto, motivazioni, difficoltà.

Quali fattori guidano la sua attività di ricerca, ad esempio, nella scelta di temi di studio, nella partecipazione/sottomissione di progetti e nelle collaborazioni con altri colleghi?

Riconosce nello sviluppo della sua carriera una figura di mentore? Può raccontare il rapporto negli anni?

Effetti delle esperienze all'estero sulla sua attività di ricerca e sulla crescita professionale.

Come sceglie di pubblicare i risultati della sua attività di ricerca?

Quali sono le aspettative sulla sua carriera: fattori determinanti e criticità.

Attraverso la strutturazione di temi di conversazione e la creazione di un ambiente di dialogo informale, gli intervistati sono stati incoraggiati a esplorare in modo spontaneo le dinamiche della valutazione nella loro carriera.

5. Principali risultati

I risultati presentano una prospettiva ricca e articolata sulle dinamiche del percorso accademico dei giovani ricercatori, evidenziando l'importanza dell'influenza del mentore e i determinanti nelle strategie di pubblicazione, nelle collaborazioni e nelle esperienze all'estero.

La seguente tabella, riprendendo la concettualizzazione presentata nel paragrafo precedente, sintetizza per elementi-chiave i principali risultati emersi a partire dalle citazioni analizzate, a seguire il dettaglio dello studio.

Tab. 2 - Sintesi dei risultati rispetto alle dimensioni analitiche

Ambiti di analisi	Principali evidenze
Percorso accademico	<ul style="list-style-type: none">Contributo specifico del mentore nel plasmare le decisioni chiave e la crescita scientifica degli intervistati. Spinta vs contrapposizione.Influenza del mentore anche su opportunità di ricerca e di finanziamento.Piccoli settori accademici italiani, domanda e offerta possono invertirsi.Vari casi di concorsi deserti per valutazioni severe mostrano le tensioni nel sistema.
Attività/interessi di ricerca	<ul style="list-style-type: none">Influenza della competizione finanziaria nella scelta degli argomenti di studio, favorendo ricerca applicata finanziata da privati.Ricerca di base sacrificata per instabilità economica profili.Passaggio da specializzazione a capacità di spaziare tra temi.Maggiore permeabilità tra ricerca e didattica, influenzata anche dalle politiche universitarie.Richiesta crescente di competenze interdisciplinari, allineamento con mercato.
Pratiche di pubblicazione	<ul style="list-style-type: none">Pubblicare è essenziale per la carriera e l'identità dei ricercatori.Le riviste indicizzate garantiscono qualità in alcuni settori.Nelle SSH i criteri di valutazione non sempre si allineano con il processo di crescita dei ricercatori.Riflessioni condivise sulla crescente pressione a pubblicare in riviste internazionali o di fascia A/ANVUR per garantire visibilità e avanzamento di carriera.

Tab. 1 - segue

Ambiti di analisi	Principali evidenze
<i>Collaborazioni accademiche</i>	<ul style="list-style-type: none"> • Dualità degli approcci: autonomia <i>vs</i> risposta a inviti esterni. • Necessità di creare canali propri per attrarre partner di ricerca diversificati. • Importanza dell'internazionalizzazione delle collaborazioni accademiche. • Ruolo delle risorse finanziarie nei partenariati accademici. • Complessità delle dinamiche: volontà personale, <i>mentorship</i>, competizione, internazionalizzazione e opportunità finanziarie.
<i>Mobilità</i>	<ul style="list-style-type: none"> • Importanza della mobilità internazionale come stimolo competitivo e trampolino di lancio. • Vantaggi e sfide psicologiche dell'esperienza all'estero. • Ruolo determinante del mentore nella facilitazione delle opportunità internazionali e nello sviluppo professionale all'estero.
<i>Valutazione</i>	<ul style="list-style-type: none"> • Importanza centrale dell'ASN nella produzione e nelle strategie di pubblicazione.
<i>Contesto accademico</i>	<ul style="list-style-type: none"> • Valutazione esterna alla propria comunità epistemica vista come condizionante. • Mancanza di una premialità chiara. • Critiche ai criteri bibliometrici nella valutazione e necessità di criteri più ampi che includano didattica e valutazione degli studenti. • Rigidità del sistema accademico italiano rispetto ad altri paesi. • Importanza crescente della valutazione della terza missione accademica. • Bisogno di una valutazione equilibrata che riconosca crescita individuale, multidisciplinarietà e esperienze internazionali.

Percorso accademico

Diversi e a volte indiretti sono gli effetti della valutazione sul percorso di carriera accademica. L'evidenza più forte restituita dalle SDV sul percorso di carriera dei giovani ricercatori è l'importanza del mentore.

L'analisi, infatti, ha identificato il contributo specifico del mentore nel plasmare le decisioni chiave e la crescita scientifica degli individui, esplorando le dinamiche del rapporto mentore-descendente e valutando le sfide e i benefici derivanti da tale guida. Le testimonianze di alcuni enfatizzano l'importanza di una guida severa, ma preziosa, nel delineare il percorso di crescita scientifica individuale:

Era il direttore del dottorato, mi ha fatto sudare fino alla fine, ma non con dispetti, nel senso di richiesta, di rigore, di correttezza. Mi ha continuamente pungolato, il che mi ha aiutato a salire di livello (RTD-A, UOMO, CENTRO, Area 7).

È la persona che mi ha instradato, perché io non capivo nulla, nel senso che sono stato un bravo studente, credo, però non sapevo cosa avrei dovuto fare, mi ha indirizzato su cosa studiare, mi ha messo alla prova, mi ha fatto scrivere, mi ha corretto, anche in modo molto duro in certi momenti, solo dopo qualche anno ho capito che fosse nel mio interesse, ma il rapporto è stato molto severo nelle sue fasi iniziali (RTD-B, UOMO, SUD E ISOLE; Area 12).

Oppure come il mentore abbia influenzato non solo la scelta del percorso accademico ma anche le opportunità di ricerca e finanziamento:

[...] il mio mentore – chiamiamolo così – il professore della tesi seguiva sempre il mio percorso, il professore [...], se ha visto qualche cosa del mio curriculum troverà il suo nome, era entrato in un PON, con [...], un progetto importante, piuttosto lungo e anche abbastanza ben finanziato, per cui mi hanno proposto il post dottorato, con un assegno di ricerca e questo è stato un lungo assegno di ricerca (RTD-B, DONNA, SUD E ISOLE, Area 8).

Io l'ho conosciuto nel 2001, quando lui era appena arrivato presso il nostro dipartimento con le lezioni di Psicologia; gli chiesi la tesi, non avevo neppure la consapevolezza dell'importanza di quella scelta, perché poi è da lì che è iniziata la mia avventura accademica con lui. Mi ha sempre spinto a fare oltre, non mi sentivo pronta, ma lui mi diede parte del suo insegnamento di Psicologia clinica, mi ero appena laureata; quindi, ancora non avevo completato neppure la specializzazione e fece in modo di avere un bando di docenza per Psicologia clinica; quindi, appena laureata sono andata ad insegnare ai miei colleghi; lui ha sempre fatto in modo che noi potessimo cavalcare un po' l'onda, anche di tematiche, oggetto di studio importanti. È stata, sicuramente, una persona importante, non solo dal punto di vista professionale, ma anche personale (RTD-B, DONNA, SUD E ISOLE, Area 11).

Alcuni intervistati, invece, sottolineano le sfide nei settori di dimensioni ridotte, dove la domanda e l'offerta possono invertirsi in maniera inusuale. La descrizione di concorsi deserti a causa di valutazioni severe offre uno sguardo penetrante sulle difficoltà e le tensioni nel sistema accademico italiano:

Io vengo da un settore molto piccolo, era tutto molto diverso, noi ci siamo trovati in una situazione dove la domanda e l'offerta era rovesciata, nel senso che ci sono stati un paio d'anni dove forse eravamo quattro abilitati come professore associato in tutta Italia. Ci sono 96 Atenei, di questi almeno 80-85 avranno Giurisprudenza, quindi è chiaro che, semmai, si è creato il mercato al contrario – “vieni da me”, “No, vieni da me”, “Vai dall'altra” – ma è una cosa rarissima, normalmente è l'esatto contrario, quindi i nostri concorsi, per assurdo, andavano deserti, perché le Commissioni sono state talmente strette o rigide nelle valutazioni che non c'erano docenti disponibili (RTD-B, UOMO, SUD E ISOLE; Area 12).

Interessi e attività di ricerca

La competizione collegata all'acquisizione di finanziamenti per la ricerca influenza in molti casi la scelta degli argomenti di studio. Se da una parte si sottolinea come la ricerca applicata finanziata dalle aziende fornisca sicurezza finanziaria, dall'altro si evidenzia che la ricerca di base in alcuni casi viene sacrificata per motivi economici legati all'instabilità dei profili:

I miei finanziamenti, ultimamente, visto che lavoro sul ..., di interesse agronomico – io lavoro al Dipartimento ..., un dipartimento molto applicativo – lavoro molto con le aziende e riesco a fare della ricerca applicata, che interessa le aziende, per cui loro mi finanziano e io, con quei soldi, a fare ricerca per loro e ricerca di base, ovviamente la ricerca di base rimane più sacrificata, però rispetto ai Grant classici, lavorare con le aziende mi dà una sicurezza diversa, perché se porto a casa il risultato sono sicura che avrò il contratto rinnovato con l'azienda (RTD-A, DONNA, NORD EST, Area 5).

Emerge anche un cambiamento nel modo in cui i ricercatori affrontano i loro temi di ricerca. La tendenza a specializzarsi su argomenti specifici per l'intera carriera viene sostituita in alcune aree da una valutazione positiva per coloro che sono capaci di spaziare tra temi diversi. Ciò conferma l'apertura crescente verso nuovi ambiti di studio per tenere il passo con nuove richieste che possano emergere anche dalle esigenze degli studenti. Il collegamento tra ricerca e didattica mostra la permeabilità fra queste attività nell'università italiana, e la possibilità di effetti ulteriori che strumenti di *policy* destinati a regolare l'attività di ricerca possono produrre in ambito accademico.

[...] prima nella mia materia c'è stata una tendenza verso lo specializzarsi su alcuni temi e farli propri per tutta la carriera – quello che ha fatto il mio maestro – si è assistito negli anni a un cambiamento, cioè ad una valutazione maggiormente positiva per chi è capace, dovendo fare docenti di tutta la materia, di spaziare tra temi diversi, adesso che sto scrivendo la terza monografia ho cercato dei macro temi, così come nei temi dei saggi, di occuparmi un po' di tutto, proprio perché effettivamente, siccome l'attività di ricerca diventa fondamentale anche per la didattica, soprattutto nei corsi di alta formazione, più che per gli studenti, dove le lezioni sono, chiaramente, un pochino più istituzionali – per capirci – questo costituisce anche l'occasione per provare ad essere, comunque, un professore preparato un po' su tutto, perché la ricerca e la didattica non sono scollegati, quindi i temi sono stati scelti, perché mi piacevano, fondamentalmente per passione, anche questo è frutto di un insegnamento (RTD-B, UOMO, SUD E ISOLE; Area 12).

Alcune riflessioni sulle carriere indicano una situazione di inerzia e immobilità che caratterizzano talvolta i percorsi accademici:

Il mio percorso all'Università è stato, veramente, faticoso, perché non avevo nessun tipo di garanzia, quindi ero lì, in qualche modo, come volontaria, anche lo stesso dottorato, vinto senza borsa. [...] sono rimasta sempre nello stesso Ateneo, nello stesso dipartimento, con lo stesso tutor, che è stato il mio punto di riferimento (RTD-B, DONNA, SUD E ISOLE, Area 11).

La crescente richiesta del mercato di individui competenti su più ambiti viene percepita dai giovani accademici come una trasformazione importante che indica come l'interdisciplinarietà sia diventata indispensabile. La ricerca di un collegamento con le esigenze del mercato spinge in alcuni settori verso un cambiamento nella selezione delle domande e degli ambiti di ricerca e un'apertura verso nuove frontiere:

[...] sicuramente la prospettiva è quella di trasformare queste ricerche, così come altre che ne verranno, in ricerche di carattere interdisciplinare, ma credo che, comunque, ci sia proprio una tendenza, anche come comunità di ingegneria in ambito accademico, a trasformare e a rendere un po' più interdisciplinare il nostro ambito, perché non se ne può fare a meno, credo che sia ormai indispensabile, ma non lo credo solo io, ma anche a livello più alto e ha l'opportunità di confrontarsi su questi argomenti. Credo che l'interdisciplinarietà sia ormai proprio richiesta dal mercato, di avere delle persone che siano non multisettoriali, ma che siano competenti su più ambiti (RTD-B, DONNA, SUD E ISOLE, Area 8).

Pratiche di pubblicazione

Il discorso sulle pratiche di pubblicazione è quello che maggiormente fa emergere gli effetti della valutazione. Pubblicare è indispensabile non solo per la carriera ma per “essere” ricercatori e ricercatrici e per costruire la propria identità. Tuttavia, le percezioni sono diverse e fortemente legate al settore scientifico di riferimento. Le riviste indicizzate sono considerate un ottimo modo di pubblicare per i giovani ricercatori di alcuni settori, in quanto garantiscono una revisione di qualità. Non è così per altri settori, in genere legati alle scienze sociali e umane, dove i criteri imposti dalla valutazione non sempre sono in linea con il naturale processo di crescita anche identitaria del ricercatore. La valutazione, quindi, viene percepita come fortemente condizionante specie se non proviene dalla propria comunità epistemica di riferimento:

I miei articoli, quelli che ho sempre pubblicato, sia io o in collaborazione, sono il lavoro di tanti anni, decine, centinaia e migliaia di ore di lavoro buttate dentro,

però con delle scoperte nuove, ad esempio nel [...] ho pubblicato un bell'articolo come primo autore, molto bello, sono molto fiera di quell'articolo, ci ho messo tanto a farlo, ci ho messo sei anni, lavorando tantissimo, però sono molto fiera di quel lavoro, è un lavoro che è citato molto, è interessante, è un lavoro che si sta confermando corretto dalla comunità scientifica (RTD-A, DONNA, NORD EST, Area 5).

A parte che è indispensabile, perché siamo valutati su quello e lo devi fare per forza, ma io lo vedo un modo per restituire alla gente quello che mi è stato dato, perché lavoriamo con i soldi pubblici, io non faccio Service, non ho ditte private che mi sostengono, facendo ricerca di base non interessa alle ditte; quindi, restituisco alla gente quello che mi viene dato, almeno cerco di restituire qualcosa in termini di conoscenza, non vorrei essere troppo supponente (RTD-B, UOMO, NORD EST, Area 5).

La riflessione sulla pressione crescente verso la pubblicazione in riviste internazionali o classificate da ANVUR come di fascia A per garantire visibilità e avanzamento di carriera, è condivisa da molti degli intervistati:

Quando io ho iniziato a pubblicare non c'era tutto il discorso delle mediane, degli indicatori, mentre adesso le riviste si trovano a dover gestire una richiesta maggiore. Io ho avuto la sensazione che si cerchi, in qualche modo, di privilegiare anche ricercatori che hanno una certa visibilità; quindi, sono riconoscibili rispetto a quella tematica (RTD-B, DONNA, SUD E ISOLE, Area 11).

Ho sempre fatto un po' da autodidatta nelle riviste italiane, non indicizzate su Scopus, poi mi sono lanciato su quelle e pubblico solo su quelle, perché su quelle c'è la valutazione (RTD-A, UOMO, CENTRO, Area 7).

Io di base mi sono orientato nel cercare di pubblicare su riviste di fascia A, perché mi serviva per l'abilitazione. [...] (RTD-B, Uomo, CENTRO, Area 12).

Al momento, gli output della mia ricerca sono essenzialmente pubblicazioni su riviste prevalentemente internazionali, perché danno più riscontro in termini di carriera, ma anche di visibilità, soprattutto di visibilità (RTD-B, DONNA, SUD E ISOLE, Area 8).

In generale la consapevolezza delle nuove sfide legate alla progettualità, alle regole della produzione scientifica, e alla volontà di diversificare le attività di ricerca, indicano una riflessione critica sull'adattamento alle richieste mutevoli del contesto accademico.

Collaborazioni accademiche

Le collaborazioni accademiche possono scaturire sia dalla volontà intrinseca del ricercatore di presentare progetti di ricerca che dall'invito esterno, come la partecipazione a progetti competitivi, primi fra tutti i

PRIN. Questo suggerisce una dualità di approcci, con alcuni ricercatori che sono in grado di iniziare collaborazioni in modo autonomo, mentre altri rispondono a inviti o opportunità esterne:

È chiaro che nella partecipazione a progetti PRIN o altre cose del genere, soprattutto quando ti unisci ad altri gruppi interdisciplinari, mi è capitato anche adesso, i nostri internazionalisti avevano necessità di un supporto ad un processo penalista e mi hanno chiesto, sul tema che avevano individuato, se mi potevano inserire nell'unità, quindi a quel punto dai la tua linea, che però è orientata da altri fattori (RTD-B, UOMO, SUD E ISOLE; Area 12).

In alcuni casi siamo stati contattati. In tempi recenti, le proposte di cui le parlavo dell'ultimo anno, a parte i PRIN, dove invece ci siamo mossi in maniera più diretta, siamo stati un po' coinvolti; anche se abbiamo diversi contatti, però i nostri contatti non arrivano a coprire quest'ampia sfera di partner (RTD-B, DONNA, SUD E ISOLE, Area 8).

I ricercatori sottolineano anche la necessità di creare canali propri per essere coinvolti, indicando che, specialmente nelle università più piccole o in contesti geografici specifici, può essere complicato attirare l'attenzione di partner di ricerca diversificati. Tuttavia, non emergono indicazioni che individuino la valutazione come un elemento che influenza questo tipo di scelte o queste opportunità:

[...] bisogna crearsi dei canali propri per riuscire a essere coinvolti, perché l'Università piccola del sud è difficile che riesca, non impossibile, si può sempre fare, però è complicato riuscire a coinvolgere tanti soggetti diversi di paesi diversi su un unico tema (RTD-B, DONNA, SUD E ISOLE, Area 11).

Alcune considerazioni riguardano l'importanza dell'internazionalizzazione delle collaborazioni accademiche. Essere parte di associazioni internazionali o partecipare a convegni internazionali viene considerato un elemento chiave nel favorire la diversificazione delle reti accademiche e nell'aprire opportunità di collaborazione:

Grazie al network diventa un pezzo scritto da ognuno e riesci poi a gestire una call in tempi "umani", poi considera che in ambito europeo sono, comunque, molto più abituati di noi a partecipare a questo tipo di call, per cui sanno già come si scrivono i progetti, cosa valorizzare, quali aspetti, sanno fare più lobbying di noi a livello di Commissione Europea – che poi è quello che serve, un po' tutto serve – soprattutto il Nord Europa, anche la Francia, hanno più capacità di fare rete istituzionale, oltre che nell'ambito della ricerca (RTD-B, UOMO, NORD EST, Area 5).

[...] come attività internazionale, noi abbiamo dei progetti di carattere internazionale e, chiaramente, pubblichiamo. Questa è una cosa che mi porto dietro da pri-

ma, con il gruppo di XXX; in quel periodo in cui avevamo tante attività, che non erano solo nazionali, avevamo progetti con diversi paesi dell'area mediterranea, la Spagna, il Marocco, oltre ad altri partner italiani, poi Algeria, Israele, in alcuni casi con i Giordania [...] All'epoca era più facile, perché avevamo dei meeting prestabiliti; quindi, ci si incontrava più volte all'anno (RTD-B, DONNA, SUD E ISOLE, Area 11).

Alcuni intervistati infine sottolineano che alcune collaborazioni sono guidate da requisiti specifici di bandi di finanziamento, indicando che la disponibilità di risorse finanziarie può influenzare significativamente la natura e la portata delle collaborazioni:

Altri bandi funzionano che sono più dedicati a noi giovani, quindi sono “esclusiva” mia, nel senso che io propongo una tematica e chiedo consiglio “ho in mente di portare questa idea, che per bando potrebbe andare bene”, quindi c’è una fase in cui si punta su un aspetto piuttosto che un altro, perché tecnicamente bisogna adattarsi ai requisiti di quel bando (RTD-A, Uomo, NORD OVEST, Area 3).

In conclusione, le testimonianze evidenziano la complessità delle dinamiche delle collaborazioni accademiche, sottolineando la necessità di considerare una serie di fattori, tra cui volontà personale, *mentorship*, competizione, internazionalizzazione e opportunità finanziarie, per comprendere appieno come i ricercatori strutturano e gestiscono le loro reti accademiche.

Mobilità ed esperienze all'estero

Gli intervistati evidenziano l'importanza della mobilità e l'esperienza internazionale è descritta da molti come una palestra stimolante e competitiva, spesso funzionando come trampolino di lancio:

Al di là del fatto di restare sempre nello stesso posto – pro e contro – ti dico come la penso: io ho avuto tantissimo dalla mia esperienza all'estero, ho vissuto, in tutto, circa un anno e mezzo a [Città statunitense], mi ha dato tantissimo quella esperienza, mi ha cambiato la carriera e anche un po' me, come essere umano, perché è stata un'esperienza molto bella e molto formativa, quindi rimanere sempre allo stesso posto ha lo svantaggio che non ti guardi intorno, non vedi altre realtà, non allarghi la tua cerchia, ecc. (RTD-B, UOMO, NORD EST, Area 5).

Mentre alcuni sottolineano i vantaggi derivanti da periodi di studio e lavoro all'estero, altri evidenziano le sfide psicologiche e il clima accademico impegnativo, che, sebbene faticoso, ha contribuito a sviluppare una resilienza notevole:

[...] il mio professore mi ha coperto con una borsa per il periodo tra la fine del dottorato e l'inizio del post-doc e sono andato in America. [...] Devo dire "beata incoscienza", perché sono passato da un gruppo di ricerca in cui tutti collaborano, nel senso che ci si confronta con le idee, ci si aiuta in qualsiasi cosa, ad un gruppo altamente competitivo, quale può essere un gruppo americano, un professore bravissimo, nel senso che dal punto di vista della scienza era il top, io ho imparato tantissimo in quell'anno di dottorato, umanamente un pochino più difficile da trattare (RTD-A, Uomo, NORD OVEST, Area 3).

In Germania sono stata quattro anni, dal 2006 al 2011, quattro anni e mezzo, prima [Città tedesca] e poi [Città tedesca]. Al [Ente di ricerca] è stata una palestra pazzesca, ci sono scienziati di tutto il mondo, un ambiente molto competitivo, dove si lavora e si lavora, è un trampolino di lancio (RTD-A, DONNA, NORD EST, Area 5).

Il ruolo del mentore emerge anche rispetto alla mobilità come determinante nella spinta e nelle possibilità di fare esperienze all'estero. Coloro che hanno avuto *mentorship* positiva vedono nell'esperienza internazionale un contributo significativo alla loro crescita professionale:

La fortuna è stata che il mio tutor mi ha sempre lanciata a fare cose, così, probabilmente non ne avevo nemmeno le capacità, però lui vedeva oltre, quindi mi lanciava. Ho partecipato al mio primo convegno che ero ancora studentessa, non mi ero ancora laureata, quando poi mi sono laureata ho iniziato a partecipare a convegni nazionali e internazionali, per cui sono stata in America, a [Città statunitense], sono stata in Austria più volte – in Austria tanti anni – poi in Italia, ho partecipato sempre, in modo continuativo [...] (RTD-B, DONNA, SUD E ISOLE, Area 11).

Contesto accademico ed effetti esplicativi della valutazione

La riflessione esplicita degli intervistati sugli effetti della valutazione sottolinea alcuni aspetti ulteriori. In alcuni casi la mancanza di premialità chiara e tangibile in relazione ad essa:

Se la valutazione comportasse poi anche una premialità, avrebbe più senso, perché quello che ci si chiede è di ottemperare a tutta una serie di parametri, sia dal punto di vista della ricerca, sia dal punto di vista della didattica, del servizio agli studenti, ecc., però non è che dice "Hai preso l'abilitazione, sei bravo, ti premio, hai fatto bene, hai pubblicato due articoli di fascia A nel corso di tre anni, quindi ti assumo", in realtà è una valutazione di cui beneficiano sostanzialmente gli Atenei e forse i dipartimenti, non sicuramente il singolo, per cui uno lo fa perché lo deve fare, non è che trovo nella valutazione il senso della mia ricerca (RTD-A, Uomo, CENTRO, Area 10).

Il più grande difetto di questo sistema di valutazione dovrebbe più incidere sugli stipendi delle persone, questo potrebbe essere proprio una rivoluzione copernicana (RTD-A, Uomo, CENTRO; Area 7).

L'importanza dell'ASN in particolare emerge come elemento centrale nella produzione accademica e sulle strategie di pubblicazione, più che nelle decisioni relative ai temi di ricerca:

[...] sotto il profilo dell'ASN, anche leggendo i criteri, che vengono indicati di solito dai Commissari, ti viene detto, sostanzialmente, “senti, noi leggiamo le monografie, poi il resto è il contorno, ma non è il piatto principale”, quindi con tutto che le riviste in fascia A prevedono il *check*, il *double-blind*, tutto quello che si vuole, il lavoro monografico è, comunque, considerato apicale rispetto a tutto il resto (RTD-B, Uomo, SUD E ISOLE, Area 12).

Le percezioni sulla valutazione comunque variano, alcune voci sottolineano la mancanza di un legame diretto tra la valutazione e il riconoscimento individuale, e altre mettono in evidenza il suo impatto sulla produzione e orientamento della ricerca:

Per fortuna è arrivata in un momento in cui iniziavo a produrre, per cui ho orientato la mia produzione anche sulla base delle soglie previste o comunque dei titoli che mi servivano per l'ASN; poi che sia giusto o non sia giusto non lo so, però di fatto è pieno di gente abilitata che sta in giro (RTD-A, Uomo, CENTRO, Area 10).

Alcuni ritengono che i criteri di valutazione basati su parametri bibliometrici possano essere considerati limitanti e deprimenti. L'inclusione della didattica e della valutazione degli studenti come aspetti importanti nella valutazione accademica suggerisce la necessità di una visione più ampia e bilanciata:

Per la valutazione, io credo che, almeno a livello universitario, io non guarderei solo i parametri bibliometrici, il numero degli articoli, perché ci deve essere un minimo; [...] ma anche la valutazione degli studenti sulla didattica è importante, cioè se uno diventa universitario deve anche insegnare, gli studenti valutano la didattica e quello lì è uno strumento importante, perché la didattica è veramente importante per un universitario (RTD-A, Donna, NORD EST, Area 5).

Chi è in contatto e collabora con accademici all'estero evidenzia la rigidità del sistema accademico italiano rispetto ad altri paesi, sottolineando la necessità di adottare criteri di valutazione più flessibili. La valutazione basata su numeri potrebbe non riflettere adeguatamente la complessità delle attività accademiche:

Io trovo che adesso, in particolar modo in Italia – lo vedo qui e ho due o tre riferimenti: l'Italia, Stati Uniti, Canada e forse anche l'Australia, perché conosco colleghi che lavorano là – il nostro sistema, rispetto a loro, è più rigido, nel senso che noi

siamo molto più abituati a mettere paletti, è sempre molto tutto complicato in Italia, se non si mettono mille regole, *escamotage* per trovare tutte le direzioni trasversali, diciamo che partiamo creativi, siamo molto rigidi sulle valutazioni e sul reclutamento, vogliamo basarci su dei criteri super oggettivi, che per natura stessa del temine sono numeri, allora io vedo che noi siamo pesantemente legati a dei numeri, mentre all'estero fanno altri tipi di valutazione (RTD-B, Uomo, NORD EST, Area 5).

Infine, la valutazione della terza missione emerge come un elemento di importanza crescente, incoraggiando i ricercatori a dimostrare il loro coinvolgimento sociale attraverso la collaborazione con le scuole e altre istituzioni:

Siamo sempre divisi tra didattica, ricerca e terza missione. Didattica e ricerca sono cose che ormai vanno da sole, la terza missione è quello che non sei obbligato a fare, ma se la fai è meglio; adesso pare che nei parametri di ANVUR vogliano valutare anche la terza missione e tutto, per cui in qualche modo la devi certificare, il fatto di avere questi tipi di lavori, di collaborazione con le scuole, è un modo per “dimostrare” che ti occupi anche del sociale (RTD-A, Uomo, NORD OVEST, Area 3).

Questi elementi in generale indicano la necessità di una valutazione equilibrata, che riconosca il valore della crescita individuale, della multidisciplinarietà e delle esperienze internazionali. Inoltre, suggeriscono che una valutazione più ampia. Inoltre, suggeriscono che una valutazione più ampia, che comprendesse ad esempio anche l'impatto sociale e culturale legato alle attività della Terza e Quarta Missione delle università, potrebbe favorire un ambiente accademico più inclusivo, stimolante e adattabile alle mutevoli esigenze non solo della ricerca scientifica ma anche della società.

6. Discussione e conclusioni

Il Capitolo ha esaminato i fattori che influenzano le carriere accademiche dei giovani ricercatori, con particolare attenzione alla valutazione delle attività di ricerca. La metodologia delle SDV informate ha consentito di esplorare esperienze personali di ricercatori e comprendere gli effetti sulle loro scelte di carriera.

Le SDV hanno inoltre catturato la diversità delle esperienze accademiche, contribuendo a delineare il quadro complesso in cui la valutazione sulla ricerca si inserisce e quale tipo di incentivi fornisce per la strutturazione della carriera accademica.

Gli intervistati riconoscono che le politiche di valutazione hanno effetti sulla produzione scientifica, e possono influire anche sulle relazioni

interpersonali, sulle strategie di pubblicazione e sulla percezione del prestigio accademico. Ma l'atteggiamento nei confronti della valutazione non è uniforme e spesso è condizionato dal settore di appartenenza e dalle pratiche presenti nelle diverse comunità epistemiche. Per i giovani ricercatori si conferma in modo più netto quanto già emerso nel Capitolo 1 di questo volume, in quanto la nozione di qualità dei diversi settori determina una diversa propensione nei confronti della valutazione esterna: quanto più essa è vicina ai parametri utilizzati dalla comunità scientifica, tanto più sarà considerata in modo positivo dai giovani ricercatori. Quanto più la valutazione impone criteri che non appartengono alla comunità di riferimento, tanto più essa sarà considerata un elemento intrusivo e lesivo della libertà accademica.

I Ricercatori a Tempo Determinato di tipo A e B, nati contestualmente agli strumenti di valutazione come la Valutazione della Qualità della Ricerca (VQR) e l'Abilitazione Scientifica Nazionale (ASN), non considerano questi strumenti come novità ma come integrati nella loro carriera. La valutazione non influenza la scelta o la riorganizzazione del proprio spazio di ricerca, ma entra chiaramente in gioco nella scelta della strategia di pubblicazione. In particolare, gli indicatori della ASN guidano le decisioni di pubblicazione per rafforzare la loro posizione nel contesto accademico. La pubblicazione è essenziale non solo per la carriera ma anche per costruire l'identità di ricercatore. Le percezioni variano tra settori, con alcuni che vedono le riviste indicizzate come l'unico modo per garantire qualità, mentre altri, specialmente nelle scienze sociali e umane, percepiscono questa scelta come limitante. La pressione a pubblicare in riviste internazionali indicizzate o di fascia A è diffusa, ed è considerata rilevante per guadagnare visibilità e per l'avanzamento di carriera.

Anche il nostro studio, in linea con quanto rilevato dalla letteratura (Van Balen *et al.*, 2012), conferma l'importanza del mentore nella costruzione della carriera accademica. Il mentore influenza le prime fasi della carriera, stimolando l'acquisizione di competenze, orientando verso attività internazionali, progetti di ricerca e aiutando nella creazione di una rete di contatti. Alcuni intervistati sottolineano anche le opportunità di ricerca e finanziamento ottenute grazie ai loro mentori. Le testimonianze enfatizzano l'importanza di una guida rigorosa, come descritto da un ricercatore che ha apprezzato l'approccio severo ma formativo del proprio mentore.

Tuttavia, la necessità di massimizzare gli indicatori di valutazione, relativi alla ASN, può portare a tensioni tra la scelta di pubblicare tenendo in considerazione le indicazioni della valutazione e l'accettazione della comunità accademica rappresentata dal mentore e dai relativi orientamenti in termini di qualità scientifica. Questa tensione, evidenziata anche da

contributi in letteratura (Sobey *et al.*, 2013; Mäkinen *et al.*, 2024), crea una contrapposizione che i giovani ricercatori devono gestire con attenzione.

La competizione per i finanziamenti, d'altra parte, è il fattore che fortemente influenza la scelta degli argomenti di studio. La ricerca applicata offre continuità e certezza di finanziamento, mentre la ricerca di base a volte deve essere sacrificata non avendo strumenti di finanziamento dedicati. Inoltre, i giovani ricercatori sono più orientati verso approcci interdisciplinari, in accordo con i più recenti orientamenti degli strumenti di finanziamento della ricerca e di didattica su temi collegati alla sostenibilità.

La mobilità internazionale è vista come fondamentale per la crescita professionale. Le esperienze all'estero offrono opportunità formative significative, anche se possono presentare sfide psicologiche. Il mentore gioca un ruolo importante nel facilitare queste esperienze, aiutando i giovani ricercatori a superare le difficoltà iniziali e a sfruttare al meglio le opportunità internazionali.

Le collaborazioni accademiche possono nascere sia dall'iniziativa personale sia da inviti esterni. La partecipazione a progetti competitivi, come i PRIN, e la necessità di creare proprie reti di contatti sono fondamentali. L'internazionalizzazione delle collaborazioni è vista come un elemento chiave per diversificare le reti accademiche e aprire nuove opportunità. Tuttavia, le risorse finanziarie possono influenzare significativamente la natura e la portata delle collaborazioni.

In conclusione, le dinamiche accademiche dei giovani ricercatori sono influenzate da una varietà di fattori sia esterni che interni, tra cui la valutazione (pressione esterna), l'influenza del mentore (pressione interna), la competizione per i finanziamenti (pressione interna/esterna), e l'importanza delle collaborazioni e della mobilità internazionale. Questi elementi delineano un quadro complesso e articolato del percorso accademico, richiedendo un bilanciamento attento tra diverse esigenze e opportunità.

3. La formazione dell'agenda di ricerca e gli effetti della valutazione esterna

di *Valentina Carazzolo*

1. Introduzione

In molti paesi OCSE, negli ultimi decenni, l'università ha sperimentato un intenso periodo di trasformazione in risposta al manifestarsi di fenomeni sia interni che esterni al sistema universitario stesso. Sin dagli anni '60, eventi come la liberalizzazione dell'accesso universitario e la riduzione delle risorse finanziarie destinate all'istruzione superiore dovute alla crisi economica del 1973, hanno indotto i legislatori ad interrogarsi su quale dovesse essere il ruolo dell'università in un contesto socioeconomico così complesso e mutevole (Paradeise *et al.*, 2009). Le risposte dei governi sono state numerose e molto diverse tra loro: tra tutte, è opportuno citare, la ridotta disponibilità di fondi per la ricerca e il passaggio ad un'allocazione delle risorse su base premiale, l'introduzione del concetto di terza missione e del ruolo attivo dell'università nel contesto economico e sociale in cui è inserita, e, per ultimo, il concetto della valutazione della ricerca.

La ricerca, qui presentata, si concentra sull'ultima delle (relativamente) nuove misure introdotte nel contesto accademico, ovvero sull'introduzione del concetto di valutazione della ricerca, con uno specifico *focus* sul contesto accademico italiano. Vale la pena, a nostro avviso, riflettere sulla valutazione della ricerca nel contesto accademico italiano per due ragioni preminent. In primo luogo, il contesto universitario italiano si configura come peculiare rispetto ai sistemi universitari delle altre maggiori economie europee. La valutazione della ricerca è stata ufficialmente introdotta relativamente tardi rispetto ad altri paesi, a seguito di un percorso lungo e complesso (con l'introduzione della Valutazione della Qualità della Ricerca, d'ora in poi VQR, nel 2010). In secondo luogo, nello stesso periodo è stato adottato uno strumento nuovo di valutazione individuale dei ricercatori per il conseguimento dell'abilitazione scientifica nazionale (ASN),

facendo sì che l'accademia italiana sperimentasse due importanti elementi di novità in poco tempo.

L'obiettivo della nostra ricerca è capire se l'introduzione della valutazione della ricerca (a livello di ateneo e a livello individuale), abbia influito non solo sulle modalità con cui i ricercatori conducono le loro attività, ma anche sulle pratiche, sulle finalità e sui temi centrali del loro *research agenda setting*, generando così un cambiamento nella “curva epistemologica” di una specifica parte della comunità accademica. Come si vedrà più nel dettaglio (nei paragrafi successivi), per ‘agenda di ricerca’ dei ricercatori si intende l’attività con la quale i ricercatori selezionano i temi su cui svolgere le proprie ricerche, e rappresenta pertanto un elemento di assoluta rilevanza per la reputazione e la carriera del ricercatore (Latour e Woolgar, 1979).

A tale scopo, ci siamo direttamente rivolti ai ricercatori attualmente attivi nel contesto universitario italiano, adottando un approccio metodologico misto, basato sulla convergenza parallela di metodi quantitativo e qualitativo (Creswell *et al.*, 2003): in un primo momento con una *survey*, condotta a livello nazionale e, in un secondo momento, con lo svolgimento di interviste informate agli esperti di valutazione nello specifico contesto italiano (Laudel e Gläser, 2007).

2. Valutazione della ricerca e professione accademica nel contesto italiano

Le proposte di ridefinizione del ruolo della ricerca nel contesto socioeconomico di riferimento (Shattock, 2014) hanno contribuito ad una riorganizzazione dell'università sia sotto il profilo amministrativo-organizzativo, sia per quanto riguarda il ruolo di ricercatori e professori. A seguito della progressiva ridefinizione del ruolo delle università, in passato spesso caratterizzato da un'autonomia e da una separazione delle istituzioni accademiche dal contesto sociale, particolarmente evidente in specifici periodi e contesti storico-istituzionali¹ (Merton, 1938; Binswanger, 2014), la maggior parte dei paesi occidentali ha introdotto una serie di riforme per la gestione del sistema universitario ispirate ai principi della dottrina neoliberale, secondo una strategia di controllo e direzione delle attività da remoto (*steering-at-a-distance*). Secondo il modello di *governance* denominato *New Public Management* (NPM), diffusosi a partire dagli anni '80 in numerosi paesi OCSE, una gestione ottimale delle risorse pubbliche sarebbe stata

1. In particolar modo nei paesi le cui tradizioni accademiche afferiscono al sistema accademico di tradizione napoleonica o humboldtiana (Bongaerts, 2022).

possibile grazie al decentramento delle responsabilità alle autorità locali che, una volta ottenuta l'autonomia finanziario-amministrativa, avrebbero dovuto ottemperare a parametri di produttività stabiliti a priori a livello centrale. Queste, una volta ottenuta tale autonomia, avrebbero dovuto rispettare parametri di produttività definiti centralmente. Tale progressivo passaggio dal metodo tradizionale di gestione della pubblica amministrazione verso l'approccio proposto dal NPM, è stato guidato da un crescente senso di insoddisfazione nei confronti delle modalità operative del settore pubblico (Hood, 1995). Il NPM, al contrario, sposta l'attenzione dal processo al risultato, ispirandosi fortemente al modello manageriale delle aziende private che operano in un contesto di libero mercato. Ne deriva la necessità di introdurre meccanismi di controllo (*accountability*) volti a monitorare e valutare l'operato del personale pubblico. Questi includono l'elaborazione di indicatori di performance, lo sviluppo di sistemi di *audit* e controllo esterni, oltre che il rafforzamento di una gerarchia organizzativa modellata sulle aziende private (Paradeise, 2009). Anche nel contesto accademico viene così introdotto l'approccio manageriale istituzionale (Enders *et al.*, 2011), per il quale lo stato valutatore (Neave, 1988) stabilisce priorità e risultati verso cui le università devono tendere, secondo parametri di efficienza e di efficienza nell'uso delle risorse disponibili.

Come detto precedentemente, l'implementazione della valutazione della ricerca nel contesto accademico italiano è avvenuta relativamente tardi rispetto alle maggiori economie europee, con un processo tutt'altro che semplice e lineare. Due sembrano le ragioni prevalenti di questo ritardo. In primo luogo, per una concezione prevalentemente elitista dell'accesso al sistema universitario (per la quale, la preparazione accademica veniva ritenuta appannaggio delle élite in preparazione dell'accesso all'amministrazione pubblica), non era avvertita l'esigenza di introdurre meccanismi che rendessero più competitiva la produzione scientifica e i livelli di performance delle università. In secondo luogo, il mercato del lavoro stesso non incoraggiava particolarmente la modernizzazione del sistema universitario, in quanto non erano richieste figure professionali particolarmente qualificate (Moscati, 2001).

L'introduzione della valutazione nel contesto accademico italiano è stata sollecitata da una doppia pressione, sia di carattere nazionale che internazionale (Colarusso e Giancola, 2020). In primo luogo, la pressione economica esercitata dall'aumentato numero di iscritti alle università, insieme alle crescenti esigenze di razionalizzazione delle risorse da devolvere all'istruzione nel complesso, hanno gradualmente sollecitato il legislatore ad adottare misure che prevedessero una ristrutturazione in termini organizzativi ed economici dell'intero apparato universitario. In secondo luogo, gli incentivi per armonizzare il sistema universitario europeo secondo i principi della

Dichiarazione di Bologna del 1999 (per sostenere la modernizzazione e la competitività delle università europee), e gli effetti del coinvolgimento delle università italiane in organizzazioni internazionali per lo svolgimento di comuni attività di ricerca (come nel caso della Commissione Europea, con i Programmi quadro dell’Unione Europea di ricerca e innovazione, Paradeise *et al.*, 2009), hanno ulteriormente stimolato una riflessione sul rapporto tra università, governo centrale ed apparato produttivo del paese.

A margine del relativo ritardo dell’Italia nell’implementazione di politiche per la valutazione della ricerca, è opportuno inoltre rilevare tre importanti peculiarità del sistema universitario italiano, in relazione al suo ruolo nel sistema produttivo del paese. Nonostante si tratti di un paese fortemente industrializzato, l’Italia investe tradizionalmente molto meno in ricerca e sviluppo (circa il 1,4% del PIL), rispetto alla spesa media dei paesi europei (2,2%)². Tale gap è particolarmente evidente se si confronta il relativamente basso numero di ricercatori per milione di abitanti in Italia (2.307 ricercatori), con lo stesso dato in altri paesi europei (Germania 5.212, Francia 4.715, Austria 5.733, e Spagna 3.001)³. Anche il tasso di investimenti nella ricerca da parte delle imprese private è relativamente basso, fatto principalmente dovuto alle caratteristiche del sistema industriale italiano, nel quale per la maggior parte si tratta di piccole e medie imprese (circa il 95%), con limitata capacità di investimento. D’altro canto, le grandi imprese coinvolgono solo marginalmente ricercatori in progetti di ricerca applicata (Potì e Reale, 2009; Rostan, 2011). In secondo luogo, lo stesso contesto culturale di riferimento sembra non considerare l’istruzione superiore come una risorsa per le imprese, tanto che si registra un tasso di occupazione dei laureati di 4,3 punti più basso della media europea (si parla all’incirca dell’82,1% di laureati con un impiego su tutto il suolo nazionale, con picchi negativi nelle regioni del Sud) (Istat, 2022). Infine, è opportuno sottolineare come il sistema universitario italiano non si sia dimostrato efficace nel ridurre le diseguaglianze sociali ed economiche presenti nel paese: il livello di istruzione dei genitori è ancora un buon indicatore del futuro livello d’istruzione dei figli (Borrelli e Stazio, 2018), con concrete conseguenze sulle future condizioni sociali ed economiche.

Due sono le fasi principali che hanno portato all’introduzione della valutazione della ricerca in Italia:

1. la prima fase, iniziata con la promulgazione della Legge n. 168 del 1989, è caratterizzata dall’assenza di un sistema strutturato nazionale
2. R&D expenditure, dal sito Eurostat: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=R&D_expenditure&oldid=645219.
3. Researchers in R&D, dal sito StatNano: <https://statnano.com/report/s90>.

- per il coordinamento degli esercizi di valutazione della ricerca, ma, al contempo, anche da un grado relativamente alto di sperimentazione da parte delle istituzioni allora attive;
2. durante la seconda fase, con l'istituzione dell'ANVUR (Agenzia Nazionale di Valutazione del sistema universitario e della Ricerca), l'università italiana si è dotata di un sistema centralizzato di coordinamento per la valutazione della ricerca.

È opportuno notare che sia durante il primo che il secondo periodo, l'approccio del legislatore è stato caratterizzato da un forte grado di sperimentazione e da un approccio che possiamo definire *trial-and-error*, avendo difatti modificato periodicamente le procedure per la valutazione della ricerca e le stesse strutture preposte (Abramo e D'Angelo, 2015c).

La prima fase è pertanto caratterizzata dalla promulgazione di leggi che avevano lo scopo precipuo di devolvere responsabilità ed autonomia agli atenei e di proporre gradualmente strutture per il monitoraggio delle attività svolte dalle università:

1. con Legge del 9 maggio 1989, n. 168⁴ si stabilì che le università erano dotate di personalità giuridica ed erano inoltre dotate di autonomia didattica, scientifica, organizzativa, finanziaria e contabile;
2. con Legge 24 del dicembre 1993, n. 537⁵, venne riformato il sistema di finanziamento per le strutture universitarie, con l'istituzione del *Fondo per il Finanziamento ordinario (FFO)*⁶; gli atenei inoltre erano tenuti a dotarsi di *Nuclei di valutazione* interni all'ateneo (dato che solo in alcune università erano già stati predisposti i nuclei di valutazione interni);
3. con Legge 59 del 1997 (Legge Bassanini) vennero poste le basi giuridiche per una più completa definizione di autonomia delle università.

Verso la fine degli anni Novanta vennero istituite le prime organizzazioni a livello nazionale preposte alla elaborazione di dati e alla formulazione di proposte per promuovere la valutazione di efficacia ed efficienza delle attività svolte nell'ambito universitario. Nel 1996 venne istituito l'Osservatorio per la valutazione del Sistema Universitario, che confluì nel CNSVU⁷

4. Legge 9 maggio 1989, n. 168:

www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1989-05-09;168#:~:text=%2D%20L'arte%20e%20la%20scienza,senza%20oneri%20per%20lo%20Stato.www.miur.it/0006menu_c/0012docume/0098normat/1668istitu.htm.

5. Legge 24 dicembre 1993, n. 537:

www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1993-12-24;537,%20art.%2022%20e%2023.

6. Il FFO prevede una quota base di finanziamento ed una quota di riequilibrio.

7. Il CNSVU faceva capo al Ministero dell'Istruzione, dell'Università e della Ricerca, ed era composto da 9 membri eletti dal Governo.

(Comitato Nazionale per la Valutazione del Sistema Universitario) nel 1999, mentre nel 1998 venne fondato il CIVR⁸ (Comitato di Indirizzo per la Valutazione della Ricerca).

Questo primo periodo di sperimentazione e studio delle procedure per la valutazione della ricerca si concluse con lo svolgimento del primo esercizio di valutazione nazionale, la VTR (Valutazione Triennale della Ricerca), che considerava la produzione scientifica realizzata durante il triennio 2001-2003. La VTR, cui potevano partecipare su base volontaria università ed istituti di ricerca, era basata interamente sulla valutazione *peer review*, ed aveva l'obiettivo di incoraggiare il meccanismo di auto-valutazione di atenei e ricercatori (Reale e Pennisi, 2013). L'esperienza della VTR, nonostante alcuni limiti⁹, ha rappresentato un'innovazione nel contesto universitario italiano (Minelli *et al.*, 2008), contribuendo alla formazione di una cultura della valutazione della ricerca, e conseguentemente, alla graduale accettazione da parte della comunità accademica della presenza di un sistema istituzionale su base nazionale. Il processo per l'organizzazione di una seconda edizione di VTR (secondo il Decreto Ministeriale dell'8 luglio 2010¹⁰, che prevedeva lo svolgimento di un esercizio di valutazione della ricerca prodotta nel quadriennio 2004-2008) fu bruscamente interrotto dall'istituzione dell'ANVUR, che avrebbe dovuto inglobare, sostituendoli, il CIVR e il CNSVU (secondo la Legge del 24 novembre 2006, n. 286¹¹). L'introduzione dell'ANVUR¹² sancisce l'inizio del processo di istituzionalizzazione del sistema centralizzato di coordinamento della valutazione della ricerca. Per le modalità piuttosto caotiche con cui si è svolto questo

8. Il CIVR era preposto, prevalentemente, alla promozione dello sviluppo della cultura della valutazione della ricerca, porre le basi per lo svolgimento di un esercizio nazionale di valutazione della ricerca, e migliorare il collegamento istituzionale tra valutazione e allocazione delle risorse (Reale *et al.*, 2007).

9. I principali limiti riguardano la non obbligatorietà dell'adesione all'esercizio di valutazione, la poca trasparenza nella presentazione delle procedure per la valutazione, e la natura discutibile dei criteri di classificazione delle università che partecipavano alla valutazione (Minelli *et al.*, 2008).

10. Decreto Ministeriale dell'8 luglio 2010:
www.gazzettaufficiale.it/eli/id/2010/10/01/10A11544/sg.

11. Legge del 24 novembre 2006, n. 286:
www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2006-11-24;286.

12. Le funzioni dell'ANVUR sono in sintesi: sviluppo e valutazione dei processi di controllo della qualità di didattica e ricerca, con un *focus* specifico sul trasferimento di conoscenze presso la comunità di riferimento; definizione delle procedure, delle soglie minime e dei metodi per la valutazione della produzione scientifica di università ed enti di ricerca; stabilire i criteri per l'allocazione delle risorse pubbliche alle università, tenendo conto dei risultati delle *performance* negli esercizi di valutazione della ricerca, e valutare l'efficacia e l'efficienza dei programmi di finanziamento pubblico.

passaggio dalla prima alla seconda fase, esso ha suscitato dubbi e perplessità sia presso la comunità accademica (Reale e Pennisi, 2013), che nel dibattito politico¹³.

Con Legge 111 del 2011¹⁴, si stabilì che l'ANVUR dovesse essere preposta alla gestione del primo ciclo della VQR, ovvero il primo esercizio di valutazione della ricerca condotto su scala nazionale, a cui erano tenute a partecipare tutte le università. Il primo ciclo della VQR era dedicato alla valutazione dei prodotti di ricerca realizzati durante il periodo 2004-2010. A questo primo esercizio di valutazione della ricerca, ne sono seguiti altri due (2011-2014, 2015-2019), mantenendo tendenzialmente invariata la struttura organizzativa¹⁵, con solo qualche modifica apportata nelle edizioni successive¹⁶ (testimone di un approccio *trial-and-error* da parte delle istituzioni, al fine di mantenere aggiornate le procedure, secondo anche i suggerimenti proposti dalla comunità scientifica stessa). La VQR ha come obiettivi principali: proporre un'istantanea completa della produttività scientifica delle università e degli enti di ricerca in tutto il territorio nazionale; stabilire una lista di università e dipartimenti, come base per l'assegnazione della quota premiale FFO; creare le precondizioni per permettere una comparazione su scala sia europea che internazionale della produttività scientifica italiana (Fondazione CRUI, 2005).

Contestualmente all'istituzione dell'ANVUR e all'organizzazione del primo ciclo di VQR, nel contesto accademico italiano viene introdotto un nuovo meccanismo di assunzione dei ricercatori. La Legge n. 240 del 2010¹⁷ introduce l'ASN, con l'obiettivo di individuare procedure standard

13. Dai vari resoconti delle Commissioni incaricate di discutere la legge emergono perplessità circa l'obiettività dei criteri su cui si fondava la costituzione stessa dell'ANVUR, ritenuti incapaci di produrre effetti concreti sulle storture e le criticità del sistema universitario italiano.

14. Legge 15 luglio 2011, n. 111:
www.gazzettaufficiale.it/eli/id/2011/07/16/011G0153/sg.

15. La VQR si struttura in tre fasi principali: durante la fase della sottomissione, università e istituti di ricerca devono selezionare i prodotti della ricerca tra quelli che ogni ricercatore ha indicato tramite la piattaforma CINECA; durante la fase della valutazione, i GEV (Gruppi Esperti per la Valutazione) definiscono i criteri di valutazione e selezionano i *referee* che supporteranno i GEV nella fase della *peer review* dei prodotti scientifici; nella fase finale, ANVUR pubblica la *Relazione finale* (basata sui report dei GEV), in cui vengono riportati i risultati della valutazione per area disciplinare e università.

16. Il numero dei Gruppi di Esperti per la Valutazione e degli esperti stessi è aumentato nelle edizioni successive; le classi di merito per la valutazione dei prodotti della ricerca sono leggermente cambiate, introducendo come criterio per la valutazione l'*Internazionalizzazione* e la *Terza Missione* (Wang, 2022).

17. Legge 30 Dicembre 2010, n. 240:
www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2010-12-30;240.

e trasparenti che potessero garantire l'assunzione di ricercatori di alto livello, ed ovviare così all'annoso problema del nepotismo dei *baroni* nelle università italiane (Marini e Meschitti, 2018; Abramo, D'Angelo e Rosati, 2014b; Moscati, 2001). Si passa così da un processo di assunzione decentralizzato e fondamentalmente in capo alle università stesse, che potevano indire direttamente i concorsi (Legge n. 210 del 1998¹⁸), ad un meccanismo che si sviluppa in due fasi principali: nella prima fase, i candidati per il conseguimento dell'ASN vengono valutati da una Commissione preposta alla valutazione dei titoli; nella seconda fase, si prevede che sia l'università ad indire concorsi per le posizioni da professore associato o professore ordinario, cui possono partecipare i ricercatori che hanno già ottenuto l'abilitazione. Per come è stata concepita la legge, conseguire l'abilitazione rappresenta la condizione necessaria ma non sufficiente per l'accesso alle posizioni da associato e ordinario, e questo aspetto ha rappresentato uno degli elementi di maggiore criticità della riforma, dal momento che in molti casi, nonostante il conseguimento dell'abilitazione, molti ricercatori non hanno mai potuto accedere a posizioni strutturate.

In conclusione, è opportuno ricordare, in breve, come l'introduzione di VQR e ASN sia stata ampiamente dibattuta nel contesto accademico. Il meccanismo VQR è stato criticato per le conseguenze che i risultati hanno sull'erogazione dei fondi della quota premiale FFO (con potenziali conseguenze negative per università e contesti già in condizioni critiche, Grisorio e Prota, 2020), per il ruolo poco chiaro in alcuni casi svolto dai GEV (Fasanella e Martire, 2017; Baccini, 2014), e sull'adeguatezza dei metodi di valutazione (indicatori, scelta dei tre prodotti da presentare fatta dalle università stesse), che potrebbero restituire un risultato distorto, in termini di qualità dei prodotti scientifici (Rossi, 2016; Abramo, D'Angelo e Di Costa, 2011, Abramo e D'Angelo, 2015; Baccini e De Nicolao, 2016). Per l'ASN, inoltre, una delle maggiori critiche riguarda il fatto che si sia introdotto un meccanismo troppo macchinoso che tende ad escludere dalla professione accademica ottimi candidati non provenienti dal mondo universitario (Abramo, D'Angelo e Rosati, 2014b; Abramo e D'Angelo, 2015a). Questi rischieranno di vedere scaduta la propria abilitazione senza essere riusciti ad accedere all'università (Patti, 2020).

18. Legge 3 luglio 1998, n. 210:
www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:1998-07-03;210.

3. Sulla *research agenda* dei ricercatori: concetti chiave, domande di ricerca e metodologia

L'introduzione della valutazione della ricerca e l'attenzione degli esercizi valutativi sulla qualità e la produttività dei ricercatori hanno suscitato molte riflessioni su come la stessa valutazione abbia potuto influenzare diversi aspetti di cui la produzione scientifica si compone, per la grande pressione a pubblicare un numero sempre maggiori di prodotti a cui i ricercatori sono sottoposti (la cosiddetta *publish or perish culture*). Da qui deriva l'obiettivo della nostra ricerca di approfondire gli effetti della valutazione su come i ricercatori selezionano gli argomenti su cui condurre le proprie ricerche e sul modo in cui orientano la propria *research agenda setting*.

Due sono le ragioni al riguardo: in primo luogo, la selezione dei domini di ricerca rappresenta un momento fondamentale nella carriera del ricercatore (specialmente di quelli più giovani). Può creare legami personali e professionali, con i quali incrementare il proprio “ciclo di credibilità” nella comunità di pari (Latour e Woolgar, 1979). D'altro canto, sebbene uno studio sugli effetti della *research agenda* possa rappresentare un'opportunità per provare a prevedere l'andamento della curva epistemologica di una determinata comunità scientifica, attualmente non ci sono studi che abbiano proposto un'evidenza scientifica solida a favore o meno dell'influenza della valutazione della ricerca sulla *research agenda* dei ricercatori (Gläser e Laudel, 2016; Whitley, 2007).

Nella sezione seguente verrà pertanto proposta una declinazione operativa del concetto di *research agenda setting*, assieme alle tipologie di domanda di ricerca che hanno supportato il lavoro, spiegando la metodologia adottata.

3.1. Concetti chiave e domande di ricerca

Il termine *agenda setting* è mutuato dalle scienze della comunicazione. In contesto accademico è stato utilizzato con diverse accezioni per indicare, tendenzialmente, la lista di argomenti che un ricercatore o un gruppo di ricercatori dovrebbe trattare. Ertmer e Glazewski (2014) per primi proposero una definizione di *research agenda setting*, indicando il processo con cui un giovane ricercatore seleziona gli argomenti di ricerca, in modo da aumentare le probabilità di fare carriera in contesto accademico. Più precisamente, le scelte che compongono il processo di costruzione dell'agenda di ricerca che il giovane ricercatore è chiamato a compiere riguardano l'identificazione dei propri interessi di ricerca, la motivazione nel fare ricerca accademica, l'indi-

viduazione della propria comunità accademica di riferimento, e la selezione delle pratiche di ricerca. Altri studi hanno contribuito all'individuazione dei fattori che favoriscono la strutturazione dell'agenda di ricerca. A grandi linee, essi possono essere raggruppati in due classi: da un lato, i fattori di carattere personale¹⁹; dall'altro, i fattori contestuali²⁰.

Considerando la definizione di Ertmer e Glazewski (2014) e i contributi disponibili sui vari fattori che concorrono al processo di creazione dell'*agenda di ricerca*, possiamo definire operativamente il processo di *research agenda setting* come una sequenza di azioni a breve e medio termine con cui il ricercatore seleziona gli argomenti su cui fare ricerca e le relative impostazioni concettuali. Gli elementi che compongono operativamente la *research agenda* sono: 1) selezione degli argomenti da trattare, 2) motivazioni di ricerca; 3) selezione delle pratiche di ricerca, e 4) la selezione della comunità di pari a cui rivolgersi.

Considerando l'elemento di innovazione rappresentato dall'introduzione nel contesto accademico italiano di VQR e ASN, la nostra ricerca è stata sviluppata a partire dalle seguenti domande di ricerca:

1. L'agenda di ricerca dei ricercatori è stata influenzata dall'introduzione della VQR?
2. Ci sono differenze tra VQR e ASN in termini di effetti sul processo di *research agenda setting* dei ricercatori?
3. È possibile individuare differenze e peculiarità tra discipline scientifiche? Facendo riferimento ai contributi disponibili in letteratura, è stato possibile formulare tre ipotesi:
 1. è ragionevole aspettarsi una certa influenza di VQR sulla *research agenda* dei ricercatori. Secondo la distinzione di Whitley (2007²¹), il sistema

19. Tra i fattori di carattere personale, è possibile individuare la preferenza personale per determinati argomenti (Polanyi, 1958), la percezione del ricercatore di cosa rappresenti la professione accademica (Åkerlind, 2008), il percorso di studi e i benefici che percepisce nello specializzarsi in un determinato ambito, in termini di riconoscimento da parte dei pari (Karvalics, 2013; Pinheiro *et al.*, 2014; Latour e Woolgar, 1979), il ruolo del mentore (Bourdieu, 1999), e la personale tendenza ad assumere un atteggiamento pionieristico nel fare ricerca scientifica (Horta e Santos, 2016 Santos e Horta, 2018; Santos *et al.*, 2020).

20. Tra i fattori contestuali è possibile individuare la collaborazione scientifica con colleghi (Santos, 2020; Latour e Woolgar, 1979; Bourdieu, 1999), il ruolo di eventuali stakeholder (Luukkonen e Thomas, 2016), il dover rispettare determinati standard e soglie di produttività (Shore e Wright, 2004), la disponibilità di fondi (Laudel, 2006).

21. Whitley (2007) distingue tra sistemi nazionali di valutazione della ricerca forti (in cui gli esercizi di valutazione vengono svolti con cadenza regolare, c'è un certo livello di chiarezza e trasparenza degli obiettivi della valutazione e delle procedure, i risultati della valutazione vengono resi pubblici, e vengono tenuti in considerazione nell'allocazione delle risorse per la ricerca), e deboli (in cui gli esercizi valutativi vengono condotti più di rado o non regolarmente, sono organizzati informalmente, e i risultati vengono di rado resi pubblici).

- di valutazione italiano può essere considerato ‘forte’ e quindi con maggiore pressione sulle *research agenda* rispetto a sistemi più deboli;
2. ASN potrebbe influenzare maggiormente il processo di strutturazione della *research agenda* dei ricercatori rispetto a VQR, dal momento che il conseguimento dell’abilitazione riguarda direttamente la carriera dei ricercatori (ed esercita una pressione istituzionale diretta sui ricercatori), mentre VQR ha come obiettivo la valutazione delle performance degli atenei e dei dipartimenti (esercitando una pressione istituzionale indiretta sui ricercatori);
 3. ci si aspetta un’influenza maggiore di VQR e ASN nelle discipline in cui la componente empirica svolge un ruolo importante (come statistica, economia, e sociologia applicata), rispetto alle discipline in cui l’argomentazione teorica è più centrale, come sociologia e scienze politiche (Bonaccorsi, 2015; Lamont, 2009).

3.2. Metodologia

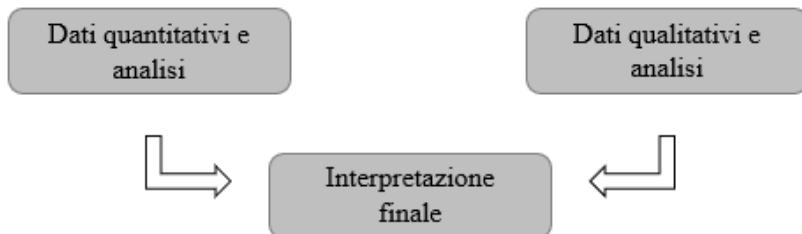
Coerentemente con l’obiettivo di ricerca, si è deciso di adottare un approccio metodologico misto, combinando i risultati di una prima fase di tipo quantitativo (*webmail survey* somministrata alla popolazione *target*), ai risultati di una seconda fase di tipo qualitativo, con ‘interviste informate’ ad esperti di valutazione della ricerca. L’approccio metodologico misto assolve ad una doppia funzione: da un lato, consente di raccogliere il maggior numero di informazioni possibili tramite la somministrazione della *survey*, al fine di rispondere alle domande di ricerca, interrogando direttamente la popolazione di riferimento; dall’altro, tramite l’analisi dei risultati delle interviste informate agli esperti, consente di cogliere elementi nuovi per la comprensione e l’interpretazione degli effetti della valutazione della ricerca sull’agenda di ricercatori. L’agenda rappresenta tutt’ora un argomento di discussione relativamente nuovo, per lo studio del quale l’utilizzo di uno strumento standardizzato come la *survey* potrebbe restituire alcuni importanti passaggi non sufficientemente approfonditi (Creswell *et al.*, 2003).

Come modello di riferimento si è preso il disegno di ricerca fondato sulla convergenza parallela (o triangolare) dei metodi quantitativo e qualitativo proposto da Creswell *et al.* (2003). Secondo questo disegno, i risultati delle due fasi della ricerca tendono a convergere verso una comprensione più composita e puntuale del fenomeno²² (figura sotto). Nella nostra ricerca, inoltre, non c’è

22. Gli autori hanno proposto una distinzione tra disegni di ricerca con approccio metodologico misto: i metodi sequenziali, in cui la raccolta dei dati avviene in momenti logici e cronologici differenti e in cui la raccolta ed elaborazione dei dati nella seconda fase rap-

un legame tra prima e seconda fase a livello di campione sottostante: la scelta dei partecipanti della seconda fase avviene, infatti, sulla base della disponibilità dei rispondenti della *survey* ad essere ricontattati, senza ricorrere a criteri di profilazione sul campione di rispondenti alla prima fase della ricerca.

Fig. 1 - Disegno di ricerca con convergenza parallela



La ricerca si è, dunque, articolata in due fasi concomitanti: la fase di raccolta di dati quantitativi (con la somministrazione della *survey*), e la fase di raccolta di dati qualitativi.

La popolazione *target* è composta di 6297 individui: professori ordinari (32,2%), associati (42,8%), ricercatori a tempo indeterminato (9,1%) e ricercatori a tempo determinato di tipo B (13,8%), afferenti alle Aree 13 (economisti e statistici) e 14 (sociologi e scienziati politici), attivi presso le università italiane, sia pubbliche che private. Abbiamo deciso di rivolgerci ai ricercatori di Area 13 e 14 in quanto, secondo la letteratura disponibile, il modo in cui strutturano la loro agenda di ricerca è tendenzialmente più sensibile all'introduzione della valutazione della ricerca (Whitley, 2007), rispetto ai ricercatori afferenti alle scienze naturali, e, inoltre perché condividono pratiche di ricerca e pubblicazione simili.

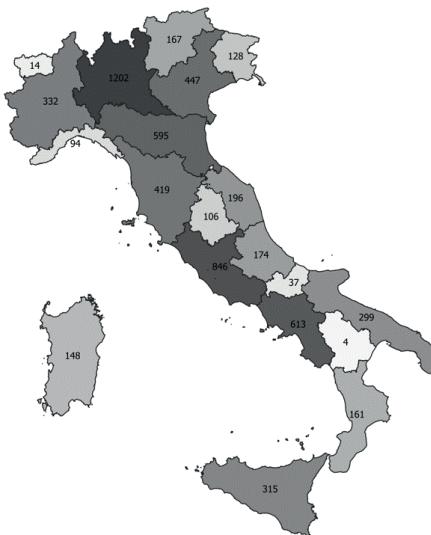
Durante la fase quantitativa, è stata effettuata una *survey online* (che comprendeva sia domande standardizzate che commenti liberi) alla popolazione *target*, composta da sociologi, economisti e scienziati politici attivi presso le università pubbliche e private italiane che, al momento della raccolta degli indirizzi e-mail istituzionali dalle università di afferenza (piattaforma CINECA, estrazione dei dati risalente ad ottobre 2021) ricoprivano le posizioni di professore ordinario, professore associato, ricercatore a tempo indeterminato e ricercatore a tempo determinato di tipo B. La scel-

presentano una specificazione ed approfondimento dell'interpretazione dei dati raccolti durante la prima fase; nei metodi concomitanti, la raccolta dei dati avviene simultaneamente dal punto di vista logico, ovvero la seconda fase non rappresenta una specificazione della prima (Creswell *et al.*, 2003).

ta della *survey* sembrava la più adatta per una serie di ragioni: la *survey* permette, in primo luogo, di rivolgervi direttamente ai soggetti dell'indagine, al fine di studiare in maniera meno distorta da altri fattori di contesto (Gläser e Laudel, 2016), l'influenza dell'introduzione della valutazione della ricerca basandosi su dati primari; in secondo luogo, l'uso della *survey* consente di raggiungere in maniera relativamente facile un numero molto alto di persone distribuite su tutto il territorio nazionale (di cui, appunto, la popolazione *target* è composta), considerando le difficoltà legate al contesto della pandemia da Covid 19, in cui si è svolta questa fase della ricerca (Babbie, 2020; Corbetta, 2014; Favata *et al.*, 2020; Lombi, 2015). In ultima istanza, la *survey* permette di affrontare con maggiore facilità un tema che potrebbe essere sensibile (Lombi, 2015; Tourangeau e Yan, 2007).

La popolazione di riferimento, distribuita su tutto il territorio nazionale (con una maggiore concentrazione in corrispondenza dei centri urbani più popolosi, come Milano, Roma e Napoli, come evidente dalla Fig. 2), è composta per il 60% circa da uomini e per il 40% da donne.

Fig. 2 - Distribuzione regionale della popolazione di riferimento



La *survey* è strutturata in tre sezioni principali:

1. nella prima sezione sono riportate domande su dati biografici del rispondente (età, posizione e anzianità accademica, genere, dimensione dell'ateneo di afferenza, settore scientifico disciplinare di afferenza), utili per le elaborazioni statistiche da svolgere in un secondo momento;

2. nella seconda sezione si chiede all'intervistato di esprimersi in merito al ruolo di fattori personali o di contesto nella creazione dell'agenda di ricerca. In particolare, come fattori personali abbiamo considerato il riconoscimento da parte dei pari, la passione per la ricerca, oltre al contributo dell'attività scientifica alla società; inoltre, come fattori di contesto abbiamo considerato la collaborazione scientifica, il ruolo degli argomenti *mainstream* e di cronaca, assieme al ruolo della partecipazione a bandi competitivi nazionali o internazionali;
3. nella terza sezione, con due domande si chiede all'intervistato di esprimere il grado di influenza di VQR e ASN sul proprio processo di costruzione dell'agenda di ricerca. In particolare, viene richiesto come VQR e ASN abbiano influito sulle diverse componenti dell'agenda di ricerca (argomenti su cui fare ricerca, motivazione nel fare ricerca, pratiche di ricerca e scelta della comunità di pari di riferimento).

La *survey* include anche delle domande aperte, per dare l'opportunità al rispondente di esprimersi liberamente su argomenti complessi: in particolare, in corrispondenza della sezione sulla partecipazione ai bandi competitivi, sulla collaborazione scientifica, su ASN e VQR. L'ultima domanda della *survey*, inoltre, richiede ai rispondenti la disponibilità a partecipare ad un'eventuale seconda fase della ricerca, inserendo il loro indirizzo mail in apposita sezione.

La *survey* è stata progettata da luglio ad ottobre 2021. A seguito di uno studio pilota svolto nel novembre 2021, a cui hanno partecipato prevalentemente colleghi ricercatori del dipartimento di Sociologia dell'Università La Sapienza, e dell'istituto IRCrES-CNR, la versione definitiva della *survey* era pronta verso la fine di dicembre del 2021. Si è scelto di utilizzare LimeSurvey come strumento di supporto e raccolta dati, contattando la popolazione *target* per e-mail (usando l'indirizzo istituzionale), tramite un sistema di invio di e-mail automatizzato, al fine di garantire ai partecipanti il massimo della privacy possibile. Nella mail di presentazione della *survey*, venivano illustrati gli obiettivi dello studio, con una breve descrizione del concetto di agenda di ricerca, in modo da facilitare la compilazione. La *survey* è iniziata il 5 gennaio 2022, registrando nella prima fase un tasso di risposta del 29,51%. A seguito dell'invio della mail di sollecito, inviata il 26 gennaio 2021 (in cui veniva esplicitato l'intervallo di tempo entro cui era possibile rispondere alla *survey*, ovvero l'11 febbraio 2021), si è registrato un tasso di risposta del 36,64%. Nel complesso, un totale di 2299 ricercatori ha risposto alla *survey*.

Durante la seconda fase, sono state svolte interviste informate con esperti di valutazione della ricerca, selezionati tra i rispondenti della *sur-*

vey che hanno espresso la loro adesione per un’eventuale fase successiva della ricerca. Si è scelto di intervistare esperti della valutazione (o anche professori e ricercatori che avessero svolto un ruolo nei vari processi valutativi) per poter coniugare il punto di vista del “valutatore” con quello del “valutato” (Laudel, 2006; Laudel e Gläser, 2007; Morawska, 2018). La traccia del questionario, composta orientativamente di dieci domande, è stata formulata prendendo in considerazione i punti più rilevanti emersi dall’analisi dei commenti liberi della *survey*, con l’obiettivo di approfondire aspetti che, a nostro parere, meritavano un ulteriore approfondimento. Dalla lista di persone che hanno manifestato la loro disponibilità per partecipare alla seconda fase della ricerca (circa 120 persone), sono state selezionati i ricercatori con una certificata esperienza pregressa nella valutazione (a livello di ateneo, come membro dei GEV, come commissario ASN), per un totale di 30 persone. I 30 ricercatori sono stati contattati la prima settimana di settembre 2022 tramite e-mail, nella quale si presentavano gli scopi e le modalità per la partecipazione all’intervista. Sono state svolte 10 interviste tra settembre e ottobre 2021, tramite piattaforma Google Meet. Si è deciso di svolgere in modalità *online* anche le interviste per ovviare a difficoltà legate agli spostamenti, oltre che per evitare problemi legati al contatto interpersonale e alla trasmissione del Covid 19 (dal momento che, nel periodo dello svolgimento delle interviste, un livello di incertezza ancora elevato in merito al virus poteva scoraggiare le persone a partecipare). Le interviste duravano mediamente un’ora, per circa dieci domande. L’intervista veniva registrata²³ per permettere la trascrizione da eseguire manualmente in un secondo momento (dal momento che i software per la trascrizione automatica delle registrazioni riescono difficilmente a riprodurre correttamente il contenuto di tracce audio registrate direttamente da contenuti multimediali).

4. Risultati della ricerca

In questa sezione vengono presentati i risultati della ricerca, della fase della ricerca quantitativa *in primis*, e della fase qualitativa successivamente, proponendo un’interpretazione congiunta, al fine di rispondere alle domande di ricerca, introdotte nel secondo paragrafo.

23. Agli intervistati erano inviati i moduli per la privacy da compilare e firmare, da consegnare prima dell’inizio dell’intervista.

4.1. Risultati della survey

4.1.1. Statistiche descrittive

Prima di passare a presentare i dati relativi all'analisi degli effetti di VQR e ASN sulla formazione dell'agenda di ricerca, è opportuno descrivere brevemente le caratteristiche della popolazione di riferimento. Il tasso di risposta registrato dalla *survey* è del 36,55%, per un totale di 2299 rispondenti, 61% uomini e il 39% donne. Questo dato è coerente con la distribuzione della popolazione di riferimento rispetto alla variabile genere, composta per il 60,98% da uomini e il 39,02% da donne. Più del 40% dei rispondenti è *professore associato* (45,2%), seguito da *professori ordinari* (34,2%), da *ricercatori a tempo determinato di tipo B* (10,4%), e, infine, *ricercatori ruolo ad esaurimento* (10,1%). Anche in questo caso, come per la variabile di genere, la composizione del campione riflette, in linea di massima, le caratteristiche della popolazione di riferimento: il 43% è composto da professori associati, il 32% da professori ordinari, mentre il 14% e l'11% sono, rispettivamente, ricercatori a tempo indeterminato e determinato. Considerando il tasso di risposta in termini assoluti, il settore scientifico disciplinare più numeroso tra i rispondenti della *survey* è Economia Aziendale (13/B), con il 31% dei rispondenti, seguito da Economia (13/A), con il 23% (che, nella popolazione di riferimento, rappresentano rispettivamente il 31% e il 23% del totale, configurandosi come i due settori concorsuali numericamente più rappresentati). Tenendo in considerazione la diversa numerosità dei diversi settori scientifico disciplinari, Sociologia (14/C) e Storia Economica (13/C) sono i settori disciplinari che hanno registrato la frequenza maggiore in termini di partecipazione. Nella *survey* veniva anche richiesto ai rispondenti di indicare le dimensioni dell'ateneo in cui sono attivi, distinguendo tra atenei con meno di 10.000 studenti, atenei con un numero di studenti tra 10.000 e 20.000 studenti, grandi atenei (tra i 20.000 e i 40.000 studenti iscritti), e mega atenei (con più di 40.000 studenti iscritti): dalle analisi, emerge che i rispondenti si distribuiscono in maniera abbastanza omogenea tra le diverse categorie di risposta, con una prevalenza, abbastanza prevedibile, dei professori attivi in mega atenei (36% dei rispondenti). Considerando la variabile relativa agli anni di esperienza accademica (*meno di 5 anni, tra i 6 e i 10 anni, più di 11 anni di esperienza*), il gruppo di ricercatori con più di 11 anni di esperienza è il più numeroso. Tale risultato è in linea sia con riferimento alla domanda sulla posizione accademica (dato che professori ordinari e associati insieme rappresentano circa l'80% dei rispondenti), che all'età media dei rispondenti (51 anni), dal momento che si registra una buona correlazione tra età anagrafica, posizione accademica e anni di esperienza.

4.1.2. Effetti dell'introduzione di VQR e ASN sulla formazione dell'agenda di ricerca

Per rispondere alle domande di ricerca²⁴ poste nel paragrafo precedente, si è deciso di utilizzare come modello la regressione logistica, dato che nella *survey* la modalità di risposta alle domande sugli effetti di VQR e ASN sulla formazione dell'agenda di ricerca è di tipo dicotomico: con il *sì* il rispondente indica che l'introduzione di VQR o ASN ha prodotto effetti sulla formazione della sua agenda di ricerca, mentre il *no* indica che VQR e ASN non hanno prodotto alcun effetto²⁵.

Analizzando la domanda in cui si chiedeva ai rispondenti di esprimersi sugli eventuali effetti dell'introduzione di VQR, emerge che per la maggioranza dei rispondenti (60,07%) non si sono registrati cambiamenti nella propria agenda di ricerca. Questo primo risultato ci permette di rispondere alla prima domanda di ricerca, e quindi di affermare che la pressione istituzionale indiretta sui ricercatori introdotta da VQR non ha prodotto effetti rilevanti sulla formazione delle agende. Dall'applicazione della regressione logistica²⁶, inoltre, emergono tre risultati: nonostante l'effetto dell'introduzione di VQR non sia rilevante per la maggioranza dei rispondenti, 1) l'introduzione di VQR tende in generale ad essere più importante per ricercatrici donne, rispetto a ricercatori uomini; 2) i ricercatori più esperti tenderebbero più probabilmente a cambiare la loro agenda di ricerca in ragione dell'introduzione di VQR; 3) i ricercatori afferenti all'area CUN 13 tendono maggiormente a cambiare la loro agenda di ricerca in ragione dell'introduzione di VQR. La sezione della *survey* in cui viene richiesto ai rispondenti di esprimersi in merito agli effetti dell'introduzione di VQR sulle diverse

24. Domande di ricerca che qui ricordiamo: 1) L'agenda di ricerca dei ricercatori è stata influenzata dall'introduzione di VQR? 2) Ci sono delle differenze tra VQR e ASN in termini di effetti sul processo di *research agenda setting* dei ricercatori? 3) È possibile individuare differenze significative tra discipline scientifiche?

25. Nel modello di regressione logistica, la variabile dipendente è l'assenza/presenza di effetti sulla formazione dell'agenda di ricerca, mentre le variabili indipendenti sono età, posizione accademica, genere, anzianità accademica, dimensioni dell'ateneo di afferenza e settore scientifico disciplinare.

26. Prima dell'applicazione della regressione logistica, si è ritenuto opportuno inserire nel dataset un'ulteriore variabile per ponderare la diversa numerosità di determinati gruppi presenti nel campione, e per controbilanciare un eventuale *bias* di selezione nei sottogruppi della popolazione di riferimento. Il peso è stato calcolato considerando il numero totale di persone per professione accademica, genere, settore scientifico-disciplinare e università di afferenza. Per l'applicazione della regressione logistica, inoltre, si sono dovuti eseguire due modelli differenti, dato che per due variabili indipendenti (ovvero la posizione accademica e l'anzianità accademica), essendo fortemente correlate, può essere molto difficile stimare il loro coefficiente di regressione in maniera affidabile.

componenti dell'agenda di ricerca (che, ricordiamo, sono la scelta degli argomenti di ricerca, la motivazione nel condurre una ricerca, la scelta delle pratiche di ricerca e la scelta della comunità di pari di riferimento) sembrerebbe confermare quanto emerso precedentemente: i ricercatori più esperti e i ricercatori afferenti all'area 13 affermano che la formazione della loro agenda di ricerca è cambiata a seguito dell'introduzione della VQR, con riferimento a tutti gli elementi di cui l'agenda di ricerca si compone.

Analizzando invece le risposte alla domanda sugli effetti dell'introduzione di ASN sulla formazione dell'agenda di ricerca, il 50,15% dei rispondenti ha risposto affermativamente. Dai dati raccolti, pertanto, emerge che la pressione istituzionale diretta esercitata dall'introduzione di ASN sui ricercatori ha prodotto maggiori effetti in termini di condizionamento sulla formazione dell'agenda. Come nel caso di VQR, dall'analisi del modello di regressione logistica sugli effetti dell'introduzione di ASN sulla formazione dell'agenda di ricerca emergono tre risultati particolarmente rilevanti: 1) anche in questo caso, le ricercatrici donne e i ricercatori giovani tendono più probabilmente a cambiare la loro agenda in ragione dell'introduzione di ASN; 2) i ricercatori afferenti alle aree CUN 13 e 14/B tendono più probabilmente a cambiare la loro agenda di ricerca. Tali risultati vengono confermati dai dati raccolti in corrispondenza delle domande sugli effetti di ASN sui diversi componenti dell'agenda: i ricercatori meno esperti e ricercatori afferenti all'area 13 affermano di aver cambiato in misura rilevante la scelta degli argomenti su cui condurre le proprie ricerche, le proprie pratiche di ricerca, e le motivazioni per le quali fare ricerca.

Come visto precedentemente, le analisi dei dati raccolti con la *survey* ci permettono di rispondere alle tre domande. In primo luogo, possiamo rispondere alla prima domanda e affermare che l'introduzione di VQR non ha prodotto effetti rilevanti sulla formazione dell'agenda di ricerca dei ricercatori, confutando così la prima ipotesi precedentemente formulata. Possiamo però anche affermare che l'introduzione di VQR ha rappresentato un fattore relativamente importante di influenza sulla formazione dell'agenda dei ricercatori per il 40% dei rispondenti alla *survey*, e ciò è comprensibile, nonostante eserciti una pressione istituzionale indiretta sui ricercatori. I risultati di VQR determinano infatti parte dell'allocazione dei fondi per le università previsti da FFO, essendo quello italiano un sistema di valutazione *forte* (Whitley, 2007; Wang, 2022). L'introduzione di VQR sembrerebbe avere infatti indotto soprattutto i ricercatori più esperti a modificare la loro agenda di ricerca. Essi sono tendenzialmente più coinvolti in posizioni di maggiore responsabilità a livello universitario e dipartimentale, e pertanto più tenuti a rispondere personalmente dei risultati del proprio dipartimento o settore di competenza. In secondo luogo, possiamo

rispondere alla seconda domanda di ricerca affermando che l'introduzione di ASN ha prodotto effetti maggiori rispetto a VQR, confermando così l'ipotesi iniziale. Secondo i dati raccolti, l'introduzione dell'ASN ha rappresentato un momento di svolta per i ricercatori che hanno dovuto conseguire l'abilitazione scientifica da associato (o per i ricercatori relativamente giovani che devono ancora conseguirla), dal momento che rappresenta il passaggio obbligato (ma spesso, purtroppo, non sufficiente) per uscire dalla precarietà e conseguire la stabilizzazione in ambiente accademico. Possiamo anche rispondere alla terza domanda, affermando che l'area CUN 13 sembrerebbe più incline a modificare la propria agenda di ricerca rispetto all'area 14, confermando la nostra terza ipotesi, secondo la quale i settori scientifici in cui la base empirica rappresenta una parte importante sono tendenzialmente più inclini, secondo la letteratura disponibile (Lamont, 2009; Bonaccorsi, 2015) ad adattarsi ai sistemi di valutazione, rispetto ai settori disciplinari in cui il dibattito teorico svolge un ruolo rilevante. A margine della discussione sui risultati della *survey*, è opportuno proporre un brevissimo spunto di riflessione sulla disparità di genere, nel contesto della valutazione della ricerca. Considerando infatti i risultati che emergono dalla *survey*, si evince che le ricercatrici tendono ad essere in generale più propense a modificare la propria agenda di ricerca rispetto ai ricercatori. Tale risultato è abbastanza prevedibile, se si considerano le maggiori difficoltà che le ricercatrici tendenzialmente riscontrano nel tentativo di conciliare il tentativo di fare carriera nel settore accademico con le esigenze della vita privata (Baldarelli *et al.*, 2016; Bozzon *et al.*, 2017).

4.2. Risultati delle interviste informate

Le interviste informate con gli esperti della valutazione della ricerca in Italia, selezionati tra i rispondenti della *survey* che hanno accettato di essere eventualmente contattati per la seconda parte della ricerca, sono state condotte da settembre a metà ottobre 2022. Le interviste condotte sono di tipo semi-strutturato, con un protocollo assunto come punto di partenza per la discussione, ma che non esaurisce necessariamente gli argomenti da affrontare. Nel protocollo di intervista sono elencati i seguenti argomenti:

1. il ruolo di VQR e ASN nella formazione dell'agenda di ricerca dei ricercatori;
2. la pressione delle università di afferenza nell'influenzare l'agenda di ricerca dei ricercatori;
3. ruolo di VQR e ASN sulla propensione a condurre ricerca multidisciplinare, sulla collaborazione scientifica e sull'internazionalizzazione della ricerca italiana;

4. ruolo di VQR e ASN nel modificare la concezione della ricerca e dei suoi stessi obiettivi, con un *focus* specifico sulla produzione scientifica dei giovani ricercatori.

Per l'analisi delle interviste (trascritte manualmente dalle registrazioni), si è proceduto inizialmente isolando gli estratti delle interviste che riguardavano l'introduzione di VQR e ASN, cercando di individuare una regolarità nelle interviste. Dalle interviste, è stato possibile individuare dei concetti chiave²⁷, sulla base dei quali gli estratti delle interviste sono stati classificati. Le parti rimanenti delle interviste sono state in seguito strutturate a seconda degli argomenti più rilevanti, consentendo il disegno di un grafico in grado di rappresentare l'atteggiamento degli intervistati e della loro organizzazione attuale: con il termine *integrati* indichiamo gli intervistati che hanno espresso un'opinione particolarmente favorevole all'introduzione di VQR e ASN, mostrando un livello mediamente alto di adattamento alle riforme e valutandone principalmente gli effetti positivi e le opportunità per la qualità e la produttività scientifica, gli *apocalittici*, invece, sono coloro che manifestano un atteggiamento particolarmente sfavorevole verso VQR e ASN, evidenziando criticità nella loro implementazione e considerando prevalentemente negative le conseguenze sulla qualità della produzione scientifica nazionale e sulle prospettive di carriera dei giovani ricercatori, e, infine, i *neutrali* sono coloro che non rilevano particolari criticità legate all'introduzione di VQR e ASN.

Di seguito vengono illustrati gli elementi più rilevanti emersi dalle interviste, riguardo all'influenza di VQR e ASN sulla formazione dell'agenda di ricerca, e, più in generale, sulla produzione scientifica.

4.2.1. Effetti di VQR

Dalle interviste, emerge un consenso generale sul fatto che l'introduzione di VQR abbia rappresentato nel contesto accademico un importante elemento di novità, stimolando una corsa alla pubblicazione che ha come primaria conseguenza quella di concentrarsi su argomenti *mainstream* o poco rischiosi, con ricadute negative sull'originalità e sulla qualità dei prodotti scientifici. L'introduzione di VQR è percepita da più intervistati come

27. I concetti chiave emersi sono: *influenza di VQR e ASN sull'agenda di ricerca; ruolo di VQR e ASN sulla ricerca multidisciplinare; effetti sull'internazionalizzazione della ricerca; effetti a lungo termine sulla produttività scientifica dei ricercatori; effetti sulla collaborazione scientifica; effetti sulla concezione della ricerca; effetti di VQR e ASN sulla carriera dei giovani ricercatori; ruolo sulla gestione e disponibilità dei fondi per la ricerca.*

inefficace rispetto ad uno dei suoi principali obiettivi, ovvero quello di stimolare un'autoriflessione sulla qualità dei prodotti scientifici. Al contrario, per alcuni intervistati sembrerebbe incoraggiare una produzione scientifica ripetitiva e meno originale nei temi e nell'approccio metodologico, per massimizzare la probabilità di pubblicazione:

L'idea che si ha è che un po' tutti dicano le stesse cose. Vengono così meno l'originalità della ricerca scientifica, visto che si fa ricerca su un tema su cui tantissimi hanno detto già tantissimo; inoltre, si deve riflettere che, come sociologi, si devono tenere in considerazione i temi che emergono, ma, con la rincorsa alla pubblicazione, molti di questi articoli mancano di una dimensione del pensiero tipico della sociologia, facendo prevalere una dimensione descrittiva, più che analitica. Come se si facesse solo della microsociologia, senza però una dimensione di tipo esplicativo (Intervistato 14_PO_5).

L'enfasi sulla quantità dei prodotti scientifici, piuttosto che sulla qualità, secondo alcuni intervistati potrebbe determinare una mutata concezione dell'approccio alla ricerca, per la quale i ricercatori hanno come obiettivo principale la pubblicazione scientifica ed il frazionamento delle fasi della ricerca in risultati pubblicabili, piuttosto che mantenere il *focus* sulle domande della ricerca e sul processo scientifico di per sé. Gli *apocalittici* della VQR criticano in particolar modo l'eccesso di enfasi, a loro parere non del tutto giustificata dalle evidenze scientifiche, per i criteri che vengono utilizzati per valutare la qualità di un prodotto scientifico (*originalità, rigore metodologico, impatto*), che potrebbero produrre distorsioni quando si conduce una ricerca, per la necessità di rispettare i suddetti criteri.

Gli *integrati*, al contrario, valutano positivamente il contributo di VQR alla “sprovincializzazione” della produzione scientifica italiana. Inoltre, da parecchie interviste emerge un’attitudine abbastanza positiva sugli effetti di VQR in merito all'internazionalizzazione della produzione scientifica, con ricadute positive sulla standardizzazione della produzione scientifica italiana, in modo da renderla più competitiva nel panorama internazionale e maggiormente aperta al dialogo con colleghi internazionali:

L'unico effetto assolutamente positivo [di VQR] è stata una maggiore spinta verso l'internazionalizzazione della nostra comunità accademica. Questo è un effetto congiunto di queste nuove forme di valutazione che sono state introdotte. Su questo siamo sicuri che è interessante l'impatto di VQR, che spinge anche le riviste italiane a internazionalizzarsi.... Comunque, la sprovincializzazione è un effetto positivo sotto tanti punti di vista (Intervistato 14_PO_F).

D’altro canto, alcuni intervistati rilevano come il dover ricondurre la produzione scientifica nei rispettivi settori scientifici disciplinari rappresenti un freno sia per la collaborazione multidisciplinare, che per la collaborazione con autori stranieri. Dalle interviste emerge molto spesso come criticità il fatto che i bandi della VQR vengono pubblicati successivamente rispetto al periodo a cui la valutazione fa riferimento, impedendo così ai ricercatori di strutturare adeguatamente strategie di ricerca e pubblicazione per massimizzare i risultati.

4.2.2. Effetti di ASN

Rispetto agli interventi degli intervistati su VQR, ASN ha stimolato riflessioni piuttosto unanimi, sulle sue conseguenze sulla produzione scientifica e la qualità della ricerca dei ricercatori che ambiscono a conseguire l’abilitazione. Emerge un consenso generale innanzitutto sul fatto che l’introduzione di ASN abbia prodotto effetti rilevanti in termini di influenza sulla formazione dell’agenda di ricerca dei ricercatori. Il meccanismo di ASN indurrebbe pertanto i giovani ricercatori ad adeguare gli argomenti e le pratiche con cui condurre ricerche alla massimizzazione delle possibilità di pubblicare un numero sufficiente di prodotti scientifici, fino al momento del conseguimento dell’abilitazione scientifica.

L’ASN, certamente, ed è evidente a tutti, ha provocato questa corsa a pubblicare; quindi, è evidente che se cominci a dire che tante pubblicazioni servono come requisito di accesso per partecipare all’Asn, è legittimo ipotizzare che ci sia o una maggiore superficialità nel trattare il medesimo tema o la tendenza a trattare tanti temi. Quindi sicuramente l’Asn ha prodotto degli effetti strutturali e questi si vedranno tra 5/10 anni quando tutti i giovani assunti cominceranno ad avere dei problemi perché stiamo crescendo dei giovani che pensano che fare il professore sia scrivere dei paper, per cui finiscono in un’istituzione e non sanno gestirla... Sul modo di fare ricerca e di concepire il mestiere, vedo degli effetti strutturali della Asn (Intervistato 14_PO_M).

Per come l’ASN è strutturata, la quasi totalità degli intervistati afferma che rende molto difficile la collaborazione scientifica (sia con colleghi italiani che stranieri) e la ricerca multidisciplinare, per la difficoltà di ricondurre i propri prodotti scientifici in un unico settore scientifico-disciplinare. Emerge spesso, da parte degli intervistati, uno sforzo molto attivo nel sostenere i giovani ricercatori durante il percorso verso l’abilitazione, viste le oggettive difficoltà che essa comporta. Anche in questo caso, gli intervistati percepiscono come un vincolo negativo i criteri preposti alla valutazione dei titoli dei candidati per il conseguimento dell’abilitazione scientifica, tali da condizionare negativamente la produzione scientifica dei giovani ricercatori.

5. Conclusioni

I risultati della ricerca sugli effetti dell'introduzione di VQR e ASN sull'agenda di ricerca dei ricercatori italiani sollevano importanti riflessioni su come la comunità accademica stia rispondendo a queste misure, a distanza di una decina di anni dalla loro implementazione. Per quanto riguarda VQR, le interviste e i commenti liberi della survey rivelano che, sebbene venga percepita come un ulteriore onere amministrativo da gestire, non ha prodotto cambiamenti significativi nelle modalità di ricerca dei ricercatori. Sebbene VQR eserciti una pressione indiretta sulla produttività, attraverso la valutazione delle performance universitarie, non ha determinato deviazioni sostanziali nella “curva della conoscenza” all'interno delle singole discipline. In altre parole, la sua influenza appare meno diretta sulle pratiche quotidiane di ricerca dei singoli ricercatori.

Al contrario, l'introduzione di ASN ha avuto effetti decisamente più concreti, soprattutto sui ricercatori più giovani, che si trovano a dover affrontare una crescente pressione per pubblicare un elevato numero di articoli scientifici in *outlet* prestigiosi. La competizione tra colleghi e la necessità di soddisfare i criteri imposti per ottenere la *tenure* accademica hanno portato molti giovani ricercatori a concentrarsi maggiormente sulla quantità che sulla qualità della loro produzione scientifica, con possibili effetti sulle loro pratiche di ricerca. Tale pressione istituzionale potrebbe influenzare in modo significativo la “curva della conoscenza”, portando a un adattamento delle *research agenda* per rispondere alle esigenze di carriera e di visibilità accademica. I giovani ricercatori sembrano adattarsi a queste pressioni adottando strategie che si concentrano più sulla quantità di *output* rispetto all'originalità o al contenuto innovativo delle loro ricerche. Questo fenomeno potrebbe limitare la diversità tematica della ricerca scientifica, a favore di tematiche più *mainstream* e riconosciute dalle comunità scientifiche e dai sistemi di valutazione accademici.

In sintesi, i risultati emersi indicano che, mentre VQR ha avuto un effetto più limitato sulle scelte di ricerca individuali, ASN ha comportato cambiamenti più evidenti, in particolare tra i giovani ricercatori. Questo fenomeno pone interrogativi sulla direzione che sta prendendo la ricerca accademica in Italia, in particolare per quanto riguarda la qualità e l'originalità dei temi trattati. Sebbene sia ancora presto per prevedere gli effetti a lungo termine, è chiaro che la pressione per conformarsi ai requisiti istituzionali sta modellando in modo significativo le agende di ricerca, con possibili implicazioni negative per l'innovazione e la ricerca di frontiera.

4. Collaborazioni di ricerca e mobilità degli accademici: sono visibili effetti della valutazione?

di *Andrea Orazio Spinello, Antonio Zinilli, Emanuela Reale*

1. Introduzione

La crescente propensione allo sviluppo di collaborazioni di ricerca e l'intensificazione della formazione di reti internazionali costituiscono due aspetti distintivi del modo di fare scienza contemporaneo (Wagner e Leydesdorff, 2005; Wutchy *et al.*, 2007). Sempre più frequentemente, gli studiosi cooperano per affrontare le sfide scientifiche nell'ambito di nuove pubblicazioni o progetti di ricerca, potendo altresì avvalersi dei progressi delle tecnologie nel campo delle comunicazioni che hanno minimizzato la rilevanza delle barriere geografiche (Nielsen, 2011). Nello stesso tempo, il mondo accademico sta sperimentando una maggiore integrazione su scala globale, risultato dell'azione congiunta di diversi fattori, tra i quali un ruolo di rilievo è assunto dalle dinamiche di mobilità degli studiosi oltre i confini nazionali (Welch, 1997; Huang *et al.*, 2013), fenomeno che ha trovato ampio sviluppo nel contesto europeo, dove la promozione della mobilità geografica è emersa tra gli elementi caratterizzanti del Processo di Bologna (European Ministers in Charge of Higher Education, 1999).

Numerosi studi hanno indagato i fattori che contribuiscono a motivare gli studiosi a intraprendere attività collaborative a livello nazionale e internazionale, o iniziative di mobilità all'estero. La letteratura ha evidenziato come la decisione di collaborare possa derivare da determinanti connesse all'interesse scientifico, a un miglioramento delle prestazioni o della reputazione dello studioso, o dipendere da elementi contestuali legati all'area di ricerca (Beaver e Rosen, 1978; Beaver, 2001; Lee e Bozeman, 2007; Kyvik e Reymert, 2017; Abramo, D'Angelo e Murgia, 2017). Per quanto concerne la mobilità internazionale, i principali fattori che giocano un ruolo nella scelta di intraprendere periodi di ricerca all'estero

includono l'incremento del capitale relazionale con la creazione di reti, l'aspirazione all'avanzamento nella carriera accademica e il desiderio di migliorare la produzione e le performance scientifiche (Jonkers e Cruz-Castro, 2013; Franzoni *et al.*, 2014; European Commission, 2017; Cañibano *et al.*, 2020). Allo stato attuale, rimane tuttavia poco approfondito il tema del possibile contributo incentivante della valutazione cui sono sottoposti gli accademici e le istituzioni di appartenenza verso queste due dimensioni dell'attività scientifica.

Il presente capitolo intende esplorare l'eventuale effetto esercitato dalla valutazione nazionale – prendendo in considerazione sia l'Abilitazione Scientifica Nazionale (ASN) che la Valutazione della Qualità della Ricerca (VQR) – nel costituire un fattore motivante per gli accademici verso l'intrapresa di collaborazioni di ricerca e di iniziative di mobilità all'estero. Il nostro obiettivo è comprendere se e in che misura la valutazione contribuisca alla promozione di entrambi questi aspetti legati all'organizzazione del lavoro scientifico, rispetto ad altri fattori che la letteratura ha individuato come influenti. A questo scopo, adotteremo un approccio focalizzato sull'analisi della percezione individuale degli stessi studiosi, tenendo in considerazione la varietà di contesti istituzionali, disciplinari e geografici in cui essi operano. La base empirica dello studio sarà costituita dai dati derivanti dalla *survey* nazionale implementata nell'ambito del progetto PRIN 2017 “Gli effetti della valutazione sulla ricerca accademica: produzione della conoscenza e problemi metodologici”, rivolta ad un campione probabilistico di accademici afferenti ad atenei italiani (Spinello *et al.*, 2021).

Nel prossimo paragrafo, presenteremo brevemente un quadro dei fattori, individuati dalla letteratura, considerati influenti per l'attivazione di collaborazioni scientifiche e l'intrapresa della mobilità internazionale da parte degli studiosi. La terza sezione metterà a fuoco l'obiettivo e l'approccio metodologico dello studio. La quarta sezione discuterà i risultati dell'indagine. Infine, saranno presentate alcune riflessioni conclusive.

2. Quadro teorico sulle dimensioni di studio

Lo sviluppo di collaborazioni di ricerca e la partecipazione a iniziative per l'internazionalizzazione individuale, quali quelle offerte dai programmi di mobilità all'estero, costituiscono due aspetti significativi, talvolta interconnessi, nell'attuale contesto dell'organizzazione del lavoro scientifico. La letteratura ha individuato alcuni fattori particolarmente rilevanti nell'indirizzare le decisioni degli studiosi in relazione a queste due dimensioni connesse alle attività di ricerca.

2.1. Collaborazioni di ricerca e determinanti individuali

La collaborazione tra studiosi può esprimersi in molte forme, che condividono tutte gli elementi di interazione e cooperazione, che possono estendersi per tutta la durata o per una specifica parte di un'attività di ricerca condivisa (Katz e Martin, 1997). Può essere instaurata tra colleghi afferenti alla medesima unità di ricerca o coinvolgere soggetti afferenti ad altre istituzioni; può formarsi tra singoli o avvenire nell'ambito di grandi gruppi di lavoro; tra scienziati del medesimo campo disciplinare o di campi diversi; assumere dinamiche gerarchiche o di pari status. Allo stesso modo, diverse sono le motivazioni che possono concorrere alla decisione di collaborare di uno scienziato.

In primo luogo, la collaborazione può essere determinata dall'*interesse scientifico*, e in particolare dalla necessità di creare sinergia tra competenze e/o campi disciplinari diversi. Uno scienziato potrebbe trovarsi di fronte a questioni eccessivamente complesse da affrontare da solo o nell'ambito di una singola disciplina. Pertanto, diventa essenziale integrare una gamma di competenze più ampia di quelle a disposizione del singolo. In questo senso, l'aggregazione di una massa critica di studiosi con conoscenze specialistiche o multidisciplinari consente di superare limitazioni e conseguire la varietà cognitiva necessaria per affrontare determinate questioni scientifiche (Heinze *et al.*, 2009; Nielsen, 2011). La necessità di amalgamare competenze diverse può accompagnarsi al bisogno di accedere a risorse e attrezzature non disponibili, in parte o totalmente, presso la propria struttura di afferenza. Questo aspetto può spingere ad intraprendere progetti con scienziati o gruppi di ricerca che dispongono di un equipaggiamento adeguato (Beaver e Rosen, 1978). L'attività di ricerca in collaborazione offre inoltre la prospettiva di ottimizzare l'efficienza delle ricerche grazie alla possibilità di distribuire le attività necessarie al conseguimento degli obiettivi scientifici. Infine, il “fattore tempo” nella produzione di *output* scientifici, in risposta a scadenze progettuali o in vista della risoluzione di sfide incombenti, può rappresentare un elemento non trascurabile nella scelta di collaborare. Attraverso l'azione collaborativa, si compie un processo di suddivisione modulare del lavoro, in cui ciascuno studioso assume la responsabilità di una particolare porzione dell'attività e può esaminare l'operato degli altri (Nielsen, 2011).

Un secondo gruppo di determinanti riguarda fattori legati a *motivazioni o strategie personali*. Il lavoro di ricerca congiunto consente di instaurare relazioni sociali con altri scienziati, incrementando la riserva di capitale sociale. Ciò può andare a soddisfare un bisogno psicologico, poiché la collaborazione, alimentando l'interazione, allevia la sensazione di isolamento (Fox e Faver, 1984). Nondimeno, la costituzione di reti relazionali collabo-

rative può essere intrapresa per rispondere a un'istanza competitiva tesa al miglioramento delle prestazioni individuali e della reputazione, in vista di benefici per la propria carriera accademica (van Rijnsoever *et al.*, 2008). Molta letteratura ha dimostrato come le collaborazioni di ricerca, sia all'interno della propria istituzione che tra soggetti afferenti a istituzioni diverse o a livello internazionale, generino un impatto positivo sulle performance scientifiche, influenzando la produttività e la qualità del lavoro di ricerca, nonché le probabilità di pubblicazione su riviste internazionali (Lee e Bozeman, 2005; Wutchy *et al.*, 2007; Abramo, D'Angelo e Di Costa, 2009b; Abramo, D'Angelo e Murgia, 2017). Nello scegliere i *partner* di ricerca, gli scienziati possono esprimere preferenze variabili: in particolare, le collaborazioni con ricercatori affermati possono incrementare la visibilità (Beaver, 2001) e quelle internazionali spesso si traducono in un aumento degli indici citazionali (Wagner e Leydesdorff, 2005). Ad ogni modo, alcuni studi dimostrano che le collaborazioni tendono ad essere avviate tra soggetti che condividono una qualche similarità, anche in termini di reputazione, con effetti di vantaggio cumulativo (Zinilli, 2016).

Infine, la propensione individuale a collaborare non può prescindere da *fattori contestuali*. Tra questi possono esservi l'eventuale pressione dell'istituzione di appartenenza, l'influenza del mentore o le caratteristiche specifiche dell'area o del settore di ricerca, cosicché ogni dominio di studi sviluppa modalità specifiche per l'instaurazione e l'evoluzione delle collaborazioni (Kyvik e Reymert, 2017). Inoltre, la costituzione di *network* può essere influenzata da caratteristiche di similarità organizzativa tra le istituzioni di afferenza, in termini di reputazione e capacità di ricerca (Zinilli *et al.*, 2023), o di prossimità geografica (Bergé, 2017).

2.2. Motivazioni per la mobilità internazionale degli accademici

Lo sviluppo di programmi di mobilità, per cui studiosi acquisiscono risorse per trasferire temporaneamente la propria attività oltre i confini nazionali della loro istituzione, ha acquistato ulteriore slancio nel contesto di una maggiore interconnessione globale e con l'implementazione di politiche della ricerca volte a rafforzare la cooperazione scientifica a livello internazionale (European Commission, 2012; Huang *et al.*, 2013).

Durante le diverse fasi della carriera, gli accademici possono decidere di intraprendere periodi di studio o ricerca all'estero, spostandosi presso altre università, verso organizzazioni pubbliche o approcciandosi al settore privato. Queste iniziative costituiscono una piattaforma per impegnarsi in progetti di ricerca scientifica congiunti, acquisire metodologie e competenze specializzate

e nuove prospettive nella ricerca, fungendo in ultimo da importanti catalizzatori per la diffusione e lo scambio della conoscenza (Welch, 1997; Kim e Locke, 2010; Teichler, 2017). La soglia dei tre mesi è solitamente presa a riferimento dagli studi internazionali per distinguere la mobilità di breve durata da quella di lunga durata (de Weert, 2013; European Commission, 2017).

Da un lato, la letteratura ha analizzato il fenomeno della mobilità degli scienziati dalla prospettiva dei vantaggi a livello sistematico, poiché i flussi verso una specifica istituzione, e in generale verso un determinato Paese o regione, fungono da indicatore di attrattività delle destinazioni della mobilità; all'altro alcuni studiosi si sono soffermati sui molteplici effetti positivi a livello individuale, che possono essere interpretati quali fattori motivanti nell'intrapresa della mobilità (De Filippo *et al.*, 2009; Jonkers e Cruz-Castro, 2013; Veugelers e Van Bouwel, 2015). Tenuto conto che lo spettro delle opportunità di mobilità contempla una varietà di esperienze diverse attivate sulla base di motivazioni diverse e intrecciate, come evidenziato da alcune *survey* condotte a livello europeo (Guthrie *et al.*, 2017; European Commission, 2017), i principali fattori di spinta che possono essere desunti riguardano la *creazione di reti internazionali*, l'*avanzamento di carriera*, l'opportunità di *incremento della produzione e delle performance scientifiche*.

Quando gli accademici trascorrono periodi presso istituzioni straniere si confrontano con ambienti professionali diversi che incrementano il capitale sociale in termini di reti relazionali, aprendo la strada al consolidamento di connessioni con colleghi internazionali. Specialmente all'inizio della carriera queste opportunità sono considerate con particolare interesse dai giovani ricercatori (Musselin, 2004), tuttavia una gamma di incentivi alla mobilità internazionale può subentrare a diverse fasi della carriera (Cañibano *et al.*, 2020). La motivazione della creazione di reti può intrecciarsi a esigenze professionali del proprio percorso scientifico, tra cui l'opportunità di accrescere reputazione e prestigio (Stephan, 2015), l'accesso a risorse altrimenti non disponibili, e aprire a possibilità di partecipazione a progetti di ricerca internazionali (Barjak e Robinson 2008).

Molti studi hanno dimostrato la relazione positiva tra mobilità degli accademici e l'attivazione di collaborazioni di ricerca (Jonkers e Cruz-Castro, 2013; Patrício *et al.*, 2018; Haupt, 2022), che possono condurre a un effettivo aumento della produttività e delle performance scientifiche (Jonkers and Tijssen, 2008; Franzoni *et al.*, 2014).

A livello individuale vi sono fattori che possono tuttavia interferire con la mobilità internazionale. Ad esempio, variabili demografiche come il genere e le circostanze familiari, come la presenza di figli, possono avere un ruolo (Ackers, 2008); e ancora, aver intrapreso periodi all'estero durante l'ultima fase dell'educazione terziaria e nelle prime fasi della carriera acca-

demica può predire l'intrapresa di periodi all'estero successivi (Reale *et al.*, 2019). Naturalmente, anche fattori contestuali vanno presi in considerazione. Alcuni studi sottolineano come i *pattern* di mobilità varino tra diverse aree e settori di ricerca (de Weert, 2013; Reale *et al.*, 2019).

3. Obiettivo e metodo della ricerca

L'analisi degli adattamenti delle pratiche di ricerca in relazione alle valutazioni esterne rappresenta un ambito d'indagine in cui molti studi tendono a rimanere a un livello teorico e le indagini empiriche spesso scelgono di approfondire aspetti specifici nelle pratiche di produzione della conoscenza (de Rijcke *et al.*, 2016). Per quanto riguarda il contesto italiano, non mancano le applicazioni di studi empirici, sia considerando l'ASN che la VQR, che prendono in considerazione aspetti quali le strategie di scelta dei siti di pubblicazione e i comportamenti citazionali degli studiosi (Scarpa *et al.*, 2018; Baccini *et al.*, 2019; Abramo, D'Angelo e Grilli, 2021). La maggior parte degli studi internazionali adopera approcci basati su metodi di analisi bibliometrica, mentre l'utilizzo della prospettiva di percezione degli studiosi come base empirica si riscontra più raramente, o comunque nell'ambito di ricerche a metodologia mista. Un esempio del primo tipo è costituito dallo studio di Cañibano *et al.* (2018), che ha adottato un approccio basato su interviste in profondità per investigare l'effetto delle valutazioni sulla ricerca in Spagna, concentrandosi su dipartimenti universitari attivi nell'ambito delle scienze storiche. Un esempio del secondo è la ricerca di Hammarfelt e de Rijcke (2015) applicata a una facoltà di un'università svedese, in cui gli autori hanno combinato l'utilizzo di metriche bibliometriche con questionari sulle pratiche di pubblicazione.

Questo contributo intende posizionarsi nell'ambito delle ricerche sugli effetti della valutazione della ricerca tese a esplorare le percezioni soggettive e il punto di vista degli studiosi. Vogliamo esplorare il possibile contributo della valutazione nazionale, sia declinata nei termini dell'ASN che in quelli della VQR, nel motivare o promuovere le decisioni degli accademici riguardanti l'attivazione di collaborazioni di ricerca e il perseguimento di iniziative di mobilità internazionale.

3.1. Rilevazione del dato

Questo studio utilizza dati provenienti dalla *survey* nazionale condotta in Italia tra ottobre 2020 e marzo 2021 nell'ambito del progetto PRIN 2017 “Gli effetti della valutazione sulla ricerca accademica: produzione della

conoscenza e problemi metodologici” (cfr. par. 4 del Capitolo 1 di questo volume e Spinello *et al.*, 2021). L’indagine è stata implementata attraverso la somministrazione online di un questionario strutturato, comprendente anche sezioni a commento libero, a un campione di 1.365 studiosi, rappresentativo della comunità accademica italiana. Sono state raccolte informazioni su comportamenti, opinioni e aspetti motivazionali riguardanti alcune dimensioni associate al lavoro scientifico, tra cui le collaborazioni di ricerca e la partecipazione a iniziative di mobilità internazionale. I partecipanti sono stati invitati a esprimere le loro percezioni nell’ambito della loro esperienza durante gli ultimi dieci anni di carriera accademica (o per l’intera durata della carriera nel caso in cui questa fosse inferiore a dieci anni), rispondendo a domande a risposta chiusa, con la possibilità di motivare o integrare le proprie risposte in modo libero al termine della compilazione di ogni sezione del questionario.

In particolare, la rilevazione delle percezioni sui fattori motivazionali relativi alle due dimensioni di studio è stata operativizzata attraverso la proposta di un set rispettivamente di nove e otto motivazioni alternative tratte dalla letteratura sul tema; è stato richiesto agli intervistati di esprimersi obbligatoriamente su *tutte* le motivazioni del set scegliendo un’opzione di una scala Likert a quattro posizioni, due negative (“per niente importante”, “poco importante”), due positive (“abbastanza importante”, “molto importante”). L’omissione dell’opzione “neutrale” è stata adottata al fine di favorire la polarizzazione delle risposte. L’obbligatorietà del posizionamento su tutti i fattori motivanti si è rivelata utile ai fini dei nostri obiettivi di ricerca, per garantire che alla valutazione fosse riservato un qualunque giudizio, negativo o positivo, consentendo di catturare l’informazione anche nel caso questa fosse percepita come un fattore incentivante non diretto (ma connesso ad altri fattori ritenuti primari), e di poter esaminare il suo contributo in rapporto a tutte le motivazioni proposte nel set.

3.2. Modalità d’analisi

Le analisi saranno prevalentemente di tipo descrittivo e forniranno prima un quadro generale su comportamenti e tendenze relativi a collaborazioni e mobilità nel contesto accademico italiano e poi indagheranno specificamente sui fattori motivazionali delle dimensioni sotto esame. Per offrire elementi di maggiore comprensione, nella presentazione dei risultati saranno riportati alcuni commenti liberi degli intervistati.

Le elaborazioni terranno conto dell’appartenenza degli accademici a quattro macroaree, che raggruppano diverse aree disciplinari CUN, vale a dire:

- *Matematica e Scienze Naturali* (MSN), categoria che comprende Matematica e Informatica; Fisica; Chimica e Scienze della Terra;
- *Scienze della Vita* (SV), che include Biologia; Scienze mediche; Agraria e Veterinaria;
- *Ingegneria e Architettura* (IA), in cui sono comprese Ingegneria civile e architettura; Ingegneria industriale e dell'informazione;
- *Scienze Sociali e Umane* (SSU), tra cui si includono Scienze dell'antichità, Filologia, Studi letterari, Storia dell'arte; Storia, filosofia, pedagogia e psicologia; Scienze giuridiche; Economia e Statistica; Scienze politiche e sociali.

Le risposte al questionario saranno analizzate rispetto ad alcune variabili strutturali, tra cui genere degli intervistati, livello di carriera (dividendo tra ricercatori a tempo determinato di tipo A e B, ricercatori a esaurimento, professori associati e professori ordinari), area geografica dell'istituzione di appartenza (Nord, Centro, Sud e Isole) e dimensione dell'ateneo come rapportata all'area CUN cui l'accademico afferisce (grande, media, piccola).

Al fine di tracciare profili di rispondenti in relazione all'influenza della valutazione nazionale sulle decisioni di collaborare e intraprendere mobilità internazionale, sarà utilizzata l'Analisi delle Corrispondenze Multiple (ACM). Questa tecnica di riduzione della complessità consente di proiettare le modalità di variabili multiple (nel nostro caso le caratteristiche individuali degli intervistati) su un numero inferiore di dimensioni o fattori, che esprimono le combinazioni tra le stesse variabili. Determinato il numero delle dimensioni di sintesi, esse sono ottenute attraverso un processo iterativo che sia in grado di riprodurre il massimo della varianza (o inerzia) totale e sia in grado di assicurarne l'ortogonalità, ossia l'indipendenza, tra loro. Attraverso un esame delle misure di discriminazione delle diverse variabili rispetto ad ognuna delle dimensioni ottenute, è possibile stabilire quali tra esse abbiano una capacità discriminante nell'interpretazione delle nuove dimensioni. A questo punto, l'analisi delle coordinate delle modalità proiettate sui piani factoriali consente di valutare quali di esse condividono caratteristiche di similarità rispetto alle variabili discriminanti che le connotano: modalità posizionate a poca distanza sul piano rispetto a una dimensione, e lontano dall'asse centrale, riveleranno un'associazione tra loro più forte (Di Franco, 2011).

4. I risultati della survey

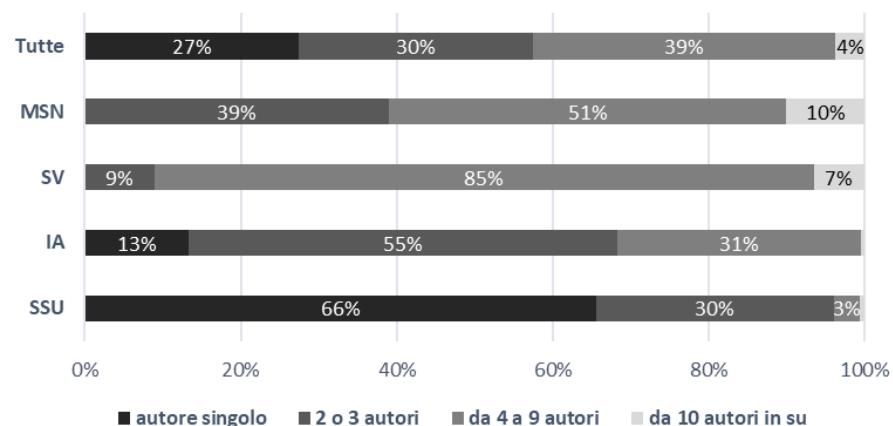
Questa sezione presenterà i risultati dell'indagine relativi ai due temi di studio trattati dal Capitolo. Si inizierà con l'analisi dei risultati riguardanti le collaborazioni di ricerca, seguita dalla presentazione di quelli legati alla partecipazione a iniziative di mobilità internazionale. Per una più appro-

fondita analisi della distribuzione dei 1.365 rispondenti rispetto alle variabili strutturali che orienteranno l'interpretazione dei risultati (vedi sezione 3) si rimanda a Spinello *et al.* (2021).

4.1. Collaborazioni di ricerca

La Fig. 1 offre un quadro della distribuzione percentuale degli intervistati in relazione al numero di autori coinvolti nella maggior parte delle loro pubblicazioni scientifiche, come indicato nelle risposte a una specifica domanda del questionario (cfr. Spinello *et al.*, 2021). Complessivamente, emerge un'importanza significativa delle collaborazioni, con il 30% degli intervistati che abitualmente pubblica i propri lavori a 2 o 3 firme, e il 39% i cui lavori coinvolgono da 4 a 9 autori. Poco più di un quarto degli intervistati (27%) ha invece dichiarato di condurre abitualmente i propri lavori in totale autonomia; tuttavia, va notato che questo dato è totalmente attribuibile alle macroaree delle SSU e di IA. Le pratiche di collaborazione si caratterizzano diversamente tra le diverse macroaree disciplinari.

Fig. 1 - Numero tipico di autori in pubblicazioni scientifiche per macroarea disciplinare ($n=1.365$)



Come anticipato, il contesto delle SSU attribuisce maggior valore all'individualità nella ricerca. La maggior parte degli intervistati (66%) dichiara di firmare solitamente i propri lavori da solo, mentre i lavori con 2 o 3 autori rappresentano la pratica usuale per il 30% degli intervistati in questa categoria. Le pratiche co-autoriali, soprattutto quando coinvolgono

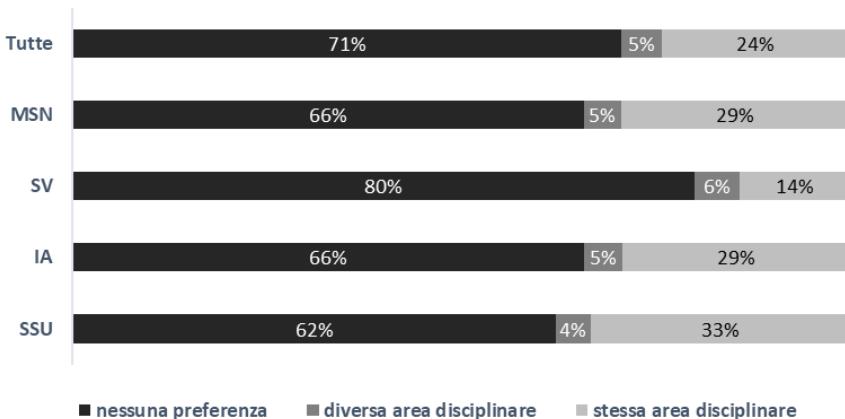
numerosi autori, a volte sono viste con sospetto da alcuni accademici della stessa macroarea, specie in vista della valutazione ASN:

La collaborazione è il sale dell'attività di ricerca per cui è auspicabile che essa venga valorizzata. Nel mio campo disciplinare però, dovrebbe essere individuato un numero massimo (astratto ovviamente) di autori di un lavoro per evitare l'effetto *free rider* ovvero che all'interno del gruppo si nasconde qualche soggetto che in realtà ha dato un contributo minimo se non inesistente. Questo numero, secondo me, equivale a tre. Oltre tre diventa soltanto una strategia di massimo risultato con minimo sforzo per avere le paventate fasce A o soglie per l'ASN (Professore associato, piccola università del Centro, macroarea SSU).

Per quanto riguarda il settore IA, si può parlare di pratiche di ricerca che tendono a un compromesso tra individualità e collaborazione su larga scala: il 13% degli intervistati ha dichiarato di non coinvolgere abitualmente co-autori e il 31% pubblica lavori firmati da 4 a 9 autori. Le SV si caratterizzano invece per un altissimo grado di collaborazione, con l'85% degli intervistati che solitamente pubblica lavori firmati da 4 a 9 autori. Nelle MSN, si osserva infine una varietà di approcci, con una distribuzione più bilanciata nel numero di collaborazioni tipiche.

Gli accademici che hanno dichiarato di svolgere lavori di ricerca solitamente con almeno un altro co-autore sono stati sondati riguardo alla preferenza rispetto alla collaborazione con colleghi afferenti alla stessa area disciplinare CUN o ad ambiti disciplinari differenti (Fig. 2). La maggioranza

Fig. 2 - Preferenze su collaborazioni con colleghi della stessa o di diversa area disciplinare per macroarea di ricerca (n=991)



degli intervistati (71%) non ha mostrato una preferenza specifica; coloro che hanno manifestato una chiara preferenza per collaborazioni all'interno della propria area disciplinare sono stati poco meno di un quarto (24%), mentre solo il 5% ha dichiarato un'esplicita predilezione verso la sinergia con colleghi di discipline afferenti ad altre aree. I dati indicano una generale flessibilità verso collaborazioni interdisciplinari, con gradi di apertura diversi a seconda della macroarea di appartenenza.

Nel campo delle SV, emerge una più forte inclinazione per la collaborazione che travalica i confini dell'area disciplinare; seguono IA e MSN, settore quest'ultimo in cui un rappresentante ha commentato, rapportando la sua preferenza alle indicazioni della valutazione, che:

L'interdisciplinarità è un parametro imprescindibile in molti settori di ricerca, ma non paga molto in termini di VQR (Professoressa ordinaria, media università del Sud, macroarea MSN).

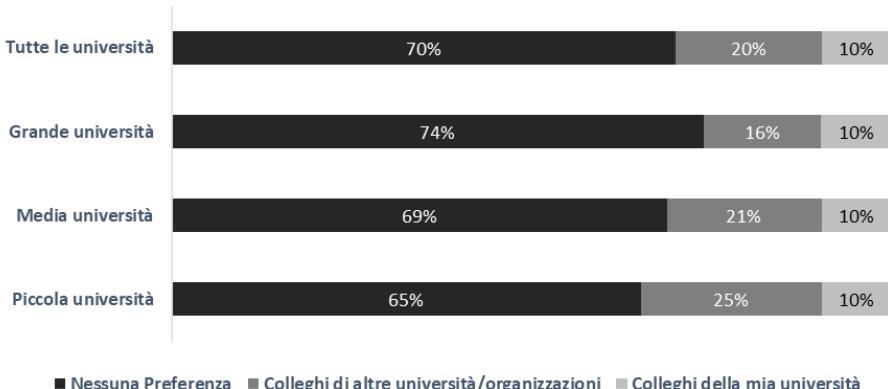
Per quanto riguarda le SSU, un terzo degli intervistati afferma di pre-diligere collaborazioni nell'ambito della medesima area di ricerca, nonostante talvolta vi siano sollecitazioni in senso opposto da parte delle proprie strutture di afferenza:

[...] Il mio Ateneo spinge alla collaborazione interdipartimentale, solo perché guarda all'esperienza delle scienze dure, ritenendola esportabile ovunque (Professore ordinario, media del Nord-Est, macroarea SSU).

La maggioranza degli intervistati (70%) non manifesta una preferenza specifica verso la collaborazione intra o interistituzionale (Fig. 3). Gli accademici afferenti ad atenei di dimensioni piccole o medie hanno dichiarato proporzionalmente una più accentuata preferenza alle connessioni con colleghi di altre università o organizzazioni (21% e 25%), rispetto a quelli afferenti a grandi università (16%). Dalle dichiarazioni libere di alcuni intervistati sembra emergere che, nella scelta di intraprendere collaborazioni interistituzionali, le precedenti esperienze in altri atenei, in Italia o all'estero, giochino un ruolo rilevante:

Ho lavorato per anni all'estero e ho costruito in tal modo collaborazioni durature di ricerca (Ricercatrice a tempo determinato di tipo B, grande università del Nord-Ovest, macroarea SSU).

Fig. 3 - Preferenze su collaborazioni intra- o inter-istituzione per dimensione università (n=991)



All'intero campione di rispondenti è stato chiesto con quale forza diversi possibili fattori, tra cui la valutazione nazionale, abbiano avuto un ruolo nelle loro decisioni di collaborare. I dati riportati in Tab. 1 indicano chiaramente che gli accademici, in un numero rilevante di casi, indicano un insieme di motivazioni concorrenti piuttosto che isolate motivazioni dominanti che li guidano nella scelta di collaborare. Per quasi il 60% degli intervistati i fattori "molto" motivanti sono almeno due, fino a un massimo di cinque; se allarghiamo lo sguardo anche a quelli "abbastanza rilevanti" addirittura il 23% ne elenca almeno sei. La valutazione si rivela rarissimamente una motivazione esclusiva (cioè, a cui gli intervistati attribuiscono un grado massimo esclusivo di importanza), e, quando indicata, entra in un set di motivazioni molto più ampio (il 59% di coloro che le attribuiscono un ruolo "molto importante" e il 92% di quelli che le assegnano un peso almeno "abbastanza importante" citano almeno altri 4 fattori incentivanti). Sulla coesistenza di diverse motivazioni e sul fatto che la valutazione possa difficilmente rappresentare una motivazione esclusiva, un intervistato si esprime così:

[...] le necessità legate ad indicatori bibliometrici e valutazioni VQR/ASN sono spesso implicitamente perseguite nelle altre scelte [...]. Alcune collaborazioni sono finalizzate alla qualità, che di per sé ha effetti positivi anche su indicatori e valutazioni, altre collaborazioni seguono invece una strategia esplicitamente orientata a necessità legate ad indicatori e valutazioni. Spesso entrambe le strategie coesistono [...] (Ricercatore a tempo determinato di tipo B, media università del Nord-Ovest, macroarea IA, per il quale la valutazione ha un ruolo "molto" importante sulle decisioni di collaborare).

Tab. 1 - Numero di motivazioni rilevanti espresse per la scelta di collaborare

Numero di motivazioni segnalate con la medesima intensità	Posizionamento su MOLTO sulla scala Likert		Posizionamento su ABBASTANZA o MOLTO sulla scala Likert	
	Tutti i rispondenti	Rispondenti che hanno indicato "Molto" su fattore VQR/ASN	Tutti i rispondenti	Rispondenti che hanno indicato "Abbastanza" o "Molto" su fattore VQR/ASN
Solo una	37% (355)	5% (6)	6% (74)	0% (0)
Due o tre	44% (420)	36% (43)	34% (449)	7% (31)
Quattro o cinque	15% (145)	34% (40)	37% (499)	31% (134)
Almeno sei	4% (36)	25% (30)	23% (310)	61% (262)

Se la valutazione rappresenta uno dei possibili fattori concorrenti per la scelta di collaborare, vediamo quanto è forte, per importanza attribuita, rispetto ad altre motivazioni. I risultati in Fig. 4 mostrano chiaramente che la necessità di seguire le indicazioni della ASN o della VQR assume, in senso generale, un ruolo non particolarmente rilevante rispetto ad altre motivazioni. Solo una parte di accademici afferenti a SV (39%) e di IA (35%) sembrano attribuire un qualche peso alle indicazioni della valutazione nelle loro scelte di collaborazione, mentre gli intervistati dei campi delle SSU e, ancor più, delle MSN appaiono in larga maggioranza indifferenti.

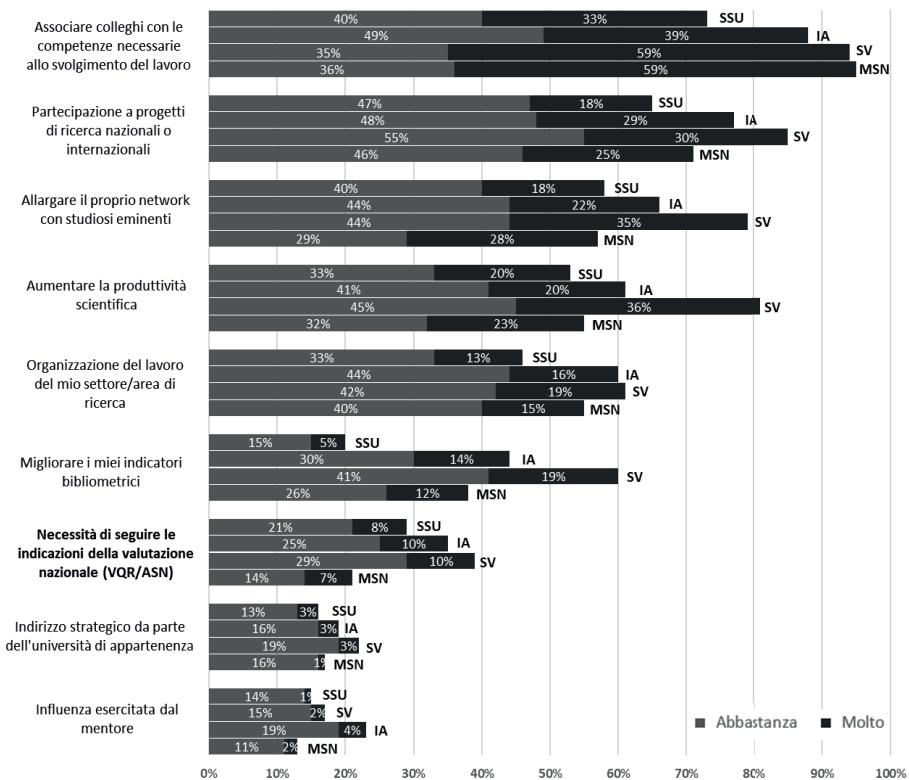
Il fattore maggiormente valorizzato in tutte le macroaree riguarda la possibilità di associare colleghi con le competenze necessarie allo svolgimento del lavoro di ricerca:

Scelgo i collaboratori sulla base della competenza e condivisione di interessi scientifici (Professore ordinario, grande università del Centro, macroarea SSU, che attribuisce un peso "per niente" importante alla valutazione).

Cerco sempre collaborazioni con colleghi italiani o dell'estero che stiano facendo ricerche che possono migliorare le nostre (del gruppo di ricerca che dirigo), incoraggiando incontri, pubblicazioni comuni ecc. (Professoressa associata, grande università del Centro, macroarea SSU, che attribuisce un peso "poco" importante alla valutazione).

Chi ha dato rilievo a questo fattore ha anche sottolineato l'importanza del fattore umano e sociale:

Fig. 4 - Motivazioni percepite per la scelta di collaborare per macroarea disciplinare (n=1.365). Nota. Per ogni macroarea disciplinare è indicata la percentuale di rispondenti che ha indicato la motivazione come "abbastanza" o "molto" importante



Lavorare con colleghi con i quali è piacevole lavorare e che possano migliorare le mie conoscenze (Professore associato, piccola università del Nord-Est, settore MSN, che attribuisce un peso “poco” importante alla valutazione).

sottolineando il:

[...] piacere nel lavoro di squadra, nel confronto (Professoressa associata, media università del Nord-Ovest, macroarea MSN, che attribuisce un peso “per niente” importante alla valutazione).

L'interesse scientifico si pone alla base delle scelte di almeno tre quarti degli intervistati, e alcuni non mancano di rilevare alcuni appunti critici verso la possibilità che la valutazione possa incidere sulle decisioni di collaborare:

Personalmente scelgo con chi collaborare in base alle sue competenze scientifiche e al rispetto reciproco, non certo per le assurdità legate all'ASN o alla VQR [...] (Ricercatore a esaurimento, piccola università del Sud, macroarea MSN, che attribuisce un peso “per niente” importante alla valutazione).

talvolta evidenziando come invece possa incidere negativamente, come quando si cerca la:

[...] possibilità di intrecciare approcci innovativi, più ampi e trasversali sulle tematiche di ricerca, anche se questo è quasi contro la logica (ristretta) della valutazione ASN (Ricercatrice a esaurimento, grande università del Nord-Est, macroarea SSU, che attribuisce un peso “poco” importante alla valutazione).

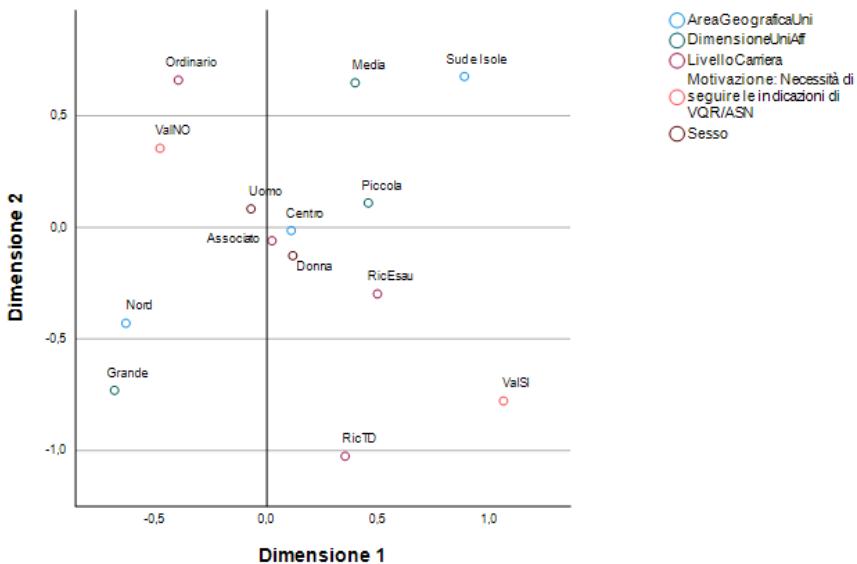
Un altro fattore particolarmente rilevante, che contribuisce ad alimentare le connessioni tra studiosi, riguarda la partecipazione a progetti di ricerca nazionali o internazionali, mentre un certo interesse di natura “remunerativa” emerge nelle attribuzioni della maggior parte degli intervistati, soprattutto quando si considera l’espansione della propria rete verso studiosi di alto livello, dal momento che essere in relazione con questi può tradursi in una sorta di garanzia per l’accesso a finanziamenti o rispetto alla possibilità di pubblicare su riviste di prestigio. Ciò naturalmente può avere un ritorno, in seconda battuta, anche in termini di valutazione, tanto che, a questo proposito, un intervistato afferma che:

È necessario adattarsi a lavorare con i grandi gruppi di ricerca consolidati, riconosciuti ed influenti, su tematiche e con metodologie *mainstream*, altrimenti è molto difficile ottenere finanziamenti e ancor di più pubblicare sulle riviste ritenute di pregio e quindi “remunerative” dal punto di vista della valutazione della ricerca (indipendentemente dagli effettivi contenuti: si valuta il contenitore per valutare il contenuto). Inoltre, se non pubblichi con “nomi” riconosciuti nella comunità, difficilmente accedi alle riviste top: quindi sei disposto ad “ospitare”, tra gli autori, studiosi considerati *mainstream*, anche se contribuiscono limitatamente alla ricerca/pubblicazione, perché sono la “chiave di ingresso” per le riviste top (Professore Ordinario, piccola università del Nord-Est, macroarea SSU, VQR/ ASN, che attribuisce un peso “molto” importante alla valutazione).

L’importanza attribuita ad effetti indiretti della valutazione, quali “migliorare i miei indicatori bibliometrici” appare abbastanza limitata, essendo valorizzata in parte da un sottogruppo di rappresentanti della macroarea SV. L’organizzazione della propria area di ricerca sembra assumere un valore solo per i settori di SV e IA. L’indirizzo strategico dell’università di appartenenza contribuisce alle decisioni per meno di un quarto del campione; ancor minore rilevanza è data all’influenza del mentore.

L'applicazione dell'ACM ha permesso di discriminare il profilo di coloro che hanno attribuito importanza alla valutazione come stimolo per la collaborazione da coloro che, al contrario, non la considerano un fattore di spinta. Attraverso questa tecnica, per ciascuna variabile considerata, viene calcolata una misura di discriminazione che ne quantifica la varianza all'interno dei fattori o dimensioni di sintesi. Questa misura varia tra 0 e 1: valori elevati di discriminazione evidenziano una significativa dispersione tra le categorie della variabile, indicando così un forte grado di discriminazione lungo la dimensione considerata. La prima dimensione sintetica (rappresentata dall'asse orizzontale in Fig. 5), che cattura poco più di un terzo della varianza (o inerzia) totale, si dimostra particolarmente efficace nel discriminare in base ai valori assunti dalla variabile "Necessità di seguire le indicazioni della VQR/ASN" (misura di discriminazione pari a 0,511). Il grado di prossimità spaziale delle modalità delle variabili relative al genere, al livello di carriera, all'area geografica e alla dimensione dell'università di appartenenza e all'influenza percepita di VQR/ASN nelle decisioni di collaborazione. Nota: l'inerzia totale spiegata dai due fattori è pari a 0,618. Il primo fattore discrimina in base all'influenza della valutazione nella scelta di collaborare (misura di discriminazione: 0,511); il secondo fattore rispetto alla dimensione dell'istituzione di appartenenza (misura di discriminazione: 0,366)

Fig. 5 - Proiezione su due fattori, ottenuta tramite ACM, delle modalità relative al genere, al livello di carriera, all'area geografica e alla dimensione dell'università di appartenenza e all'influenza percepita di VQR/ASN nelle decisioni di collaborazione. Nota: l'inerzia totale spiegata dai due fattori è pari a 0,618. Il primo fattore discrimina in base all'influenza della valutazione nella scelta di collaborare (misura di discriminazione: 0,511); il secondo fattore rispetto alla dimensione dell'istituzione di appartenenza (misura di discriminazione: 0,366)



risultano maggiormente influenzati dalla valutazione (31% del campione) è molto associato con l'essere un accademico di una università localizzata nel Sud, e lievemente associato con l'essere afferente a un'università piccola o media e a trovarsi ad un livello di carriera più basso rispetto a professore associato. Coloro che attribuiscono nessuno o scarso valore alla valutazione nella promozione delle loro collaborazioni sono accademici al livello di carriera più avanzato, afferenti alle grandi università del Nord. Il genere non sembra influenzare in modo significativo l'orientamento positivo o negativo nei confronti della valutazione nel promuovere le collaborazioni.

4.2. Mobilità internazionale

La Tab. 2 fornisce un quadro sulla percezione degli accademici rispetto al grado di supporto offerto dalle università italiane per le iniziative di mobilità internazionale e scambio di ricercatori. La maggioranza degli intervistati percepisce un sostegno abbastanza incisivo verso queste iniziative (53% buono, 20% forte). Tuttavia, si possono riconoscere differenze in base alla dimensione e all'area geografica dell'università di appartenenza. Le grandi università del Nord e le medie del Centro sono percepite come quelle che forniscono il miglior livello di supporto, mentre sono soprattutto

Tab. 2 - Percezione sul supporto dell'università di afferenza verso la mobilità internazionale (n=1.365)

Dimensione università di afferenza	Area geografica	Supporto nullo o scarso	Buon livello di supporto	Alto livello di supporto
Tutte	Tutte	27%	53%	20%
Grande	Nord	25%	51%	24%
	Centro	31%	54%	15%
	Sud e Isole	38%	46%	16%
Media	Nord	33%	49%	18%
	Centro	23%	49%	29%
	Sud e Isole	25%	57%	19%
Piccola	Nord	27%	55%	18%
	Centro	24%	58%	18%
	Sud e Isole	30%	50%	20%

Note: supporto nullo o scarso = l'università sostiene per niente o poco; buon livello = l'università sostiene "abbastanza"; alto livello di supporto = l'università sostiene "molto".

i rappresentanti delle università del Sud, soprattutto quelli delle grandi e delle piccole, a lamentare uno scarso sostegno. A questo proposito, un'intervistata osserva in particolare che:

La mia università se possibile ostacola, anziché aiutare. Soggiorni all'estero quasi sempre a mie spese per ricerca, seminari e congressi (Professoressa associata, media università del Sud, macroarea SSU).

La Tab. 3 offre una panoramica relativa ai periodi di soggiorno all'estero per ricerca superiori a un mese effettuati negli ultimi 10 anni dagli accademici. Dai dati emerge che la maggioranza degli intervistati (56%) non ha trascorso alcun periodo di mobilità nel periodo considerato. Si riscontra una differenza di genere abbastanza significativa che si manifesta a prescindere dalla dimensione dell'università ed è particolarmente evidente nelle piccole università (62% di donne che non hanno intrapreso alcuna esperienza di mobilità contro il 50% di uomini). In genere, chi effettua esperienze di mobilità tende a ripeterle, soprattutto nelle grandi e medie università, nella maggioranza dei casi fino a tre volte. Tra i motivi della mancata mobilità, dai commenti dei rispondenti emergono il carico didattico e il carico familiare:

Il carico didattico impedisce la mobilità internazionale (Professore associato, piccola università del Sud, macroarea SSU).

La mia mobilità internazionale è limitata sia dagli obblighi didattici, che familiari (presenza in famiglia di due bambini sotto i dieci anni). Questo è un criterio che a mio parere andrebbe inserito nella valutazione (a prescindere dalla mobilità internazionale), dato che la produttività scientifica dei genitori è necessariamente

Tab. 3 - Periodi di soggiorno all'estero per ricerca superiori a un mese negli ultimi 10 anni (n=1.365)

Dimensione università di afferenza	Genere	Nessuna esperienza di mobilità	Solo 1	Fino a 3	Fino a 5	Oltre 5
Tutte	Tutti	56%	17%	15%	5%	7%
Grande	Donna	57%	14%	15%	5%	9%
	Uomo	54%	17%	18%	6%	5%
Media	Donna	58%	14%	18%	6%	5%
	Uomo	56%	18%	14%	3%	9%
Piccola	Donna	62%	16%	10%	5%	7%
	Uomo	50%	21%	15%	7%	7%

inferiore a quella dei colleghi senza figli (Ricercatrice a tempo determinato di tipo A di una piccola università del Nord-Est, macroarea SSU).

In alcuni casi, la valutazione stessa emerge come fattore limitante per la mobilità:

Da precario, in un settore prettamente interno, non avevo il tempo per soggiornare all'estero, dovendo impiegare tutto il mio tempo nella didattica e nel superamento delle soglie (Ricercatore a tempo determinato di tipo B, media università del Sud, macroarea SSU).

A chi ha segnalato almeno un'esperienza di mobilità all'estero è stato richiesto di indicare in che misura vari potenziali fattori, tra cui la valutazione nazionale (in questo caso con riferimento alla sola ASN), abbiano influenzato le decisioni di intraprendere questo tipo di esperienza (Tab. 4). Gli intervistati che hanno indicato una sola forte motivazione sono stati più numerosi rispetto a coloro che ne hanno citate due o tre. Tuttavia, complessivamente, più della metà del campione ha segnalato motivazioni multiple, sebbene in minor numero rispetto ai fattori di spinta valorizzati nel caso delle collaborazioni (vedi Tab. 1), anche considerando un grado di rilevanza inferiore (motivazione “abbastanza importante”). La valutazione ASN è stata segnalata raramente come motivazione dominante, più spesso come parte di quattro o cinque motivazioni principali, suggerendo che essa costituisce in qualche caso un fattore concorrente.

Tab. 4 - Numero di motivazioni rilevanti per la scelta di intraprendere mobilità internazionale

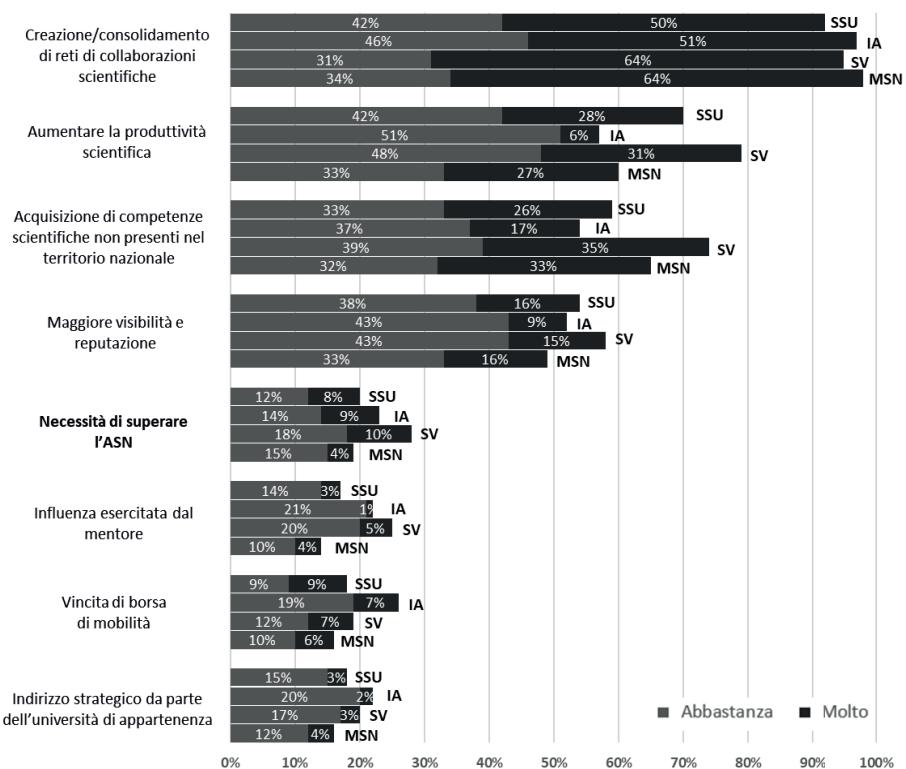
Numero di motivazioni segnalate con la medesima intensità	Posizionamento su “Molto” sulla scala Likert		Posizionamento su “Abbastanza” o “Molto” sulla scala Likert	
	Tutti i rispondenti	Rispondenti che hanno indicato “Molto” su fattore ASN	Tutti i rispondenti	Rispondenti che hanno indicato “Abbastanza” o “Molto” su fattore ASN
Solo una	44% (195)	17% (8)	6% (35)	1% (1)
Due o tre	42% (187)	17% (8)	41% (249)	6% (8)
Quattro o cinque	11% (50)	45% (21)	39% (233)	43% (57)
Almeno sei	3% (12)	21% (10)	14% (86)	50% (67)

I dati mostrano che il fattore nettamente più significativo, secondo i rappresentanti delle diverse macroaree disciplinari, è la “creazione o il consolidamento di reti di collaborazioni scientifiche”. Questo risulta particolarmente rilevante per le macroaree MSN e SV, con il 64% degli intervistati che lo giudica “molto” rilevante. Un’intervistata commenta che:

Grazie alle esperienze di mobilità ho potuto creare collaborazioni stabili che mi hanno permesso di crescere professionalmente e coltivare giovani talenti in un ambiente internazionale (Professoressa associata, grande università del Nord-Ovest, macroarea MSN, che attribuisce un peso “per niente” importante alla valutazione).

Segue l’“aumento della produttività scientifica”, un aspetto considerato di grande importanza soprattutto per le SSU (70%) e le SV (79%).

Fig. 6 - Motivazioni percepite per l’intrapresa di mobilità internazionale per macroarea disciplinare (n=607). Nota. Per ogni macroarea disciplinare è indicata la percentuale di rispondenti che ha indicato la motivazione come “abbastanza” o “molto” importante



L’“acquisizione di competenze non presenti sul territorio nazionale” si pone quale terza motivazione maggiormente valorizzata, risultando un fattore particolarmente rilevante nel contesto delle SV. Coloro che hanno segnalato questo elemento hanno contestualmente rilevato la possibilità di “utilizzare risorse analitiche complesse” o “utilizzare le risorse dei centri di ricerca” o “accesso a laboratori e strumentazione non presenti nella mia università e/o nel territorio nazionale”.

Come nel caso delle collaborazioni, il fattore valutazione – in questo caso segnalato dalla “necessità di superare l’ASN” – risulta nel complesso poco riconosciuto, con segnalazioni intorno al 20%, con una punta del 28% per il settore SV, concentrate maggiormente sull’opzione “abbastanza importante”. Questo dato si affianca ancora una volta alla scarsa importanza attribuita ad altre pressioni esterne come l’influenza del mentore e ancor più alla pressione dell’istituzione di appartenenza, sottolineando la limitata rilevanza assegnata agli aspetti istituzionali nel processo decisionale relativo alla mobilità internazionale. A questo proposito, un intervistato afferma che la mobilità è utile per:

[...] eseguire delle ricerche possibili solo presso istituzioni estere! (uno non deve necessariamente fare delle ricerche “in funzione di qualcos’altro” (ASN, VQR, pressioni universitarie), ma per scelte individuali autonome (Professore associato, media università del Sud, macroarea SSU, che attribuisce un peso “per niente” importante alla valutazione).

Come per l’analisi sulle collaborazioni, anche in questo caso col contributo dell’ACM (Fig. 7) è stato possibile tracciare un profilo dei pochi intervistati che hanno attribuito un certo peso alla necessità di conformarsi ai criteri ASN quando decidono di intraprendere periodi di ricerca all’estero (modalità: “abbastanza” e “molto” importante), rispetto a coloro che non ritengono affatto questo fattore rilevante (modalità: “per niente” e “poco” importante). Esaminando il primo fattore della Fig. 7, che raggruppa circa un terzo della varianza e riesce a discriminare i due gruppi soprattutto rispetto ai valori assunti dalla variabile “Necessità di superare l’ASN” (misura di discriminazione: 0,642), emerge che chi indica la valutazione come rilevante (modalità: ASNSi) non ha un profilo ben definito, anche se risulta essere lievemente associato ad un’università del Sud ad un livello non avanzato di carriera. Al contrario, coloro che non considerano rilevante la valutazione (modalità: ASNNo) sono i professori ordinari di grandi università del Nord.

Infine, la Fig. 8 restituisce un quadro della percezione degli intervistati riguardo agli effetti positivi percepiti sulle carriere scientifiche derivanti

Fig. 7 - Proiezione su due fattori, ottenuta tramite ACM, delle modalità relative al genere, al livello di carriera, all'area geografica e alla dimensione dell'università di appartenenza e all'influenza percepita di ASN nelle decisioni di mobilità. Nota: l'inerzia spiegata dai due fattori è pari a 0,649. Il primo fattore distingue in base all'influenza della valutazione nella scelta di collaborare (misura di discriminazione: 0,642); mentre il secondo rispetto alla dimensione dell'istituzione di appartenenza (misura di discriminazione: 0,505)

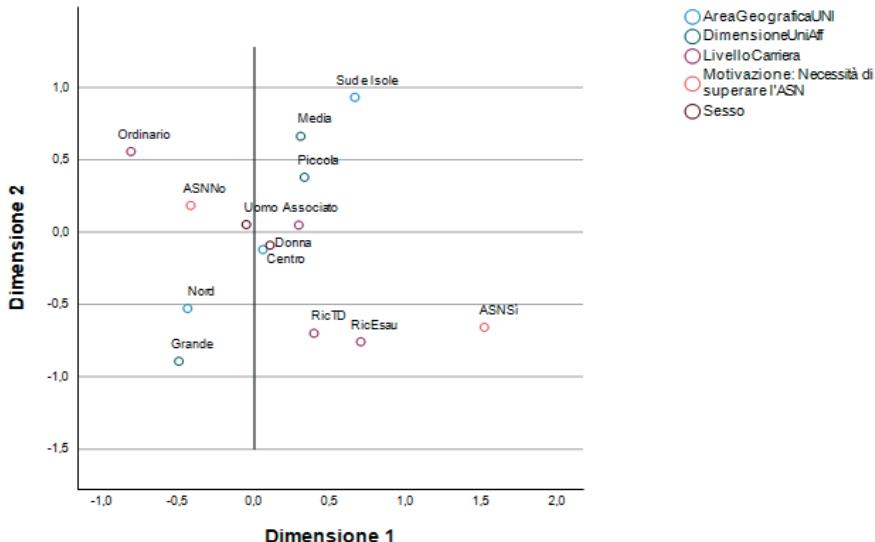
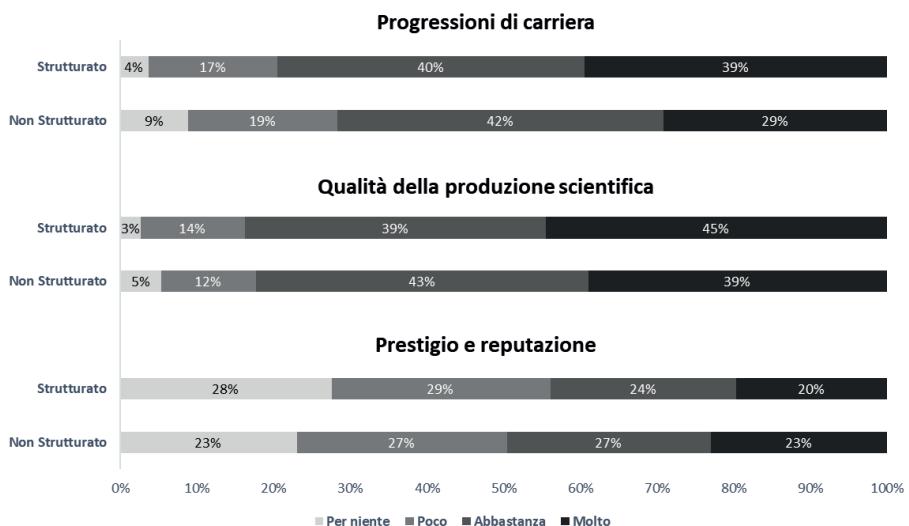


Fig. 8 - Effetti positivi percepiti a seguito delle esperienze di mobilità internazionale sulla carriera scientifica (n=607)



dalle esperienze di mobilità internazionale. Le percezioni sono analizzate distinguendo tra esperienze del personale strutturato, come ordinari, associati o ricercatori a esaurimento, e non strutturato, ossia ricercatori a tempo determinato di tipo A e B. La maggior parte degli intervistati ha riconosciuto un effetto positivo significativo sulla carriera scientifica, soprattutto in termini di qualità della produzione e progressione di ruolo. Questa percezione è più spiccata tra il personale strutturato che tra i non strutturati, con rispettivamente il 79% e l'84% di soddisfazione contro il 71% e l'82%.

Infine, emerge una divisione all'interno del campione riguardo agli effetti su prestigio e reputazione, con il 57% del personale strutturato che non associa questo effetto, tra tutti i possibili, all'esperienza di mobilità internazionale, mentre i non strutturati sembrano percepire un impatto leggermente più positivo in questo ambito.

5. Riflessioni conclusive

Attraverso l'analisi dei dati tratti dalla *survey* condotta nell'ambito del progetto PRIN 2017 “Gli effetti della valutazione sulla ricerca accademica: produzione della conoscenza e problemi metodologici”, questo contributo ha inteso approfondire la percezione degli accademici italiani riguardo a comportamenti, preferenze e aspetti motivazionali legati a due dimensioni dell'attività di ricerca che stanno assumendo un rilievo sempre maggiore nella scienza contemporanea: le collaborazioni tra studiosi e la mobilità internazionale. Posizionandosi nell'ambito degli studi che mirano a esplorare l'influenza delle politiche e dei criteri di valutazione della ricerca nella configurazione delle dimensioni del lavoro scientifico (cfr. de Rijcke *et al.*, 2016), il lavoro ha esplorato in particolare l'eventuale ruolo della valutazione nazionale tra i fattori che motivano o promuovono l'intrapresa di collaborazioni scientifiche e la partecipazione a iniziative di mobilità verso l'estero.

Relativamente al primo aspetto, le analisi presentate hanno rivelato che gli accademici adottano un approccio motivazionale complesso nella formulazione delle loro scelte di collaborazione. Questo risultato corrobora le evidenze presenti in letteratura, che indicano un'eterogenea gamma di possibili fattori che influenzano tali decisioni (cfr. Beaver e Rosen, 1978; Lee e Bozeman, 2007; Kyvik e Reymert, 2017; Abramo, D'Angelo e Murgia, 2017). Nello specifico contesto degli accademici italiani, emerge chiaramente che la valutazione nazionale della ricerca è un *driver* relativamente marginale nell'avvio di attività di ricerca congiunta. Solo una minoranza degli accademici, indipendentemente dalla loro area disciplinare di affe-

renza, ritiene importante l'adesione alle indicazioni fornite dall'ASN e dalla VQR come elemento da considerare nelle proprie decisioni di collaborazione, evidenziando come un incentivo in questa direzione non derivi in modo significativo dal riferimento a criteri o indirizzi delineati dalle politiche di ricerca. Gli accademici hanno manifestato una chiara preferenza per aspetti legati all'interesse scientifico e mirati all'implementazione di strategie personali, come la prospettiva di ampliare il proprio *network* coinvolgendo studiosi di alto livello. Quest'ultimo aspetto può essere associato a un più facile accesso a finanziamenti nell'ambito della presentazione di progetti e alla pubblicazione su riviste di prestigio, contribuendo in generale a un aumento della produttività. Questi effetti a cascata, influendo sugli indicatori di prestazione, possono produrre un impatto indiretto sulla valutazione, suggerendo una possibile influenza "di secondo livello" di ASN e VQR nelle strategie decisionali orientate alla collaborazione. Tuttavia, il ruolo della valutazione, anche come motivazione *concorrente*, emerge solo in modo limitato dall'analisi delle risposte del campione intervistato, rivelandosi solo quando i fattori di spinta sono riportati in numero superiore a tre. La modesta percezione dell'influenza della valutazione nazionale non solo pone interrogativi riguardo all'efficacia delle politiche valutative nel promuovere la collaborazione tra accademici, ma evidenzia come fattori di spinta derivanti da pressioni esterne esercitino una scarsa presa sulle attività di ricerca. A conferma di ciò, i dati indicano che anche l'eventuale pressione dell'ateneo di appartenenza e l'influenza del mentore assumono uno scarso peso motivante nella percezione degli intervistati.

Riflessioni analoghe possono essere estese relativamente al tema della mobilità internazionale. L'analisi ha infatti mostrato che la valutazione nazionale viene raramente annoverata dagli accademici tra le motivazioni per la decisione di intraprendere mobilità verso l'estero, nonostante gli stessi accademici riconoscano il rapporto positivo tra essa e prospettive di progressione di carriera. Ciò che è emerso invece come particolarmente motivante, in linea con quanto riportato in letteratura (cfr. Jonkers e Cruz-Castro, 2013; Franzoni *et al.*, 2014; European Commission, 2017), riguarda la possibilità di creare o consolidare reti di collaborazioni scientifiche, insieme ad altri aspetti quali la prospettiva di aumento della produttività scientifica e l'acquisizione di competenze non presenti sul territorio nazionale. Una volta di più, esplorando le percezioni soggettive degli accademici, è emerso come fattori esogeni abbiano un impatto limitato sul processo decisionale e, in particolare, l'azione della valutazione nazionale sembra non tradursi in effetti significativi di incentivo verso le iniziative di mobilità. Al contrario, in alcuni casi essa è percepita addirittura come un ostacolo: può infatti richiedere tempo eccessivo per il raggiungimento delle

soglie bibliometriche, limitando le opportunità di svolgere esperienze di studio all'estero o di partecipare ai bandi relativi a queste iniziative.

Complessivamente, i risultati dello studio suggeriscono che, sia nell'attivazione di collaborazioni che nell'intrapresa di iniziative di mobilità, individualità e autonomia si rivelano elementi cruciali, tenuti in grande considerazione nella percezione degli accademici. Essi hanno manifestato un approccio mirato a ottenere condizioni ideali per svolgere le proprie ricerche e raggiungere obiettivi strategici personali, anziché orientato verso un allineamento a criteri considerati desiderabili dalle politiche di ricerca. Tuttavia, analizzando il profilo dei rispondenti che in numero limitato hanno indicato l'importanza del fattore “valutazione” nelle loro decisioni, è da segnalare come un qualche indirizzo a caratteristiche di desiderabilità possa influenzare accademici ai livelli meno avanzati di carriera, afferenti a piccoli atenei, principalmente localizzati nel meridione.

In conclusione, il contributo valorizza l'impiego di approcci di studio basati sull'esplorazione delle percezioni degli studiosi come base empirica per la valutazione degli effetti delle politiche della ricerca sul lavoro scientifico, e fornisce indicazioni utili all'orientamento delle stesse politiche. Per incoraggiare ulteriormente la partecipazione attiva in attività collaborative e la mobilità all'estero, considerata la scarsa capacità incentivante delle politiche valutative, si dovrebbe insistere su leve che valorizzino l'autonomia e stimolino l'interesse scientifico. Ciò potrebbe avvenire attraverso il rafforzamento di strumenti orientati a stabilire reti di lunga durata con colleghi di istituzioni internazionali, che promuovano anche l'abbattimento dei confini disciplinari. In questo senso, le politiche nazionali della ricerca potrebbero mettere a sistema specifici programmi di finanziamento a progetto, alternativi rispetto ai già esistenti bandi europei e a quelli di cooperazione internazionale, con regole di accesso che agevolino la partecipazione dei giovani ricercatori e tengano conto degli squilibri territoriali. Se invece si vuole immaginare una valutazione che incentivi verso collaborazioni e mobilità nella ricerca, occorrerebbe strutturare indicatori in grado di valorizzare elementi di collaborazione internazionale ed eventuali periodi di mobilità all'estero in relazione alla qualità degli *output* scientifici prodotti. Questo obiettivo in una valutazione di tipo istituzionale, come la VQR, può essere perseguito adottando un approccio non legato alla valutazione delle performance delle università, ma alla capacità trasformativa della ricerca.

5. Parliamo tutti la stessa lingua? Riflessioni sulla valutazione della ricerca e i suoi meccanismi

di *Eliana Alessandra Minelli, Maria Rucsandra Stan*

1. Una sfida ontologica

La valutazione della ricerca è un concetto sfidante dal punto di vista ontologico. Cosa valutiamo? Qualsiasi prassi valutativa è il riflesso di una visione della realtà e di una scelta epistemologica, da cui scaturisce la particolare forma di conoscenza scientifica. Fin dall'antichità la natura della conoscenza rappresenta un tema arduo: il tentativo di definirla è quasi aporetico, ma certamente oggi torna alla nostra attenzione di fronte a una complessità crescente della conoscenza, nella cui produzione sarà sempre più coinvolta anche l'intelligenza artificiale. Soprattutto, valutare la conoscenza e il modo in cui la si costruisce richiede innanzitutto di chiarire cos'è la conoscenza, oggetto dell'azione valutativa.

Ritorna di attualità il Teeteto di Platone, in cui Socrate è impegnato in un dialogo con il giovane Teeteto alla ricerca di un significato condiviso della conoscenza: «Orbene, quando di qualche cosa uno riesca a formarsi la opinione vera, ma senza ragione, l'anima sua, riguardo a codesta cosa, è nel vero, ma non la conosce, perché non ha conoscenza di una data cosa chi non è capace di darne e riceverne ragione; chi invece riesca ad annettervi e cogliervi la ragione, costui è capace di tutto codesto, ed è anche, rispetto al conoscere, nella condizione migliore»¹.

Il dialogo tra i due si rinnova quotidianamente nella nostra società, dove la multiformità della conoscenza impone di riconoscere «l'immenso territorio dei significati del sapere», non sempre chiaramente definiti (Palumbo, 2002). Interrogato da Socrate su cosa sia la conoscenza, Teeteto dapprima risponde che «la conoscenza è sensazione», che sottende una

1. Platone, *Opere complete*, vol. II, trad. it. Manara Valgimigli, Laterza, Bari 1998, cap. XXXIX.

nozione di verità che si può conoscere solo passivamente attraverso l'impressione prodotta sull'animo umano. Ma Socrate, insoddisfatto, con l'esercizio della maieutica sprona il giovane a superare l'approccio relativista e lo conduce ad affermare che «la conoscenza è opinione vera» e che «la conoscenza è opinione vera di cui si sa rendere ragione», sottolineandone invece la forma attiva e richiamando l'essere umano alla “responsabilità della costruzione del sapere”. L'esito del dialogo riportato da Platone è però aporetico, così come ancora oggi lo è il dibattito sul tema, perché il concetto stesso di verità richiede di essere rifondato in un mondo attualmente dominato dai ‘mi piace’ e dai *ranking*, per ogni ambito dell'esperienza umana, in cui la prima definizione di conoscenza offerta da Teeteto – “la conoscenza è sensazione” – si riafferma prepotentemente. La conoscenza-sensazione riceve validazione e si trasforma in verità attraverso il criterio del gradimento: una posizione oggi diffusa, che esalta il soggettivismo e riduce la valutazione a stime e giudizi personali transitori e non comparabili. Ma è anche una riaffermazione del principio sofista che l'uomo è la misura di tutte le cose, che Platone voleva invece confutare a favore di una concezione attiva della conoscenza, oltre le sensazioni, verso il riconoscimento della “verità”, prerogativa del filosofo.

Oggi quindi, come forse mai prima, la produzione della conoscenza ha bisogno di un'attività valutativa rigorosa per affermarsi come “opinione vera di cui si sa rendere ragione”.

La valutazione della ricerca scientifica, ovvero di quella particolare attività che associa una entità simbolica o “valore” alla produzione di conoscenza (Mari, 2003), che deve corrispondere ai criteri di verità accettati dalla comunità degli esperti di riferimento, rappresenta il tema centrale per lo sviluppo della conoscenza e per il benessere della società.

Mari (2003) osserva che la misurazione è un particolare tipo di valutazione, in cui un misurando riceve un numero che lo quantifica e lo qualifica, e in un contesto di relativismo estremo essa perderebbe il suo specifico valore epistemico. La Fisica in questo senso è il paradigma di tutte le scienze per la sua caratteristica di misurare obiettivamente quantità e quindi definire relazioni tra esse. Ma questo è un privilegio di cui godono solo alcuni domini della conoscenza, lasciando i ricercatori in altre aree nell'aporia di definire cosa e come valutare.

L'ambito delle scienze sociali rappresenta un territorio molto delicato dal punto di vista epistemologico, in cui l'attività di indagine, anche solo l'osservazione diretta condotta dai ricercatori, interferisce con il funzionamento del sistema, caratterizzato da una combinazione dinamica di comportamenti individuali e collettivi (Mayo, 1933). È noto, infatti, come nelle organizzazioni oggetto di osservazione, l'applicazione di teorie manageriali generi profezie

che si autoavverano (Micheli e Mari, 2014) proprio perché tale applicazione influenza pensieri e atteggiamenti delle persone coinvolte, a differenza dei sistemi fisici che non sono sensibili alla disposizione dell'animo umano.

Valutare la ricerca scientifica, dunque, richiede di definire il “cosa” e il “come”, di riconoscere e rendere esplicita la posizione ontologica ed epistemologica in cui l’azione valutativa si radica e di esprimere un valore condivisibile e utile per orientare opinioni, sforzi e risorse.

Questo capitolo si propone di tracciare l’evoluzione del dibattito sul tema della valutazione tra diverse comunità scientifiche attraverso l’analisi delle pubblicazioni accademiche e di esplorare gli effetti delle prassi di valutazione attualmente applicate in Italia sul modo di generare nuova conoscenza, attraverso uno studio di come le pratiche di valutazione hanno influito sullo sviluppo di un settore scientifico-disciplinare non bibliometrico nel periodo 2000-2020. Ne emerge un quadro intricato, in un contesto caratterizzato da complessità crescente, che richiama la necessità di un approccio olistico e multidimensionale alla valutazione della qualità della ricerca scientifica.

2. Evoluzione del concetto valutativo e relative prassi

La valutazione della ricerca si è evoluta nel tempo, in relazione alla complessità crescente delle forme di ricerca scientifica e alla loro rilevanza sociale.

Nei secoli, la dimensione intrinseca della qualità è stata predominante nelle università, dove era valutata tradizionalmente attraverso l’autovalutazione (*self-evaluation*) e la valutazione tra pari (*peer evaluation*). Indicatori di caratteristiche intrinseche di qualità riguardano le componenti essenziali di un testo. Questi fattori, come la coerenza, la consequenzialità, la validità dei risultati sono parte integrante della costruzione e della valutazione della qualità. L’autovalutazione (Van Vught e Westerheijden, 1994, p. 366) è sempre stato un meccanismo cruciale che ha permesso agli accademici di accettare e affidarsi a un sistema di valutazione della qualità. Tuttavia, negli anni ’80 del XX secolo la qualità estrinseca è emersa come forma di politica pubblica e ha preso il sopravvento, segnando il passaggio a un concetto di valutazione della qualità legato alla responsabilità nell’uso delle risorse pubbliche assegnate alle università per adempiere la propria fondamentale missione di produrre conoscenza e tramandarla alle nuove generazioni. Le caratteristiche estrinseche sono definite come fattori esterni che non sono inerenti a «elementi della qualità di un testo, ma hanno una correlazione positiva con giudizi basati su tali elementi» (Bridges, 2011, p. 32) e sono legate alle ricadute sulla società.

Da quell'epoca, la crisi del *welfare state* ha indotto i governi a porre sotto controllo se non a ridurre la spesa per l'educazione terziaria allo scopo di contenere il carico fiscale che lo stato sociale, nei paesi europei, aveva fatto lievitare. La valutazione della qualità ha assunto quindi il significato di garanzia che le istituzioni universitarie, nonostante la riduzione dei fondi pro capite, continuassero a fornire una educazione di adeguato livello (Westerheijden, 2007, p. 75). Negli anni '80 dello scorso secolo, nel Regno Unito si affermò dunque il *New Public Management* come modello di gestione della cosa pubblica caratterizzato da deregolamentazione e allo stesso tempo da una necessaria responsabilizzazione (Westerheijden, 2007, p. 77). La relazione tra le istituzioni universitarie e i governi nazionali cambiò di fronte alla pressione dell'università di massa, passando da un modello caratterizzato dal controllo diretto sulle istituzioni alla supervisione indiretta (Rosa e Amaral, 2007, p. 183). Di conseguenza, in molti paesi la valutazione fu progressivamente introdotta per verificare il livello di qualità e divenne il perno del sistema di responsabilizzazione delle università. A questo processo è legata l'introduzione dell'Assicurazione della Qualità come meccanismo per controbilanciare la maggiore autonomia dell'università, rispetto al sistema precedente caratterizzato da omogeneità e controlli diretti dello stato sulle singole istituzioni (Rosa e Amaral, 2007, p. 183). La letteratura ha indicato questo cambiamento come la nascita dello «Stato Valutativo» (*Evaluative State*) (Neave, 1988; 1998; 2012).

La principale caratteristica dello «Stato Valutativo» è la razionalizzazione e la redistribuzione delle funzioni tra centro e periferia (Neave, 1988), che ha stabilito un nuovo modello operativo in cui le funzioni di definizione, implementazione, interpretazione e verifica non sono più concentrate a livello centrale di ministero, ma sono suddivise e assegnate a diverse agenzie e a livelli decisionali differenti nell'ambito del sistema universitario. Nel complesso, lo «Stato Valutativo» riflette il tentativo di superare i tradizionali modelli di valutazione e di stimolare risposte più rapide e precise da parte delle istituzioni universitarie, grazie a strumenti di valutazione più sofisticati e di maggiore respiro che in passato.

In questa direzione si inserisce anche il Processo di Bologna (1999) con la costruzione di uno spazio europeo della formazione superiore (*European Higher Education Area – EHEA*) che ha coinvolto le istituzioni universitarie dei paesi aderenti all'accordo (anche al di fuori del contesto europeo) per assicurare la qualità nella trasmissione della conoscenza nella consapevolezza che l'educazione è un fattore fondamentale per il miglioramento delle risorse cognitive (e dunque del capitale umano) e per assicurare la prosperità della società (Faber e Westerheijden, 2011).

Tra le missioni dell'università, sicuramente quella della ricerca ha un valore critico per lo sviluppo nazionale e regionale e dunque l'allineamento agli obiettivi delle politiche pubbliche rappresenta una finalità della valutazione della ricerca scientifica sempre più importante. Innanzitutto, la limitatezza delle risorse pubbliche a disposizione delle università impone una allocazione efficiente tra progetti, gruppi e iniziative alternative (Gläser e Whitley, 2007). In secondo luogo, il contesto competitivo internazionale focalizza l'attenzione sugli esiti dei processi di valutazione e sulle graduatorie internazionali (*ranking*) degli enti di ricerca soprattutto nei paesi più avanzati (Hazelkorn, 2015; Sayed, 2019). Infine, la crisi pandemica da Covid-19 ha acuito la sensibilità pubblica verso la ricerca scientifica, enfatizzando la necessità di documentare e comprovare gli effetti delle attività degli scienziati e dei ricercatori sugli individui e sulla società.

Alla base dei sistemi di valutazione della ricerca scientifica rimane la *peer review*, cionondimeno vi è una tendenza a usare metadati come complemento ai processi di revisione tra pari, i quali di fatto si stanno affermando come indicatori di qualità, come emerge nelle varie applicazioni nell'ambito del sistema della ricerca accademica. Dunque, un progressivo spostamento dalla tradizionale *peer review*, come «gold standard» della valutazione della ricerca, all'affermazione della superiorità delle misure bibliometriche (Abramo e D'Angelo, 2011). Gli indicatori bibliometrici sono comunemente usati come approssimazioni affidabili della qualità della ricerca e della produttività (Margherita *et al.*, 2022) e assumono un ruolo significativo nell'orientare le decisioni a livello politico e accademico, in particolare rispetto all'assegnazione delle risorse e alle promozioni. La letteratura ha comunque messo in luce che tali indicatori rappresentano in realtà una misura complementare che mitiga i limiti della valutazione tra pari (Aksnes, 2005; Higgins *et al.*, 1990) e contribuisce al processo complessivo di valutazione ma non ne sono una componente esclusiva. La legittimità e l'affidabilità degli indicatori bibliometrici sono comunque oggetto di dibattito (Bornmann e Daniel, 2007; Osterloh e Frey, 2015; Waltman, 2016; Waltman *et al.*, 2011; Wouters, 1999), al punto da aver indotto a dubitare della loro capacità di misurare la qualità scientifica (Cronin, 1984), nonché le differenze a livello disciplinare e sociale (Seglen, 1997). Sono state osservati effetti perversi sul sistema di produzione della conoscenza (Cagan, 2013; Hicks *et al.*, 2015; Weingart, 2005) ed effetti sulla distribuzione delle risorse che hanno ridotto l'attenzione ad altre dimensioni della qualità, come il valore sociale della ricerca (Aksnes *et al.*, 2019).

I meccanismi e le prassi di valutazione sono quindi oggetto di discussioni tra comunità accademiche di diversa tradizione che meritano una attenzione e una riflessione approfondita per comprendere meglio come

valutare la ricerca scientifica, a fronte della consapevolezza che la qualità della ricerca è un concetto multidimensionale che differisce in base all'ambito disciplinare e alla missione stessa della ricerca (Aksnes *et al.*, 2019). La comunità scientifica a livello internazionale sta affrontando la sfida di stabilire una definizione di valore universale della qualità della ricerca e di buone pratiche scientifiche, in un contesto in cui la società si attende una ricerca scientifica di eccellenza ma anche di elevato impatto sociale. Diverse iniziative transnazionali stanno nascendo per affrontare la questione: tra esse la *San Francisco Declaration on Research Assessment* (DORA) propone raccomandazioni che sostengono l'uso di indicatori qualitativi di impatto della ricerca, come gli effetti su politiche e pratiche. Analogamente, il Leiden Manifesto for Research Metrics propone alcuni principi guida per valutare la produttività della ricerca e propone approcci più sostenibili e integrali (Hicks *et al.*, 2015; Kanchan e Krishan, 2019; Margherita *et al.*, 2022). Recentemente, la *Coalition for Advancing Research Assessment* (CoARA) propone un insieme di principi per promuovere il cambiamento nelle pratiche di valutazione della ricerca, dei ricercatori e degli enti di ricerca «con l'obiettivo generale di massimizzare la qualità e l'impatto della ricerca»².

3. Il dibattito accademico³

Per avere un quadro più approfondito dell'evoluzione del dibattito sulla valutazione della ricerca, è stata condotta una revisione sistematica della letteratura prodotta da differenti comunità accademiche con lo scopo di tracciare l'evoluzione della percezione, dei metodi e dei meccanismi della valutazione della qualità della ricerca, nonché delle caratteristiche della qualità, in una prospettiva interdisciplinare.

L'approccio utilizzato, chiamato *Systematic Literature Network Analysis*, rafforza i metodi tradizionali di revisione, superando il rischio di distorsioni generate da percezioni individuali (Alerasoul *et al.*, 2021; Comerio e Strozzi, 2019) poiché i paper emergono attraverso l'analisi del *network* di pubblicazioni, fornendo una visione imparziale dello stato attuale della ricerca.

La ricerca è stata effettuata sul database Scopus, considerando le aree delle scienze sociali e del management in quanto più ricche di contributi e solo pubblicazioni in lingua inglese. Il database Scopus è stato preferito ri-

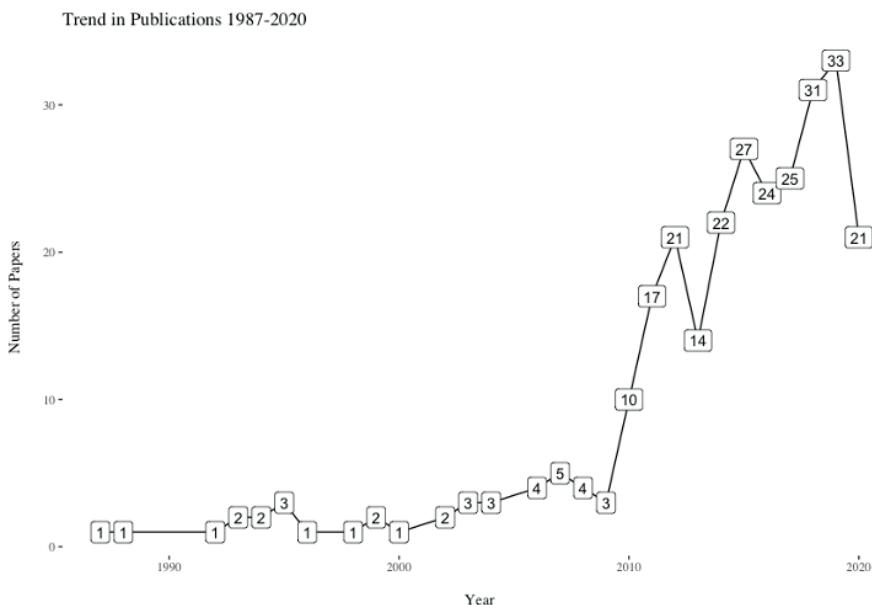
2. Coalition for Advancing Research Assessment (CoARA) Link: <https://coara.eu/agreement/the-agreement-full-text/> (Accesso: 02 agosto 2023).

3. Elaborazione da Stan, M.R. (2023). *Insights into evaluation of research quality: what role for the researcher?*. Tesi di dottorato. LIUC-Università Cattaneo.

spetto ad altre basi di dati, grazie alla sua maggiore estensione nelle aree di interesse e al *focus* scientifico (Balstad e Berg, 2020; Noruzi, 2005; Zhao e Strotmann, 2015).

La ricerca effettuata ha ottenuto 284 lavori pubblicati su riviste internazionali e *conference paper* dal 1987 al 2020 che sono stati analizzati per comprendere l'evoluzione del dibattito sulla valutazione della ricerca. La tematica ha cominciato a svilupparsi a partire dalla fine degli anni '80 dello scorso secolo ma è solo dal 2011 che l'attenzione è cresciuta significativamente, come indicato nella Fig. 1.

Fig. 1 - Sviluppo delle pubblicazioni nel periodo 1987-2020



Analizzando i paesi da cui provengono i contributi, si osserva che al vertice si trova l'Italia, seguita dal Regno Unito e dagli Stati Uniti, indizio di un contesto particolarmente sensibile al dibattito. Una analisi preliminare condotta sulla base del Global Citation Score (GCS), che identifica i lavori più influenti nell'arco temporale considerato, ha evidenziato un quadro delle criticità e delle sfide che riguardano gli attuali dei sistemi di valutazione e le implicazioni di una prevalenza degli indicatori quantitativi nella misurazione delle varie dimensioni della qualità della ricerca (Aksnes *et al.*, 2019; Costas *et al.*, 2015; Ellegaard e Wallin, 2015; Hassan *et al.*, 2017; Haunschild *et al.*, 2019). In particolare, sono evidenziati

i limiti associati all'uso degli indicatori citazionali come misura della qualità della ricerca e la necessità di applicare metodi diversificati, sia quantitativi che qualitativi, nonché l'opportunità di considerare il contesto della ricerca. Inoltre, sono studiate le relazioni tra *altmetrics* (metriche alternative rispetto ai tradizionali indici bibliometrici) e citazioni, e il differente impatto, accademico o sociale, che esse misurano. Altri studi sottolineano l'importanza di considerare il contesto più ampio in cui i lavori sono citati e la qualità della letteratura che riporta le citazioni (Bornmann *et al.*, 2018), ma anche fenomeni di inflazione delle citazioni (Petersen *et al.*, 2019). Infine, è importante considerare la relazione tra il modo di fare ricerca e i risultati scientifici e sociali attesi ai fini della valutazione dei progetti di ricerca e i benefici potenziali di approcci inter e transdisciplinari per affrontare tematiche di ricerca sociale sempre più complesse (Newig *et al.*, 2019).

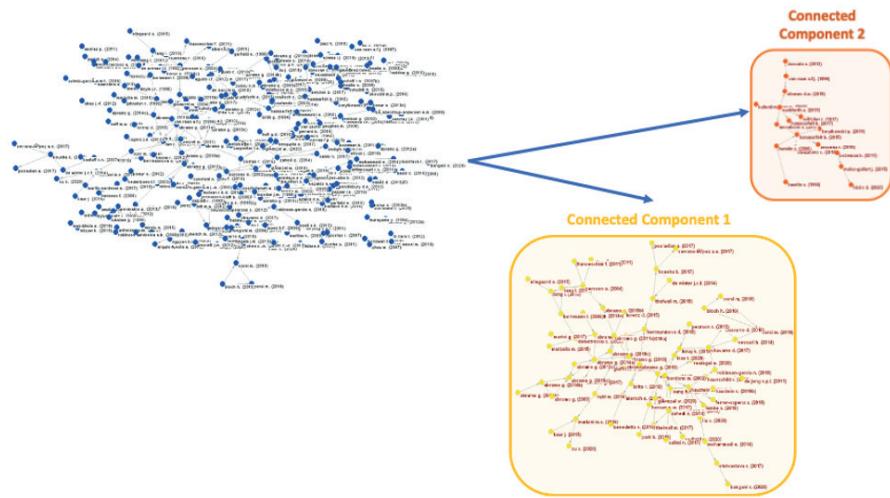
I risultati più interessanti, anticipati nei *seminal paper* (i.e., i lavori più influenti) individuati con il GCS, sono forniti attraverso la *Citation Network Analysis* (CNA), basata sulla rete di citazioni emergente nel campione di lavori scientifici selezionati nella banca dati Scopus. Lo scopo di questa analisi è l'identificazione di traiettorie di sviluppo all'interno della letteratura (Strozzi *et al.*, 2017) sulla valutazione della ricerca. Selezionando parole chiave specifiche⁴, sono emersi due componenti principali (*connected component*), che includono complessivamente 90 documenti, all'interno della più ampia rete citazionale, che comprende 284 nodi.

L'individuazione delle due componenti principali è stata realizzata con il software Pajek, mentre le traiettorie (*main path*) di ciascuna componente sono state evidenziate mediante *Vos Clustering method*, basato sul software VOSviewer (Van Eck *et al.*, 2010; Waltman *et al.*, 2010).

Le traiettorie (*main path*) sono quindi estratte da ciascuna componente per tracciare e comprendere l'evoluzione della tematica, sulla base del contenuto dei documenti considerati (Lucio-Arias e Leydesdorff, 2008; Niazi *et al.*, 2019).

4. La stringa di parole chiave finale include la seguente combinazione di termini: i) “accademico”, “scientifico” e “lavoro” che indicano il tipo di risultato della ricerca; e ii) “impatto” e “qualità”, che si riferisce a ciò che è misurato/valutato nel processo di valutazione della ricerca. La ricerca è stata limitata ad articoli e atti di convegni, poiché i primi rappresentano gli elementi costitutivi della letteratura e sono state considerate solo le aree di ricerca relative alle scienze sociali e alla gestione aziendale. Sono stati considerati tutti gli anni, tranne il 2021 per la sua scarsa rilevanza al momento dell'estrazione e presi in considerazione solo i documenti in lingua inglese. Il database Scopus è stato quindi selezionato in base alla seguente stringa: (“accademico” OR “scientifico” OR “lavoro”) AND {valutazione della ricerca} AND (“impatto” OR “qualità”).

Fig. 2 - Rete citazionale dei 284 nodi e i suoi due maggiori componenti connessi



Ciascuna componente principale contiene i contributi di comunità diverse ed è caratterizzata da temi specifici.

Da un lato, si individuano i lavori degli “informetrici” (che comprendono anche i bibliometrici e gli ‘scientometrici’), interessati alla misurazione delle caratteristiche estrinseche della qualità; dall’altro i contributi degli scienziati sociali che concentrano l’attenzione sugli effetti della valutazione.

3.1. Prima componente principale e suo sviluppo

La prima componente principale comprende sottogruppi di pubblicazioni che affrontano specifici temi raggruppabili in tre grandi aree tematiche.

Il primo tema riguarda gli sforzi per superare le distorsioni che gli indicatori bibliometrici inducono nella valutazione della ricerca e il tentativo di andare oltre i sistemi di valutazione basati sulle metriche. Gli indicatori bibliometrici sono comunemente usati nell’ambito della selezione, della promozione del personale accademico e delle decisioni di finanziamento di progetti e pertanto è fondamentale generare consapevolezza dei limiti e delle situazioni che ne influenzano la significatività e la capacità di fornire un riscontro affidabile (Abramo, D’Angelo e Di Costa, 2019a; Franceschini e Maisano, 2011; Persson *et al.*, 2004; Xu *et al.*, 2020). Inoltre, gli indicatori bibliometrici tradizionali non sembrano in grado di misurare in modo soddisfacente l’impatto sociale ed economico che la ricerca può avere (Abramo, 2018).

Le metriche citazionali tradizionali, come il conteggio delle citazioni e l'H-index, non catturano completamente la qualità o la pertinenza degli articoli che citano. Inoltre, sembrano insufficienti per valutare la produttività e l'impatto scientifico perché non considerano il contributo relativo del ricercatore o i modelli di collaborazione all'interno di una comunità scientifica (Persson *et al.*, 2004). Pertanto, sono necessari metodi alternativi (Albarán *et al.*, 2011). Alcuni studiosi (Ellegaard e Wallin, 2015; Hiruy *et al.*, 2019) incoraggiano l'uso della bibliometria insieme a metodi qualitativi per misurare con precisione l'impatto della ricerca. Molti autori evidenziano nei loro lavori (Ferrara *et al.*, 2015; Mariani *et al.*, 2015; Nykl *et al.*, 2014; Xu *et al.*, 2020) la sfida dello sviluppo di metriche e metodi di valutazione efficaci per cogliere l'impatto scientifico e la produttività. La costante ricerca di un nuovo metodo potrebbe portare a miglioramenti (Giuffrida *et al.*, 2019) o potrebbe anche tradursi in un errore di valutazione (Abramo e D'Angelo, 2016a) che diminuisce la capacità di catturare ricerche significative (Abramo, D'Angelo e Di Costa, 2019a). Per affrontare queste carenze, alcuni autori (Bornmann *et al.*, 2008) promuovono l'implementazione di buone prassi nella formulazione di nuovi indicatori bibliometrici e nella loro interpretazione.

Inoltre, alcuni autori si sono concentrati sulla complessità e sulle sfide della valutazione dell'impatto sociale e della qualità della ricerca e sulla necessità di un approccio più articolato e specifico rispetto al contesto, che consideri la diversità dei risultati della ricerca e il loro impatto su vari tipi di pubblico (de Jong *et al.*, 2011).

Altri studiosi hanno cercato di far luce sul fatto che l'eccessivo affidamento alle metriche (Tran *et al.*, 2020) può plasmare le priorità e le pratiche della ricerca, compreso l'attuale contesto competitivo nell'editoria scientifica che ha un impatto negativo sullo sviluppo scientifico in alcune aree geografiche, come l'America Latina (Vessuri *et al.*, 2014), e il perpetuarsi di pregiudizi sistematici e opportunità limitate di gruppi minoritari di ricercatori in Italia (Corsi *et al.*, 2018, 2019). Questi critici hanno evidenziato la necessità di fornire una visione più completa dell'impatto della ricerca che vada oltre le metriche tradizionali, come le citazioni o il Journal Impact Factor, che possono essere integrate con le *altmetrics* (Robinson-Garcia *et al.*, 2018).

La seconda area tematica di interesse riguarda proprio la rilevanza delle *altmetrics* nella misurazione dell'impatto sociale della ricerca. Negli ultimi anni, il crescente utilizzo dei social media tra i ricercatori ha dato origine alle *altmetrics* come metodo di valutazione della ricerca. Questi indicatori basati sul web, su piattaforme come Mendeley (Thelwall, 2017) o ResearchGate (Shrivastava e Mahajan, 2017), misurano l'impatto scientifico attraverso azioni come *download*, "Mi piace" e condivisioni di pubblicazioni scientifiche. A causa dell'eterogeneità di utilizzo tra i ricercatori

(Lemke *et al.*, 2018), le *altmetrics* hanno una elevata capacità di riflettere i diversi giudizi degli utenti, anche al di fuori della comunità scientifica, diventando un indicatore dell'impatto sociale (Haustein, 2016). Le *altmetrics* possono operare a diversi livelli, da quello istituzionale a quello del singolo articolo e distinguere tra la semplice condivisione di informazioni bibliografiche e la discussione sull'essenza di un manoscritto accademico (Haustein, 2016). Alcuni ricercatori hanno confrontato *altmetrics* e indicatori di citazione (Ferrer-Sapena *et al.*, 2018; Haunschmid e Bornmann, 2018; Zahedi *et al.*, 2014), sostenendo l'idea che «non riflettono lo stesso tipo di impatto» (Costas *et al.*, 2014). Le *altmetrics* possono rilevare il trasferimento di conoscenze tra discipline scientifiche (Mohammadi e Thelwall, 2014), fornire opportunità per la valutazione della ricerca nei paesi periferici (Shrivastava e Mahajan, 2017) e supportare processi di pianificazione, definizione del budget e allocazione delle risorse per biblioteche, università e organizzazioni di ricerca (Sedighi, 2020). Le *altmetrics* mirano, inoltre, a offrire una valutazione integrativa della ricerca nell'era digitale, integrando la tradizionale revisione tra pari e le misure bibliometriche per comprendere tutti gli aspetti di un risultato di ricerca nel tempo (Aung *et al.*, 2019; Bangani e Onyancha, 2021; Hassan *et al.*, 2017; Park e Park, 2018; Salimi, 2017). Tuttavia, le *altmetrics* presentano anche carenze e rischi (Regan e Henchion, 2019), che richiedono un'attenta interpretazione e applicazione nel contesto della valutazione della ricerca (Glänsel e Chi, 2020; Liu *et al.*, 2020).

Infine, la terza macroarea tematica individuata nella prima componente principale si concentra sul sistema di valutazione della produttività scientifica in Italia, sia a livello individuale che istituzionale, con specifica attenzione alle sfide poste dal modello nazionale di valutazione della ricerca e all'Abilitazione Scientifica Nazionale (ASN).

La misurazione della produttività della ricerca dovrebbe considerare sia aspetti quantitativi, sia qualitativi, insieme alle differenze tra i campi della ricerca (Abramo, Cicero e D'Angelo, 2011a; Abramo e D'Angelo, 2014a).

A livello istituzionale, la letteratura evidenzia una criticità del sistema di ricerca italiano nel fatto che l'output della ricerca è altamente concentrato in poche istituzioni, mentre la maggior parte delle istituzioni produce pochissima ricerca (Abramo, D'Angelo e Di Costa, 2009a; Abramo e D'Angelo, 2015b). A livello individuale le criticità individuate sono principalmente due. Da un lato, la produttività della ricerca ha una distribuzione ineguale tra discipline e campi di ricerca, dove pochi ricercatori producono la maggior parte degli articoli (Abramo, D'Angelo e Soldatenkova, 2017). D'altra parte, la tempistica dell'osservazione delle citazioni ha un impatto significativo sulle classifiche, con i ricercatori che hanno pubblicato di recente che tendono a posizionarsi meglio (Abramo, Cicero e D'Angelo,

2012a). L'Abilitazione Scientifica Nazionale, che valuta il merito scientifico dei candidati ai fini della loro promozione, si basa su indicatori bibliometrici che non sempre sono in grado di restituire una misurazione corretta della qualità della ricerca e dell'impatto che questa ha nel campo specifico (Demetrescu *et al.*, 2020a; Marzolla, 2015).

L'enfasi sugli indicatori bibliometrici nel sistema di ASN ha condotto a modelli di promozione nelle università italiane che favoriscono la produttività rispetto alla *seniority* accademica (Abramo e D'Angelo, 2015b; Marini, 2017).

Osservando la prima componente (*connected component*) da un punto di vista cronologico, attraverso l'estrazione del *main path*, si evidenzia la traiettoria di evoluzione della ricerca (Colicchia e Strozzi, 2012) rispetto alle tre macroaree tematiche appena illustrate.

Grazie al software Pajek, emerge un percorso di sviluppo della conoscenza (*knowledge path*) che si articola in due rami che si dispiegano lungo un arco temporale che inizia nel 2002. Il primo ramo si radica nel lavoro di Fernández, Morillo, Bordons, e Gómez (2002), che innesca il dibattito sull'appropriatezza del *journal impact factor* (JIF) come metodo di valutazione, mentre il secondo ha origine da lavori più recenti che focalizzano l'attenzione sull'impatto accademico e sociale e sull'ecosistema della ricerca.

All'interno del gruppo di pubblicazioni con la tradizione più consolidata, si evidenzia il discorso sugli effetti dell'introduzione degli indicatori bibliometrici nell'evoluzione dei processi di valutazione della ricerca, da cui si sviluppano contributi che affermano l'importanza di riconoscere la multidimensionalità della qualità e i limiti degli indicatori quantitativi, arrivando a riconoscere la necessità di costruire sistemi di valutazione articolati che tengano in considerazione le diverse dimensioni.

Il secondo ramo, di più recente sviluppo, sostiene la necessità di valutare in modo differenziato l'impatto sociale della ricerca, distinguendolo dall'impatto accademico (*scholarly impact*) e proponendo l'uso delle metriche alternative (*almetrics*) come forma complementare di valutazione (Co-
stas *et al.*, 2015; Haunschild e Bornmann, 2018; Haustein, 2016; Haustein *et al.*, 2016; Lemke *et al.*, 2018). Questa traiettoria di sviluppo apre anche il dibattito sulle criticità dell'ecosistema delle pubblicazioni scientifiche (Chavarro *et al.*, 2017, 2018) e sostiene la necessità di un approccio più completo e inclusivo alla valutazione della ricerca, integrando l'approccio quantitativo della misurazione della qualità (Reategui *et al.*, 2020). L'uso esclusivo di alcune metriche, che possono essere facilmente manipolate, può portare a una focalizzazione ristretta su alcuni tipi di ricerca (Vessuri *et al.*, 2014).

3.2. Seconda componente principale e suo sviluppo

La seconda componente principale di contributi collegati all'interno del campione, costituito da 17 paper, si è sviluppato precedentemente, tra il 1993 e il 2019. Analizzando il contenuto dei paper, è possibile individuare due aree tematiche rilevanti.

Il primo ambito tematico riguarda la distanza epistemica tra la qualità della ricerca e i sistemi di valutazione basati sulla performance, che influenzano le prassi di ricerca (Korytkowski e Kulczycki, 2019). Metriche come il *Journal Impact Factor* possono influenzare la scelta delle domande di ricerca per aumentare la probabilità di pubblicare su giornali ad alto impatto (Rushforth e de Rijcke, 2015). In particolare, alcune discipline come biomedicina ed economia sembrano focalizzarsi esclusivamente su pubblicazioni in riviste posizionate a livello superiore nei *ranking* internazionali (Hammarfelt, 2017).

Nel complesso, i sistemi di valutazione della ricerca influenzano le prassi promuovendo determinati tipi di ricerca o di metodologie (Hammarfelt e de Rijcke, 2015). L'introduzione di un sistema di valutazione che pone l'accento sull'internazionalizzazione e sull'impatto accademico ha portato a conseguenze indesiderate, tra cui la creazione di una cultura del tipo “pubblica o svanisci” (*publish or perish*) che privilegia la quantità rispetto alla qualità, alla fine «aumentando il numero di pubblicazioni su riviste internazionali» (Korytkowski e Kulczycki, 2019) e portando a un'omogeneizzazione dei risultati della ricerca (Kaltenbrunner, 2018) poiché i ricercatori mirano a produrre un lavoro conforme ai criteri dell'esercizio di valutazione nazionale (ad esempio, REF per il Regno Unito, oppure VQR per l'Italia) (McCulloch, 2017). Inoltre, in molte discipline, come le scienze umane e sociali, l'attuale sistema di valutazione ha portato a una perdita di diversità epistemica, poiché i ricercatori sono indotti a pubblicare più frequentemente e solo su riviste indicizzate internazionali, scoraggiandoli quindi dal perseguire progetti non convenzionali (Cañibano *et al.*, 2018). Sebbene un approccio equilibrato quali-quantitativo alla valutazione sia stato proposto fin dall'inizio di questo dibattito (Van Raan, 1996), la cultura derivante dai sistemi di valutazione della ricerca basati sulla performance ha concentrato l'attenzione solo su alcune dimensioni della qualità, trascurando l'impatto sociale (Müller e de Rijcke, 2017).

La seconda area tematica è la più antica e pone l'attenzione su complessità, multidimensionalità e dinamismo del concetto di qualità della ricerca, la cui valutazione è soggetta a interpretazioni individuali e norme delle singole discipline. La distinzione tra caratteristiche intrinseche ed estrinseche della qualità viene richiamata per comprendere e

cogliere la multidimensionalità e le diverse sfumature del concetto di “qualità” nella prospettiva della valutazione della ricerca (Aksnes *et al.*, 2019; Zuccala, 2012). La rilevanza dei fattori intrinseci di qualità, come l’originalità e il rigore metodologico (Hemlin, 1993) è stata progressivamente sostituita dall’esternalizzazione delle misure di qualità per promuovere la responsabilizzazione e la competizione nel settore accademico (Hemlin e Rasmussen, 2006). Il concetto di qualità si è evoluto nel tempo, mentre gli studiosi sottolineano l’esigenza di sviluppare metodi di valutazione in grado di catturare i più ampi benefici sociali e pubblici che la ricerca genera (Bozeman e Sarewitz, 2011). L’approccio alla valutazione è legato agli obiettivi che intende perseguire, che possono essere (i) formativi, con un *focus* sul miglioramento della pratica e con risultati specifici che sono orientati al contesto e al processo (Hemlin e Rasmussen, 2006; Molas-Gallart, 2015); o (ii) sommativo, con particolare attenzione al controllo e alla distribuzione e con risultati sintetici facilmente verificabili e comparabili tra i soggetti (Bozeman e Sarewitz, 2011; Molas-Gallart, 2015). In considerazione dei limiti dell’attuale sistema di valutazione incentrato sulla valutazione attraverso metriche, gli studiosi ricordano la necessità di ampliare gli obiettivi della valutazione verso lo sviluppo e l’impatto sociale (Arocena *et al.*, 2019), suggerendo approcci più orientati al contesto per cogliere le differenze tra i diversi campi del sapere oppure al processo per valutare gli input in termini di risorse dedicate, connessioni con gli stakeholder, e sviluppo delle metodologie. Con un approccio più attento ai dettagli e meno focalizzato sulla generalizzazione dei risultati, alcuni studiosi ritengono possibile costruire un sistema di valutazione in grado di mappare il valore pubblico che le diverse discipline forniscono alla società (Hicks e Holbrook, 2020; Molas-Gallart, 2015).

Dal punto di vista dell’evoluzione di questa componente principale, il secondo *knowledge path* si sviluppa a partire dalla definizione di “qualità”, proseguendo con l’esplorazione degli effetti degli indicatori bibliometrici sul processo di generazione della conoscenza e conseguentemente nella valutazione della ricerca, per approdare al riconoscimento della multidimensionalità della qualità e all’affermazione dell’inadeguatezza degli indicatori quantitativi.

Più specificamente, gli attuali sistemi quantitativi considerano le caratteristiche estrinseche della qualità, valutate indipendentemente dagli assunti epistemologici del testo (che costituiscono invece caratteristiche intrinseche della qualità). L’obiettivo di questo secondo gruppo di contributi riguarda dunque come affrontare i limiti degli attuali sistemi di valutazione proponendo affinamenti e sistemi complementari.

4. Uno sguardo d'insieme

Lo sviluppo dei due cluster di pubblicazioni dimostra come il processo di valutazione della ricerca richiami una pluralità di dimensioni, legate all'ambito disciplinare di riferimento, e nessun singolo metodo o misura possa essere di per sé pienamente soddisfacente ma al contrario sia opportuno riconoscere la complementarità degli strumenti di valutazione.

Da un lato si osserva l'attenzione di alcune comunità scientifiche verso il miglioramento di sistemi basati sulle performance, a cui corrispondono approcci quantitativi alla valutazione della ricerca; dall'altro emergono le istanze delle comunità *epistemic-oriented* (ovvero, le discipline che riflettono sui fondamenti filosofici della ricerca) che focalizzano l'attenzione sugli effetti che il sistema attuale esercita sul processo di creazione della conoscenza, nonché sulla necessità di diffondere adeguata consapevolezza dell'impatto sociale della scienza.

L'analisi della letteratura rivela che le comunità scientifiche, indipendentemente dalla disciplina, enfatizzano sempre più le caratteristiche estrinseche della qualità e concentrano gli sforzi nel miglioramento degli indicatori bibliometrici, come effetto dei sistemi di valutazione che promuovono l'obiettivo di una distribuzione efficiente delle risorse per la ricerca. Originariamente intesi e sostenuti come complemento al processo di *peer review* nella valutazione della ricerca (Van Raan, 1996), i sistemi quantitativi hanno perso nel tempo questa funzione fino ad affermarne la superiorità rispetto ai metodi qualitativi (Abramo e D'Angelo, 2011). In questo contesto, i ricercatori hanno iniziato a pensare attraverso gli indicatori. Di conseguenza, le metriche hanno contribuito a plasmare il “come”, il “cosa” e il “perché” della ricerca (Müller e de Rijcke, 2017). Tuttavia, gli approcci quantitativi tengono conto delle caratteristiche estrinseche della qualità della ricerca (ad esempio, misurate attraverso indicatori bibliometrici), che sono valutate senza riguardo alle affermazioni epistemologiche del testo (cioè, le caratteristiche intrinseche della qualità) (Hammarfelt, 2017). Pertanto, secondo alcuni autori, questi «dispositivi di giudizio» (*judgment device*) (Karpik, 2010) non possono da soli valutare la ricerca. L'assunzione di criteri comuni per valutare la ricerca non implica un consenso sull'applicazione di tali criteri, poiché le affermazioni sulla “qualità della ricerca” sono uniche per ciascun campo (Hammarfelt, 2017).

5. Produttività: ma a quale prezzo?

Alla luce di queste analisi della letteratura, per avere concretezza degli esiti dell'applicazione dei criteri di valutazione della qualità in un contesto reale, ancorché limitato e non paradigmatico, abbiamo indagato⁵ gli effetti delle prassi valutative sullo sviluppo di un settore scientifico-disciplinare non bibliometrico nel periodo 2000-2020, attraverso l'evoluzione qualitativamente delle pubblicazioni dei ricercatori che afferiscono all'ambito disciplinare oggetto di studio.

Lo studio si basa sull'osservazione della dinamica degli indicatori quantitativi di produttività degli accademici (numero di pubblicazioni e citazioni) in un periodo caratterizzato dall'attuazione delle riforme dei sistemi di valutazione a livello istituzionale (VQR) e individuale (ASN).

Ai fini della valutazione, il sistema italiano prevede la distinzione tra discipline bibliometriche, tipicamente le c.d. scienze “dure” alle quali si applicano strumenti matematico-statistici standardizzati (generalmente, ma non esclusivamente le aree CUN 1-9), e discipline non bibliometriche (prevalentemente le aree CUN 10-14 con eccezioni), ovvero le scienze umane e sociali, per le quali l'applicazione degli indicatori bibliometrici risulterebbe problematica perché caratterizzate da tipologie di pubblicazione più eterogenee (per tipologia, lingua, internazionalizzazione) e per la convivenza di paradigmi di ricerca differenti nell'ambito della stessa disciplina.

Per le aree non bibliometriche, l'ANVUR compila e aggiorna costantemente una lista di riviste cosiddette di Classe A, che sono collocate a livelli elevati nelle banche dati delle riviste a livello internazionale. In particolare, per l'area 13, che include le Scienze economiche e statistiche, il criterio predominante ai fini della classificazione in fascia A è esplicitamente «la presenza e il posizionamento delle riviste nelle principali banche dati internazionali»⁶.

Il sistema italiano, per la sua impostazione e complessità, rappresenta tra i paesi europei ed extra-europei un caso interessante perché è l'unico in cui alcune transizioni importanti nella carriera accademica dei ricercatori nell'ambito delle aree bibliometriche sono definite esclusivamente

5. In questa sede è presentata una rielaborazione, a cura delle autrici, dello studio: Za S., Minelli, E.A. & Resca A. (2021). *Publication trends in Organization Studies in Italy: a discussion about higher education policies and rankings*, in ITAIS 2021 Proceedings. 14.

6. ANVUR – “Chiarimenti sul regolamento di classificazione delle riviste” pubblicato il 14/09/2016.

Link: www.anvur.it/news/chiarimenti-sul-regolamento-di-classificazione-delle-riviste/#:~:text=Una%20rivista%20di%20classe%20A,dei%20singoli%20contributi%20effettivamente%20pubblicati.

sulla base di indicatori bibliometrici (Baccini *et al.*, 2019). In un certo senso anche nelle aree non bibliometriche come l'Area 13 è avvenuta una 'bibliometrizzazione', attraverso il progressivo innalzamento del numero di articoli richiesti per superare i concorsi e perché la necessità di pubblicare nelle riviste di classe A ripropone un criterio assimilabile all'IF o all'indicizzazione delle riviste nelle banche dati di riferimento per le discipline bibliometriche. Il risultato a livello di sistema è una indubbia esasperazione del fenomeno *publish or perish*, con effetti "inflattivi" e comportamenti adattivi se non opportunistiche anche nelle discipline non bibliometriche.

Nell'Area 13 (Baccini, 2014), caratterizzata da una notevole eterogeneità delle discipline che ne fanno parte, il settore 13/B3 rappresenta un ambito di interesse specifico proprio per la sovrapposizione tra la tradizione della disciplina, basata su un ruolo preminente della *peer review*, e il "modello valutativo basato sull'impatto, vale a dire la presenza e il posizionamento delle riviste nelle principali banche dati internazionali" (cfr. nota 5). Quest'area offre quindi la possibilità di cogliere e riflettere su alcuni effetti dell'applicazione di criteri bibliometrici nella valutazione della qualità della ricerca in un ambito disciplinare con una tradizione non bibliometrica. Lo studio ha quindi riguardato il trend delle pubblicazioni degli accademici del settore concorsuale 13/B3-Organizzazione Aziendale avvenute esclusivamente su riviste di Classe A, secondo la definizione dell'ANVUR, come *proxy* di ricerca eccellente. I risultati dell'indagine appaiono degni di ulteriore riflessione perché indicano, nel contesto indagato, effetti almeno in parte inattesi.

5.1. *Metodo*

Per rilevare e interpretare le tendenze di pubblicazione sono stati utilizzati semplici strumenti statistici, come medie, percentuali e altre metriche descrittive (El Gibari *et al.*, 2018; OECD, 2008) e sono stati considerati alcuni semplici indicatori, basati su osservazioni di fatti a intervalli regolari (ad esempio, il numero di citazioni per articolo), che indicano posizioni relative e permettono di rilevare le direzioni del cambiamento. Combinando indicatori basati sia su valori assoluti (numero di pubblicazioni, di riviste o di accademici, ecc.) che su valori relativi (come le pubblicazioni per accademico e per paese di origine del ricercatore, le citazioni per articolo, ecc.) è possibile fornire un quadro comprensivo della produzione scientifica e dunque della disciplina accademica (Rauhvargers, 2011).

In questo studio, i dati sono stati elaborati grazie al pacchetto software “R” per l’analisi bibliometrica (www.bibliometrix.org/) integrato dalla codifica per l’elaborazione delle figure. Lo studio si è avvalso delle informazioni disponibili sulle banche dati Scopus e ANVUR, che assicurano la correttezza dei dati per quanto riguarda le attribuzioni di pubblicazione ai docenti del settore scientifico-disciplinare considerato.

La figura 3 rappresenta il modello teorico proposto: in primo luogo, è stato individuato il campione, sulla base della banca dati del Ministero dell’Università e della Ricerca⁷, costituito dai docenti del settore Organizzazione Aziendale (indicatore in valore assoluto) suddivisi in professori ordinari, professori associati, professori a contratto, ricercatori e le due tipologie di ricercatore a tempo determinato (t.d.a ricercatore; t.d.b ricercatore). Una volta selezionati i componenti del campione, a ciascuno di essi è stato assegnato in modo univoco il codice identificativo Scopus per l’attribuzione delle pubblicazioni collegate. Ciò ha richiesto un controllo dettagliato e manuale di ogni codice identificativo Scopus. Dei 201 studiosi individuati, solo 190 sono presenti nel database Scopus con un proprio *ID number*. Il database Scopus è stato preferito come punto di riferimento per questo tipo di ricerca grazie dell’ampia copertura di riviste *peer reviewed* (Meho e Yang, 2007).

È stato quindi estratto il campione delle riviste di Classe A⁸ all’interno delle quali selezionare le pubblicazioni dei docenti del settore Organizzazione Aziendale.

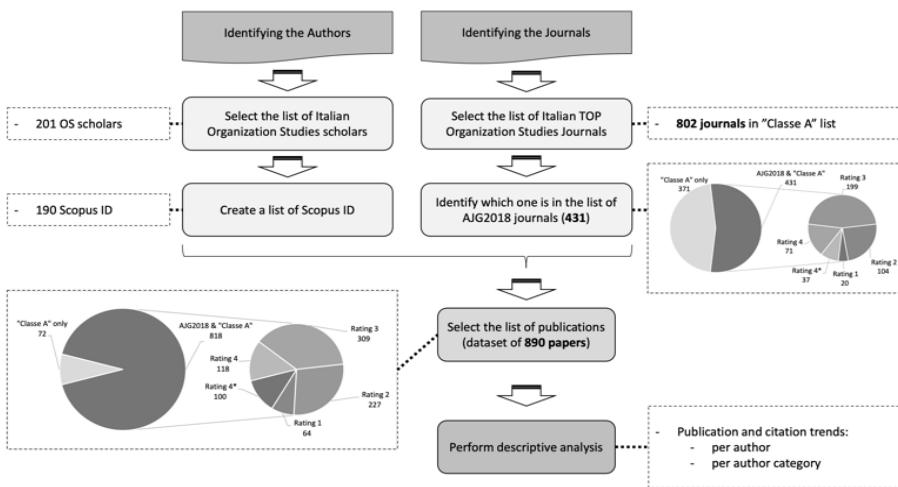
Per una ulteriore conferma della qualità delle riviste di Classe A, l’elenco è stato confrontato con quello definito dalla Chartered Association of Business Schools (ABS) nella Academic Journal Guide (AJG) del 2018 (www.charteredabs.org). Le riviste AJG rappresentano un punto di riferimento per la disciplina degli studi organizzativi, sia dal punto di vista citazionale che di impatto della ricerca. La comparazione tra le 802 riviste di Classe A e le 1582 riviste AJG ha permesso di ricavare 431 riviste presenti in entrambe le liste. In generale, è possibile affermare che le riviste qualificate come Classe A includono i giornali più prestigiosi a livello internazionale nell’area delle scienze sociali e in particolare dell’organizzazione.

Il modello teorico è completato da ulteriori indicatori (indicatori di valore relativo) per analizzare l’andamento delle pubblicazioni e delle citazioni per autore e per categoria di autori (Za e Braccini, 2017), concentrandosi in particolare sulle code delle distribuzioni dei casi (Fig. 3).

7. Dati estratti da <https://cercauniversita.cineca.it/php5/docenti/cerca.php>.

8. Elenco consultato il 4 marzo 2021.

Fig. 3 - Modello teorico di analisi bibliografica applicato al settore concorsuale 13/B3 Organizzazione Aziendale (fonte: Za et al., 2021)



5.2. Evidenze emerse dallo studio del settore Organizzazione Aziendale

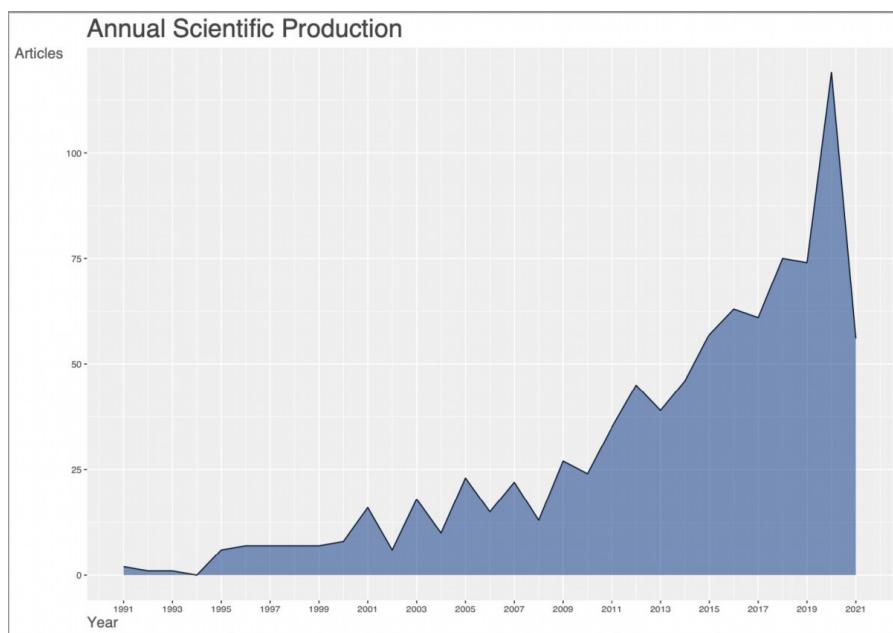
Grazie alla banca dati Scopus, è stato possibile ricavare una serie temporale di pubblicazioni comparse nel campione delle riviste di Classe A selezionate e fare una prima analisi dello sviluppo della ricerca nell'ambito dell'Organizzazione Aziendale. Ai 190 accademici dotati di *ID Scopus* afferiscono 890 pubblicazioni su riviste di Classe A nel periodo considerato (1991-2021). Di queste, 812 compaiono su riviste classificate AJG.

La Fig. 4 evidenzia un incremento significativo nel numero di pubblicazioni su riviste ad alto impatto nel periodo considerato.

Dopo una curva sostanzialmente piatta fino alla fine degli anni '90, una linea spezzata ha segnato la direzione all'inizio del nuovo millennio. Lo zig-zag è continuato anche nell'ultimo decennio. Tuttavia, il numero di pubblicazioni per anno è più che raddoppiato dal 2015 al 2020, passando da 57 a 119. Il fatto che all'inizio del 2021, quando l'indagine empirica si è conclusa, 47 articoli fossero già stati pubblicati, indica che il trend di pubblicazione era costantemente in crescita.

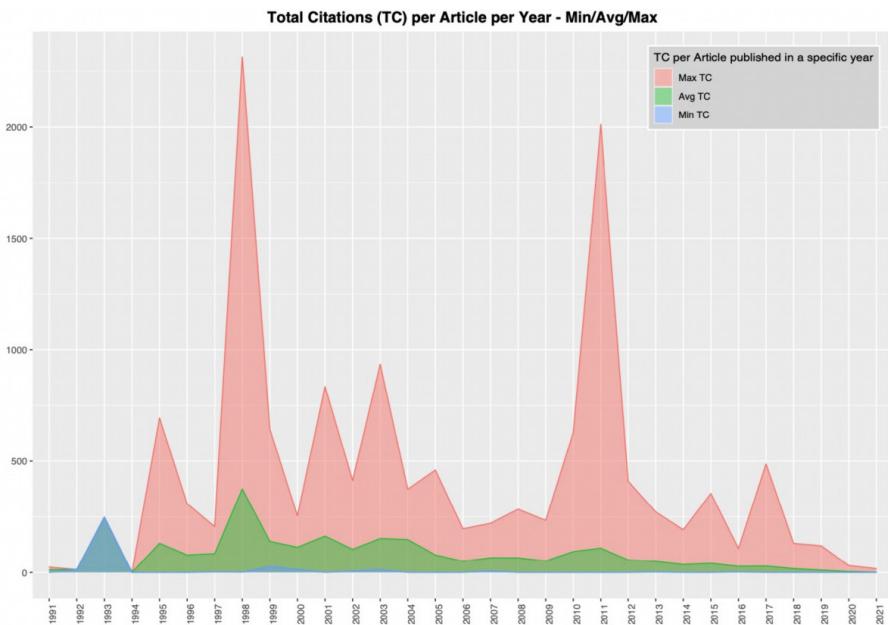
Tuttavia, la numerosità delle pubblicazioni di Classe A, rilevante sia per la carriera dei ricercatori che per i finanziamenti dei progetti di ricerca universitari, non permette di cogliere l'impatto scientifico della ricerca, rilevabile invece attraverso il numero di citazioni di queste pubblicazioni. A differenza dell'osservazione delle pubblicazioni in un determinato arco temporale il cui andamento è facilmente intelligibile, l'assegnazione delle

Fig. 4 - Numero di pubblicazione su riviste di Classe A in Organizzazione Aziendale dal 1991



citazioni a un anno specifico non è così semplice. Infatti, gli articoli di riviste, come qualsiasi altro prodotto di ricerca, possono essere citati per un lungo periodo di tempo. Per affrontare questo problema, sono stati presi in considerazione tre fattori. In primo luogo, è stata considerata la media delle citazioni per articolo pubblicato in un determinato anno. Possiamo aspettarci che i primi articoli pubblicati abbiano una media inferiore di citazioni all'anno, ad eccezione degli articoli seminali, poiché è pratica comune citare le pubblicazioni relativamente più recenti. Questa pratica dimostra la familiarità degli studiosi con prodotti di ricerca aggiornati e quindi la capacità di produrre conoscenza nuova. Questo comportamento però può essere controbilanciato dal fatto che gli articoli hanno bisogno di tempo per essere conosciuti e citati in lavori successivi. In secondo luogo, si osserva la pubblicazione con il maggior numero di citazioni in un determinato anno. Lo scopo di questo indicatore è identificare le cosiddette "opere seminali" che vengono spesso citate molto tempo dopo la pubblicazione. In terzo luogo, abbiamo anche deciso di individuare le pubblicazioni con il minor numero di citazioni all'anno per capire il loro livello di impatto rispetto agli altri due indicatori.

Fig. 5 - Massimo numero, media e minimo di citazioni per articolo pubblicato per ogni anno (fonte: Scopus)



La Fig. 5 suggerisce che i lavori seminali sono stati pubblicati principalmente nella seconda metà degli anni '90 e nei primi anni del nuovo millennio, ad eccezione di una pubblicazione del 2011 che ha già raggiunto circa 2.000 citazioni. È probabile che una pubblicazione del 2017 diventi seminale, visto che ha già raggiunto le 500 citazioni. In relazione alle citazioni medie delle pubblicazioni, possiamo osservare che l'inizio del millennio e la fine del primo decennio sono caratterizzati da opere di notevole impatto. Senza considerare gli ultimi anni, poco rilevanti per via del tempo che una pubblicazione impiega ad affermarsi nel panorama della ricerca, si può prevedere che la maggior parte delle pubblicazioni dell'ultimo decennio non vedrà un aumento del tasso di citazione negli anni a venire.

L'indicatore del minimo di citazioni viene aggiunto per delineare il profilo della ricerca a impatto limitato, poiché solo all'inizio del millennio non esistevano pubblicazioni prive di citazioni. È interessante notare che dal 1992 al 1994 i tre indicatori di citazione si sovrappongono, segnalando un livello omogeneo di qualità nella produzione della ricerca negli studi organizzativi.

In conclusione, il confronto tra tasso di pubblicazione e tasso di citazione suggerisce che ad un aumento del numero di pubblicazioni nell'ultimo decennio non corrisponde un aumento del numero di citazioni, e si può anche ipotizzare che vi sia stata forse una diminuzione negli anni.

Questa osservazione induce a riflettere sul comportamento degli accademici italiani nel periodo di attuazione delle riforme del sistema di valutazione della ricerca universitaria. Conseguentemente, è stata condotta una analisi delle pubblicazioni rispetto alla composizione dei membri accademici della comunità scientifica di Organizzazione Aziendale.

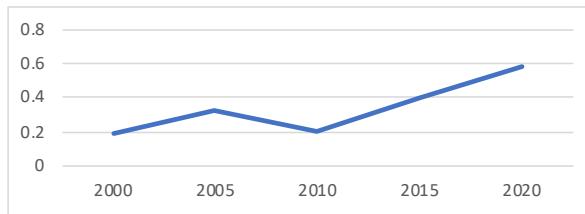
La Tab. 1 distingue tra professori ordinari, associati, straordinari, ricercatori e i due contratti di ricercatore t.d.a e t.d.b introdotti dalla legge 240/2010.

Tab. 1 - Professori afferenti al settore concorsuale “Organizzazione Aziendale”

Position	2000	2005	2010	2015	2021
Professore ordinario	5	11	20	29	63
Professore associato	17	28	28	67	84
Professore straordinario	1	6	1	0	5
Ricercatore	18	25	72	32	10
Ricercatore t.d.a				4	15
Ricercatore t.d.b				9	24
<i>Total</i>	<i>41</i>	<i>70</i>	<i>121</i>	<i>141</i>	<i>201</i>

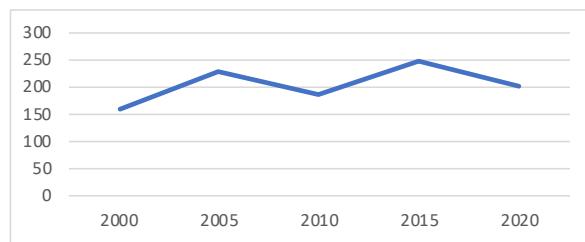
I dati mostrano un aumento di cinque volte del numero di accademici nel periodo osservato. Tuttavia, il numero di pubblicazioni è aumentato in modo più che proporzionale. Nel 2000, un totale di 41 accademici ha scritto 8 articoli sulle riviste di Classe A, ma nel 2020 gli articoli sono diventati 119 a fronte di 201 professori. Ciò significa che il rapporto delle pubblicazioni per un singolo anno passa da 0,19 a 0,58 per accademico, suggerendo che una percentuale maggiore di accademici sia riuscita a pubblicare (Fig. 6).

Fig. 6 - Numero medio di pubblicazioni in riviste di Classe A per accademico e per anno nel settore “Organizzazione Aziendale”



L'analisi delle citazioni per accademico rappresenta un passaggio cruciale in questa discussione. Nella Fig. 7 il numero totale di citazioni negli anni 2000, 2005, 2010, 2015 e 2020 è stato suddiviso per il numero di accademici afferenti al settore negli stessi anni. Si osserva come, nonostante una tendenza in leggera crescita e due picchi nel 2005 e 2015, l'incremento percentuale del numero medio di citazioni per accademico sia inferiore all'aumento del numero medio di pubblicazioni nello stesso periodo, suggerendo che all'incremento quantitativo della produttività accademica non corrisponde un analogo effetto di impatto della ricerca.

Fig. 7 - Numero medio di citazioni di articoli in riviste di Classe A per accademico e per anno nel settore “Organizzazione Aziendale”



Dunque, a una maggiore produzione scientifica in termini quantitativi non sembra corrispondere una maggiore significatività delle nuove conoscenze.

Un ulteriore approfondimento è stato condotto nell'ambito del settore restringendo l'analisi ai primi 10 autori con il maggior numero di pubblicazioni e il maggior numero di citazioni che rappresentano il 5% della forza lavoro accademica totale. Dal 1991 al 2020, 331 articoli su 834 pubblicazioni sono attribuibili ai primi 10 autori. In altre parole, 191 accademici sono autori di circa il 60% del numero totale di pubblicazioni e 10 autori

hanno prodotto il resto. Nello stesso periodo, il numero totale di citazioni era 41.175, mentre 29.569 citazioni erano appannaggio dei 10 autori con il maggior numero di citazioni. Ciò significa che solo il 30% circa delle citazioni fa riferimento a 191 accademici che hanno pubblicato il 60% degli articoli. La stessa analisi è stata svolta con riferimento a un arco temporale più limitato, ovvero dal 1991 al 2015. Su 442 articoli apparsi dal 1991 al 2015, 130 pubblicazioni ovvero circa il 29% del totale sono state realizzate dai primi 7 autori (che analogamente all'analisi precedente rappresentano circa il 5% dei 141 accademici in forza nel 2015). Pertanto, 134 studiosi hanno scritto circa il 61% delle pubblicazioni in linea con i risultati del 2020, con una concentrazione di autorialità comparabile. Per quanto riguarda le citazioni, nello stesso periodo, i primi 7 studiosi sono stati citati 10.222 volte su 35.077 citazioni, circa il 29% del totale. In questo caso il contrasto con il 2020 è significativo. Dal 1991 al 2015 le citazioni delle pubblicazioni dei 134 studiosi di Organization Studies (corrispondenti al 61% della produzione scientifica del periodo) sono state circa il 71% del totale, contro il 30% circa delle citazioni riferite a 191 studiosi nel 2020 (corrispondenti però a circa il 60% della produzione accademica), suggerendo uno spostamento da un sostanziale equilibrio tra pubblicazioni e citazioni nel 2015 ad una concentrazione delle citazioni nel 2020 su alcuni autori.

È opportuno chiedersi qual è il motivo di questo cambiamento radicale tra il 2015 e il 2020, per quanto riguarda la distribuzione delle citazioni. Negli ultimi anni, alcune università, per competere nel settore dell'alta formazione a livello globale, hanno reclutato alcuni accademici di fama internazionale che hanno all'attivo decine di pubblicazioni su riviste di Classe A che sono state citate migliaia di volte. Diversamente dal tradizionale approccio accademico, la strategia di ricerca perseguita dalle università ha avuto quindi un impatto significativo sugli indicatori di output della ricerca del settore Organizzazione Aziendale.

I risultati suggeriscono che, nel periodo considerato, le riforme introdotte in Italia hanno modificato i comportamenti non solo degli accademici ma anche delle istituzioni, almeno nell'ambito del settore concorsuale considerato. Le evidenze raccolte dimostrano un incremento del volume delle pubblicazioni sulle riviste di Classe A, dove l'accettazione degli articoli è subordinata a un rigoroso processo di *peer review* che è una prassi focalizzata sulle caratteristiche intrinseche della qualità. Tuttavia, la qualità misurata sulla base del numero delle citazioni, che ne rileva l'impatto sulla comunità scientifica di riferimento (caratteristica estrinseca della qualità), non è cresciuta in modo corrispondente.

Non solo: il sistema ha insegnato alle istituzioni (almeno in questo caso) come le regole introdotte possano essere interpretate in modo stra-

tegico e diventino una leva importante per vincere la sfida della produttività e migliorare l'impatto scientifico (a condizione però di avere sufficienti risorse).

6. Conclusioni

Questo capitolo si propone di sollecitare alcune riflessioni sul bilanciamento delle dimensioni della qualità nella valutazione della ricerca.

Ogni campo del sapere parla una propria lingua: la sfida è trovare metodologie di valutazione rispettose degli assunti ontologici ed epistemologici, dei significati e delle finalità di ogni ricerca, senza pretesa di generalizzazione e omologazione delle prassi.

La valutazione della qualità della ricerca è un obiettivo complesso e sfumato che richiede dunque un approccio olistico e multidimensionale, che bilanci le caratteristiche intrinseche ed estrinseche della qualità. L'uso di metriche tradizionali non permette di valutare adeguatamente la coerenza ontologica ed epistemologica. E questo problema, apparentemente lontano dall'esercizio quotidiano della ricerca, ha un rilievo crescente rispetto ai progetti scientifici più complessi, nei quali i campi del sapere si intersecano per costruire risposte soddisfacenti ai bisogni di conoscenza dell'umanità.

Ne consegue che, nell'ambito della ricerca interdisciplinare, che presuppone prospettive epistemologiche complesse e assume importanza crescente in molti campi della ricerca scientifica attuale (si pensi ad esempio alle relazioni tra organizzazioni e mercato del lavoro, oppure alla medicina del territorio e alla epidemiologia dopo l'esperienza della pandemia nei paesi europei), è necessario approfondire la riflessione sulle scelte ontologiche ed epistemologiche che il disegno della ricerca definisce ma che non sempre sono adeguatamente colte dagli attuali sistemi di valutazione della ricerca.

Il cambiamento istituzionale nei metodi di reclutamento e progressione della carriera accademica, nonché di valutazione della produttività scientifica delle università in Italia ha dato risalto a questa tematica, favorendo un allineamento del comportamento degli accademici italiani (si vedano anche i capitoli seguenti). Lo stimolo delle riforme ha indotto una maggiore produttività scientifica, il cui valore intrinseco viene accertato attraverso la *peer review* e sancito grazie alla pubblicazione su riviste di Classe A.

Tuttavia, almeno nell'ambito disciplinare considerato in questo studio, a una produttività maggiore non corrisponde un analogo incremento delle citazioni, caratteristica estrinseca della qualità che segnala invece l'interesse per lo specifico prodotto scientifico all'interno della comunità accademica. Emerge dunque un disallineamento delle dimensioni della qualità: a un

contenuto ritenuto scientificamente rilevante (perché pubblicato in riviste di Classe A) non sempre corrisponde un interesse nella comunità accademica e conseguentemente nel dibattito scientifico. Questa osservazione, scaturita dall'analisi di un singolo settore concorsuale particolarmente esposto al dibattito sulle prassi valutative, sollecita una maggiore attenzione a eventuali fenomeni inflattivi, che potrebbero verificarsi anche in altri settori come conseguenza di una logica *publish or perish* (Baccini *et al.*, 2019).

Le riforme del primo ventennio nell'università italiana hanno favorito quindi la produttività accademica nell'ambito dei giornali di Classe A, riorientando gli sforzi dei ricercatori dalle monografie alle riviste internazionali, dove i processi di *peer review* assumono un ruolo fondamentale nella definizione degli standard di qualità di una pubblicazione scientifica (si veda il capitolo 6). Ma qui entrano in gioco anche le logiche che orientano le scelte delle case editrici, a cui sono demandate le decisioni su cosa pubblicare e, indirettamente, sull'evoluzione di una disciplina.

6. Effetti della macchina editoriale sullo sviluppo delle discipline accademiche: un approccio neo-istituzionalista

di *Maria Rucsandra Stan, Eliana Alessandra Minelli*

1. Università di massa e produttività accademica

Partendo dalle riflessioni sui meccanismi che attualmente dominano i sistemi di valutazione della ricerca, questo capitolo propone un approfondimento sulle logiche istituzionali che intervengono nel sistema di produzione della conoscenza scientifica sotto un duplice profilo: sia rispetto ai processi di pubblicazione, che rappresentano il meccanismo fondamentale di valutazione della qualità di un prodotto scientifico, sia rispetto ai processi di allocazione delle risorse, che rappresentano l'effetto della valutazione della ricerca scientifica. Gli studi condotti hanno permesso di comprendere come le logiche istituzionali dominanti orientino i processi di generazione di nuova conoscenza attraverso dinamiche complesse e come plasmino i confini delle discipline scientifiche.

La nascita dell'università di massa, aperta a fasce sempre più ampie di popolazione, ha generato negli anni '80 dello scorso secolo l'esigenza di un controllo più attento dell'uso delle risorse pubbliche destinate alla formazione terziaria e ha portato all'introduzione di sistemi di finanziamento della ricerca basati sui risultati. Conseguentemente, metriche quali gli indici citazionali e le classifiche (Hazelkorn, 2011), riguardanti riviste e istituzioni, hanno conosciuto un notevole sviluppo per sostenere e fluidificare i meccanismi di produzione della conoscenza scientifica e per assicurarne qualità, diffondendosi universalmente poiché transnazionale è l'autorità della scienza (Whitley, 2003; Whitley *et al.*, 2010). Lo scopo di questo complesso sistema è dunque valutare la qualità della ricerca realizzata con i fondi dei contribuenti e allo stesso tempo conseguire un equilibrio tra obiettivi di ricerca e finalità pubbliche (Gläser e Whitley, 2007). Classifiche di tutti i tipi sono diventate strumenti cruciali per i ricercatori per raggiungere i livelli di produttività e di qualità prescritti e ottenere il

riconoscimento del valore scientifico del proprio lavoro ai fini della carriera. La competizione per attrarre le risorse economiche e conseguire riconoscimenti professionali avviene sul terreno delle riviste che si posizionano a livelli più elevati nei *ranking* internazionali (Tennant *et al.*, 2017). Le classifiche assumono dunque un significato centrale nel contesto del mercato globalizzato della conoscenza (Castells, 2000), in cui la conoscenza stessa è percepita come una *commodity* (Slaughter e Leslie, 1997).

Certamente la creazione di indicatori di performance oggettivi ha permesso di misurare i risultati della ricerca scientifica e valutare le prestazioni individuali in modo più accurato, ma anche di presentare alla comunità dei contribuenti un quadro informativo più trasparente e comprensibile, che dovrebbe migliorare la fiducia verso il settore della ricerca (Townley *et al.*, 2003). Ne risulta un sistema ampiamente diffuso, basato su indicatori bibliometrici che definiscono l'impatto ai fini della valutazione della ricerca (Abra-mo e D'Angelo, 2011), ma che riducono l'attenzione rispetto ad altre dimensioni della qualità (Aksnes *et al.*, 2019). Se, da un lato, la confortante certezza dei numeri induce ad attribuire fiducia a classifiche e metriche, dall'altro non mancano prese di posizione critiche (Aksnes, 2005; Higgins *et al.*, 1990), in considerazione della multidimensionalità della qualità della ricerca (Aksnes *et al.*, 2019). Esse invocano la necessità di incorporare altri fattori nel processo di valutazione attraverso un bilanciamento dei metodi utilizzati.

Il sistema ha così 'educato' ricercatori e istituzioni ad accettare i *ranking* come indicatori di qualità e ad elaborare e perseguire strategie per adeguarsi a tali criteri di eccellenza nella ricerca, al fine di sopravvivere e ottimizzare il ritorno del proprio investimento di tempo ed energie (Levitt e Dubner, 2005). Ciò ha indotto i ricercatori a concentrare l'impegno nell'ambito di certi paradigmi del proprio campo di ricerca (Cunliffe, 2022), limitando in ultima istanza la diversità epistemica di cui dispongono (Kaltenbrunner, 2018; Rushforth e de Rijcke, 2015).

Attualmente, in molti paesi i sistemi di valutazione della produttività individuale ai fini della progressione di carriera, indipendentemente dalle procedure o dalle istituzioni coinvolte nel processo di valutazione, si basano sulle pubblicazioni nelle riviste a più alto impatto (Tennant *et al.*, 2017). Di conseguenza, in tutto il mondo gli accademici considerano la pubblicazione dei risultati delle ricerche come parte della propria attività e condizione per lo sviluppo della carriera.

Oltretutto, il passaggio all'università di massa non ha fatto aumentare solo i numeri della popolazione studentesca, ma ha necessariamente accresciuto le dimensioni del corpo docente, in particolare nelle posizioni ai livelli iniziali della carriera, per fare fronte alla crescente domanda di formazione. Soprattutto in Italia, la crescita della *faculty* ha reso più compe-

titiva e meno scontata la mobilità in ascesa, minando il tradizionale potere delle oligarchie accademiche (Giglioli, 1979). Dalla fine del secondo millennio, il numero dei ricercatori, postdoc, *part-time* o *full-time* in posizioni non stabili si è incrementato significativamente, esacerbando la competizione per pubblicare sulle riviste più prestigiose, secondo la logica *publish or perish*. Dunque, appare necessario studiare i meccanismi delle pubblicazioni scientifiche e approfondire il ruolo dell'*editor* e le logiche sottese al processo di pubblicazione. La scelta di concentrare l'indagine sugli *editor*, anziché sui publisher, si basa sul ruolo centrale che gli *editor* rivestono nel processo di *peer review*. Essi selezionano i revisori, valutano le loro recensioni e prendono decisioni cruciali sull'accettazione o il rifiuto degli articoli, influenzando direttamente la qualità scientifica delle pubblicazioni. Al contrario, i publisher si occupano principalmente di aspetti commerciali e logistici non intervenendo nei processi di *peer review*.

Il capitolo si apre con una prima riflessione sul ruolo degli *editor* delle riviste scientifiche, responsabili del funzionamento di un meccanismo fondamentale per il sistema di valutazione della qualità della ricerca, la pubblicazione, che interagisce con i processi di carriera degli accademici.

2. Lo sviluppo del sistema delle pubblicazioni¹

Il contributo alla valutazione apportato dalle riviste scientifiche può essere classificato in due categorie o fasi: ex-ante ed ex-post rispetto alla pubblicazione del manoscritto, ciascuna delle quali richiede approcci e metodi diversi. Nella fase di pre-pubblicazione, la valutazione si focalizza sulle caratteristiche intrinseche della ricerca ed è realizzata per mezzo del processo di *peer review*. Delicato e complesso, soggetto a distorsioni, è comunque l'unico metodo che consente la valutazione delle caratteristiche intrinseche. La fase di post-pubblicazione invece mette in rilievo la misurazione delle caratteristiche estrinseche della ricerca, misurate sia in termini di impatto accademico, ad esempio attraverso indici delle citazioni e JIF, sia sociale, attraverso le *altmetrics*. Attraverso questi due passaggi, il sistema associa alla pubblicazione un certo livello di prestigio, spendibile per l'avanzamento della posizione accademica degli autori.

In questo quadro, la tendenza ad associare il prestigio accademico alle pubblicazioni sulle riviste si è potuta consolidare grazie alla storica disponi-

1. Elaborazione da Stan, M.R. (2023). *Insights into evaluation of research quality: what role for the researcher?*. Tesi di dottorato. LIUC-Università Cattaneo.

nibilità degli accademici a svolgere gratuitamente il lavoro di revisione tra pari (*peer review*). Ne beneficiano le società commerciali che pubblicano le riviste scientifiche e prosperano in una posizione di mercato dominante, legittimandosi grazie alla funzione assunta rispetto ai meccanismi di valutazione della qualità della ricerca (Baldwin, 2017a; Baldwin, 2017b; Fyfe *et al.*, 2017).

Il processo di revisione tra pari rappresenta il motore delle pubblicazioni scientifiche e definisce la qualità e il potenziale impatto di un manoscritto (Spier, 2002). Tale processo assicura quindi lo sviluppo della disciplina grazie alla trasmissione della conoscenza attraverso le riviste (Tennant *et al.*, 2017). Dunque la *peer review* esercita una funzione di verifica del lavoro dei ricercatori estremamente cruciale che deve essere svolta con grande accuratezza in considerazione della rilevanza delle conseguenze che essa comporta per la disciplina e per il destino accademico dei ricercatori. Nel corso del tempo, la pressione a raggiungere elevati standard accademici è aumentata, rendendo le carriere accademiche molto più sfidanti che nel passato. In questo contesto, l'*editor*, quale arbitro del meccanismo di pubblicazione, gioca un ruolo critico specialmente durante la fase di valutazione iniziale del manoscritto, nella quale prende la decisione di inoltrarlo ai revisori oppure di rigettarlo, e perciò diventa importante studiare quali fattori possono influire sulla sua scelta (Casnici *et al.*, 2017; Hamilton *et al.*, 2020; Pontille e Torny, 2015; Ross-Hellauer *et al.*, 2017), che indirettamente si riverbera anche sulla carriera accademica degli autori.

L'obiettivo di questo studio è quindi comprendere non solo come il sistema di valutazione dei manoscritti influenza il modo in cui la conoscenza viene prodotta (per esempio, quali paradigmi vengono privilegiati) ma anche gli effetti sulle carriere degli accademici. Come è già stato osservato nel precedente capitolo, il contesto in cui i meccanismi di pubblicazione si inseriscono è caratterizzato dalla rilevanza di indicatori quantitativi per la valutazione della produttività scientifica, che hanno avuto ampi effetti sulla comunità scientifica e sulle modalità di conduzione della ricerca (Shapiro, 2017). In Italia, se da un lato emerge l'enfasi sulla necessità di produrre e pubblicare un numero crescente di manoscritti, indotta dalle procedure dell'ASN per stimolare una maggiore produzione in termini quantitativi (*publish or perish*), dall'altro non sempre a tale incremento corrisponde un impatto accademico significativo, che può risultare addirittura decrescente (si veda il capitolo 5). Inoltre, la pressione a pubblicare su riviste ad alto *impact factor* può portare a una maggiore omologazione della ricerca a certi modelli, in cui solo alcuni paradigmi, argomenti e metodi sono considerati validi (Cunliffe, 2022).

2.1. La peer review nella letteratura

Il ruolo della *peer review* si è evoluto nel tempo. Storicamente, ha avuto il ruolo di custode della conoscenza per la comunità accademica. Oggi, tuttavia, agisce come meccanismo di controllo della qualità per determinare l'idoneità di un manoscritto a essere integrato nella conoscenza accademica esistente «imponendo indirettamente quale rivista sia adatta per quello specifico pezzo di ricerca» (Tennant *et al.*, 2017, p. 4).

Prima del XIX secolo, la *peer review* era un processo di collaborazione tra autori ed editori di pubblicazioni, che prevedeva principalmente scambi amichevoli sotto forma di lettere con lo scopo di selezionare contenuti da pubblicare a causa della limitatezza dei canali di distribuzione disponibili (Baldwin, 2017b). Con l'avvento dell'università di massa e l'espansione delle riviste scientifiche, il processo di revisione tra pari si è evoluto per concentrarsi maggiormente sulla valutazione dell'integrità scientifica (Csizar, 2016) in conseguenza dell'aumento del volume e della complessità del materiale inviato agli editori. Ciò ha portato all'esternalizzazione del processo di revisione, con il conseguente coinvolgimento di «revisori esterni provenienti dalle società sponsorizzatrici delle riviste» (Tennant *et al.*, 2017, p. 5). Poiché il prestigio accademico è stato sempre più associato alle pubblicazioni sulle riviste, la revisione tra pari ha assunto il significato di giudizio oggettivo e processo di costruzione del consenso. Gli editori commerciali hanno approfittato della convinzione tradizionalmente diffusa tra gli accademici, secondo cui la *peer review* è un obbligo volontario degli studiosi, per legittimare le proprie riviste e stabilire una posizione dominante nel mercato (Baldwin, 2017a, 2017b; Fyfe *et al.*, 2017). Tuttavia, ciò ha comportato il passaggio da un processo collaborativo tra studiosi a un processo più standardizzato governato da un'industria editoriale altamente competitiva (Larivière *et al.*, 2015). La *peer review* si è inoltre integrata con l'avanzamento della carriera accademica, poiché le università e gli istituti di ricerca, nonché i sistemi di *Higher Education*, privilegiano le pubblicazioni sulle riviste e il loro impatto come misura di prestigio nell'ambito della competizione accademica internazionale (Tennant *et al.*, 2017; Thornton, 2004).

La letteratura ha indagato vari aspetti del processo di *peer review*, dai fattori che influenzano la probabilità di accettazione di un paper (Hing *et al.*, 2011; Northam *et al.*, 2014) agli aspetti etici (Nicholas *et al.*, 2017; Wager *et al.*, 2009) nonché il ruolo degli *editor*, fondamentale nell'era delle riviste ad alto *impact factor*.

Il ruolo dell'*editor* è centrale nel processo di *peer review* perché assume la funzione di custode delle caratteristiche, e dunque del prestigio, della rivista. La sua azione si esercita attraverso una prima scrematura delle

proposte non conformi rispetto ai canoni di qualità previsti per la pubblicazione (Hall *et al.*, 2015; Ross-Hellauer *et al.*, 2017). L'*editor* può decidere quindi se rigettare immediatamente il testo oppure inoltrarlo ai revisori, da cui sorge l'esigenza di individuare i ricercatori più competenti rispetto alla tematica e alla metodologia (Popkin, 1989). Sono stati indagati anche i criteri adottati per selezionare i revisori nonché le modalità di formazione della decisione finale sul manoscritto (Donnellon *et al.*, 1986; Hamilton *et al.*, 2020; Mullen *et al.*, 2013) e di risoluzione di eventuali contrasti tra revisori (Squires, 1991). Il ruolo dell'*editor* appare in tutta la sua delicatezza poiché è responsabile del flusso di comunicazione tra revisori e autori, comprese eventuali raccomandazioni (Popkin, 1989; Ross-Hellauer *et al.*, 2017) che possono influenzare decisioni e scelte degli autori.

La letteratura non ha trascurato neppure il problema dei comportamenti inappropriati, che possono manifestarsi sia prima sia dopo la pubblicazione (Shelomi, 2014, p. 52), screditando il processo di *peer review* (Horbach e Halffman, 2018, p. 4). Inoltre, limiti di tempo, impreparazione dell'*editor* su specifici temi, indisponibilità di revisori competenti, ma anche conflitti di interesse (Charlton, 2004; Rosenblum *et al.*, 2020) possono influenzare l'imparzialità e l'equilibrio del processo di *peer review* in modo diretto (Epstein, 1995) o indiretto (Kassirer e Campion, 1994; Reardon, 2008; Street *et al.*, 1998). La letteratura riconosce i limiti intrinseci del processo e la sua soggettività (Margalida e Colomer, 2016; Walker e Rocha da Silva, 2015) e individua nell'*editor*, attraverso il suo impegno a minimizzarli, il garante della qualità scientifica per il progresso della conoscenza (Hall *et al.*, 2015).

Tuttavia, la letteratura non ha indagato con sufficiente attenzione le logiche istituzionali che influenzano il processo decisionale degli *editor* e dunque orientano le scelte di pubblicazione le quali, come è stato sottolineato, rappresentano un fattore cruciale per determinare l'avanzamento di carriera a livello individuale ma anche la base per la valutazione della produttività delle istituzioni universitarie e degli enti di ricerca.

2.2. *Un approccio neo-istituzionalista per studiare il ruolo degli editor*

La nostra ricerca si basa sulla concettualizzazione delle logiche istituzionali sviluppata dal filone di ricerca che da Friedland e Alford (1991) giunge alla teorizzazione proposta da Thornton e Ocasio (1999). Essa integra la dimensione strutturale, normativa e simbolica delle istituzioni e le definisce come un insieme di pratiche e di artefatti simbolici a disposizione delle organizzazioni e degli individui per elaborare significati e modelli di riferimento che «condizionano le scelte degli attori per il *sensemaking*, il vocabolario

che utilizzano per motivare l'azione e il loro senso di sé e l'identità» (Thornton, Ocasio, e Lounsbury, 2012, p. 2). Le logiche istituzionali si manifestano in modi differenti nelle organizzazioni in relazione al tempo storico e al contesto culturale, sono politicamente difese e si riproducono attraverso concrete relazioni sociali che hanno contenuti strumentali e rituali. Perciò danno senso alle attività quotidiane degli individui e forniscono un modello di comportamento a livello organizzativo (Canhilal *et al.*, 2016).

Normalmente nello stesso ambito convivono logiche multiple che possono influenzare il comportamento degli individui e delle organizzazioni, e possono portare al conflitto (Greenwood *et al.*, 2011) o alla cooperazione tra le logiche (Boch Waldorff *et al.*, 2013). La pluralità di logiche presenti nelle organizzazioni offre quindi la possibilità di disporre di schemi cognitivi alternativi (Fligstein e McAdam, 2012) che costituiscono un fondamento per orientare conseguentemente le azioni dei diversi sottogruppi che coesistono nell'ambiente organizzativo.

La complessità delle organizzazioni moderne favorisce la compresenza di logiche differenti e richiede la capacità di gestirle (Greenwood *et al.*, 2011; Reay e Hinings, 2009). Tra le organizzazioni più complesse dal punto di vista delle logiche emergono le istituzioni universitarie (Cai e Mountford, 2022), protagoniste di un ecosistema dell'*Higher Education* in cui il settore dell'editoria, che assicura la fluidità del sistema delle pubblicazioni, ricopre un ruolo centrale.

Il mercato delle pubblicazioni collegato all'*Higher Education* è già stato oggetto di uno studio secondo la prospettiva delle logiche istituzionali (Thornton, 2004), che ha evidenziato l'influenza di due logiche prevalenti: da un lato quella di mercato, caratterizzata dall'obiettivo di profitto; dall'altro quella accademico-professionale, orientata alla qualità accademica.

La decisione finale in merito alla pubblicazione è definita dalle logiche dominanti dell'organizzazione e del campo disciplinare (Thornton e Ocasio, 2012) e incorpora nel manoscritto oggetto di pubblicazione una visione di qualità della ricerca (Primack *et al.*, 2019; Watson e Xue, 2019).

La nostra ricerca ha pertanto indagato quantitativamente gli effetti delle logiche istituzionali nei processi editoriali delle riviste di riferimento di due settori concorsuali, rispettivamente uno bibliometrico e uno non bibliometrico, nell'ambito del sistema universitario italiano, con particolare attenzione alla fase iniziale del processo di revisione dei manoscritti.

A differenza dello studio della Thornton, questa ricerca riconosce la complessità del ruolo dell'*editor* delle riviste accademiche e la sua duplice natura (Serpa *et al.*, 2020) di responsabile della pubblicazione di una rivista (di cui è proprietaria una società commerciale a cui l'*editor* risponde) e di scienziato (che ha come riferimento il proprio settore disciplinare).

Quindi gli *editor* sono ricercatori che appartengono a una comunità scientifica, di cui condividono prassi di comunicazione accademica e idee (Borgman, 1989) e sono allo stesso tempo i custodi della conoscenza scientifica disseminata attraverso le pubblicazioni accademiche (Tennant *et al.*, 2017).

La nostra ricerca ha analizzato l'influenza di quattro ordini di istituzioni (accademico, professionale, burocratico, e mercato), ciascuno dei quali assume specifiche connotazioni rispetto agli ambiti in cui si manifesta all'interno dei processi di *peer review*, ovvero rispetto agli obiettivi della revisione, alle fonti di legittimazione e ai sistemi di sostegno.

Ai fini della nostra indagine sull'influenza delle logiche istituzionali, la duplice natura, di ricercatore e di responsabile della rivista, del ruolo che l'*editor* incarna richiede di considerare distintamente gli ambiti accademico e professionale. La logica accademica si fonda sul valore intrinseco della generazione e della disseminazione di nuova conoscenza (Brew, 2003) e sul riconoscimento di specifici doveri da parte dei membri della comunità scientifica per favorire lo sviluppo della conoscenza, come la revisione tra pari. La logica professionale invece si richiama ai principi mertoniani di distacco dall'oggetto della propria attività, ovvero al disinteresse emotivo ed economico rispetto al processo di *peer review*, nonché allo scetticismo, che stimola l'analisi e l'esercizio di spirito critico (Merton, 1973b). Tuttavia, nella specifica circostanza le due logiche sono fortemente collegate in quanto la logica professionale si radica nella comunità accademica nel momento in cui si determina la qualità della ricerca (K.M. Kallio *et al.*, 2021, p. 140).

Oltre a queste due logiche, l'ordine burocratico definisce un approccio alla qualità basato sui concetti di correttezza, imparzialità, parità di trattamento, definiti da Weber (Meyer *et al.*, 2014) e si focalizza sul rispetto delle norme, procedure e criteri di qualità definiti dalle riviste. Infine, la logica di mercato (Thornton e Ocasio, 2012) enfatizza la produzione, il valore e l'utilità della ricerca (Gupta, 2016). Questa prospettiva si focalizza sulla qualità estrinseca e sui rankings e persegue l'obiettivo principale del conseguimento di risultati finanziari (Baffy *et al.*, 2020).

2.3. *La metodologia e il campione*

La ricerca si basa su una *survey* proposta *online*, costruita sulla base di elementi già testati provenienti da rilevazioni condotte nell'ambito di precedenti ricerche (Cerejo, 2014; COPE, 2020; Hamilton *et al.*, 2020; Hing *et al.*, 2011; Northam *et al.*, 2014; Ross-Hellauer *et al.*, 2017), integrati da alcune domande scaturite dal colloquio con gli *editor* durante alcune interviste preliminari. Complessivamente, sono stati sviluppati 25 item per

indagare i fattori che influenzano la valutazione iniziale di un manoscritto da parte dell'*editor*.

Preliminarmente al lancio del questionario online, è stato condotto uno studio pilota basato su interviste cognitive per verificare tutti gli elementi del questionario. Sei *editor* accademici sono stati selezionati in base alla rappresentatività delle riviste selezionate nel loro ambito scientifico, nonché alla loro disponibilità a partecipare al progetto. Questa fase di verifica si è resa necessaria data la complessità del ruolo dell'*editor*.

Ciascuna parte del questionario è stata strutturata in modo da rappresentare le quattro logiche identificate in letteratura: accademica (ad esempio, riferimento adeguato alla letteratura precedente o metodi rigorosi); professionale (ad esempio, allineamento della metodologia tra rivista e articolo o accuratezza dei contenuti); burocratico (ad esempio, conteggio delle parole o stile di riferimento); e logica di mercato (ad esempio, argomento caldo o reputazione dell'autore). Le risposte, ove opportuno, sono state strutturate secondo una scala Likert a 7 punti, con una risposta aggiuntiva “Non lo so” al fine di migliorare l’accuratezza, l’onestà e la completezza dei dati raccolti, in particolare per domande complesse o tecniche o per le quali gli intervistati potrebbero non avere sufficiente esperienza o competenza (Wetzelhütter, 2020). La scala dispari è stata ritenuta più appropriata per consentire agli intervistati di esprimere un’opinione neutrale e ridurre il rischio di risposte acquiescenti (Komorita e Graham, 1965; Pearse, 2011). Solo le prime due domande della *survey* erano obbligatorie, specificamente la disciplina di riferimento e il tipo di revisione tra pari adottata, in considerazione della varietà di sfumature che il ruolo dell'*editor* può assumere nella realtà editoriale.

Nel dicembre 2021, un campione di oltre 1.500 accademici che ricoprivano incarichi di *editor* nelle riviste internazionali selezionate è stato invitato a partecipare alla *survey* online. Il campione è stato costruito sulla base di due discipline, che fanno riferimento rispettivamente a un settore concorsuale bibliometrico e a un altro non bibliometrico secondo la normativa italiana. Il campione include un numero uguale di riviste per ciascuna area, individuate come Disciplina A per il settore concorsuale non bibliometrico e come Disciplina B per il settore bibliometrico.

Ogni settore concorsuale prevede indicazioni specifiche in merito ai requisiti di pubblicazioni su riviste scientifiche per partecipare all’Abilitazione Scientifica Nazionale (ASN), il meccanismo attraverso cui avviene la valutazione della produttività scientifica ai fini dell’avanzamento di carriera degli accademici italiani. Tuttavia, questo è solo un passaggio preliminare all’inserimento nei ruoli accademici permanenti, in quanto rappresenta la condizione per partecipare ai concorsi banditi a livello locale nelle università italiane.

In particolare, nel caso dei settori non bibliometrici, le pubblicazioni che possono essere presentate dagli accademici per la valutazione devono necessariamente comparire sulle riviste incluse nelle liste prodotte dall'ANVUR per ciascuna area concorsuale. Le riviste sono raggruppate in fasce di merito in base a criteri di eccellenza riconosciuti a livello internazionale, tra i quali emergono il rigore del processo di *peer review* e il prestigio di cui godono tra la comunità scientifica di riferimento.

La principale differenza nella categorizzazione delle riviste tra discipline bibliometriche e non-bibliometriche consiste nel fatto che nel primo caso il riferimento è alle basi dati internazionali, mentre nel caso delle discipline non bibliometriche l'inclusione delle riviste delle liste è il frutto di un accordo che coinvolge le società scientifiche a livello nazionale, ancorché la legge italiana sottolinea che solo i risultati significativi ed eccellenti a livello internazionale dovrebbero essere presi in considerazione per la valutazione.

Quindi, nei processi di valutazione a livello individuale (ASN) e istituzionale (VQR) è riconosciuto e incorporato il meccanismo della valutazione della produzione scientifica compiuta dagli *editor* attraverso il processo di *peer review*.

2.4. I risultati dell'analisi

L'obiettivo della nostra ricerca è identificare i fattori che influenzano le decisioni degli *editor* nelle fasi iniziali del processo di pubblicazione e svelare i presupposti sottostanti che, in ultima analisi, legittimano la qualità di un manoscritto.

Alla *survey* hanno risposto in modo significativo 122 *editor*. È stata quindi condotta una analisi fattoriale esplorativa del manoscritto utilizzando gli standard SPSS per individuare i fattori che influenzano la decisione editoriale sui manoscritti scientifici.

L'analisi fattoriale ha rivelato quattro componenti che intervengono nella valutazione iniziale del manoscritto da parte dell'*editor*, che riflettono rispettivamente la logica del mercato, la logica professionale, la logica burocratica e quella accademica. Precedenti ricerche nel campo dell'istruzione superiore (Canhilal *et al.*, 2016) e del mercato editoriale dell'istruzione superiore (Thornton, 2004) hanno indicato l'esistenza di molteplici logiche che modellano il contesto.

I risultati della *survey* proposta agli *editor* accademici forniscono preziosi spunti per un'ulteriore riflessione sulla valutazione dei manoscritti e sull'evoluzione della definizione di qualità nella ricerca pubblicata.

La logica del mercato è caratterizzata da obiettivi ben definiti che mirano a migliorare le prestazioni della rivista, come il raggiungimento di un più elevato del JIF, favorendo fattori come argomenti di tendenza o reputazione dell'autore. La logica burocratica è focalizzata sulla definizione e sul rispetto di linee guida e politiche per le riviste accademiche (conteggio delle parole o titolo). La logica accademica riflette le posizioni della comunità accademica cui afferisce l'*editor*, come le scelte relative ai riferimenti alla letteratura, mentre la logica professionale affonda le sue radici nelle responsabilità inerenti al ruolo dell'*editor* considerando l'allineamento metodologico tra manoscritto e requisiti della rivista.

L'analisi di affidabilità dei quattro fattori ha rivelato che solo la logica mercato e la logica professionale sono risultate accettabili dal punto di vista statistico. Tuttavia, l'esiguità del numero di elementi che costituiscono le componenti accademiche e burocratiche del questionario suggerisce l'opportunità di ulteriori approfondimenti.

Inoltre, rispetto alla selezione dei revisori, le loro caratteristiche personali non risultano statisticamente rilevanti quando si tratta della valutazione di un manoscritto. Ciò fornisce un certo grado di affidabilità e obiettività del processo, rafforzando l'idea di un meccanismo di controllo basato sulle competenze dei revisori. Infatti, gli *editor* di entrambe le discipline tendono ad allineare la selezione dei revisori con quella degli esperti del settore. In particolare, la Disciplina B (bibliometrica) pone una forte enfasi su questa prospettiva in quanto seleziona i revisori anche in base al fatto che siano citati all'interno del manoscritto. Infine, lo studio mostra che gli editori si aspettano che i revisori adempiano alle proprie responsabilità fornendo commenti contestualmente appropriati e rifiuti giustificati, confidando nella deontologia per supportare il processo di valutazione del manoscritto.

Dunque le logiche dominanti sono quella del mercato e quella professionale, come indicato dai valori alfa di Cronbach. La letteratura ha evidenziato che nelle organizzazioni che sperimentano logiche multiple il comportamento individuale e organizzativo ne è influenzato (Greenwood *et al.*, 2010). La logica del mercato suggerisce che gli *editor* potrebbero essere inclini a scegliere un manoscritto sulla base di alcune variabili identificate all'interno di tale componente: per esempio potrebbero essere sensibili ad argomenti considerati interessanti nel momento attuale, che richiamano maggiormente l'attenzione. Il predominio della logica del mercato implica che la qualità del contenuto del manoscritto proposto sia meno enfatizzata nella decisione finale (Gupta, 2016) e siano privilegiati altri criteri collegati alla collocazione della rivista nei ranking e agli interessi finanziari (Baffy *et al.*, 2020).

La relazione tra le due logiche dominanti suggerisce una possibile cooperazione di logiche (Boch Waldorff *et al.*, 2013), dove la logica del mercato implica una preferenza per articoli che aumentino la performance della rivista in termini di JIF, e la logica professionale sostiene gli standard di performance della rivista attraverso la valutazione selettiva dei manoscritti. Inoltre, come confermato da Mullen *et al.* (2013), è importante che gli *editor* distinguano chiaramente tra la loro identità di ricercatori all'interno della comunità accademica (cioè, logica accademica) e il loro ruolo di *editor* (cioè, logica professionale) nel giudicare il contenuto del manoscritto.

In sintesi, i risultati dell'analisi fattoriale suggeriscono che gli *editor* intervistati riconoscono la distinzione dei ruoli, con le logiche di mercato e professionali che risultano più influenti, e le logiche burocratiche e accademiche che hanno un ruolo secondario e sono meno impattanti. Tuttavia, a causa delle dimensioni ridotte del campione, non è possibile indagare ulteriormente per capire se le logiche burocratiche e accademiche rimangono in secondo piano a causa della consapevolezza da parte degli *editor* del proprio ruolo, oppure se fungono da strutture concorrenti per un sottoinsieme di partecipanti (Fligstein e McAdam, 2012) per facilitare il cambiamento quando necessario.

Nel complesso, questi risultati fanno luce sul processo decisionale e sui criteri utilizzati dagli *editor* per valutare la qualità dei manoscritti: in realtà è proprio l'azione dell'*editor*, che avvia, coordina e legittima questa fase del processo, che rappresenta il perno del sistema di valutazione su cui si fondano le decisioni di avanzamento di carriera, a livello individuale, e le valutazioni sul livello di qualità raggiunto dalle università e dagli enti di ricerca, a livello istituzionale. Per questa ragione è opportuno indagare questa fase del processo di generazione della conoscenza, da cui scaturisce il riconoscimento della conoscenza vera come espressione della responsabilità delle comunità scientifiche nella costruzione del sapere.

3. Le implicazioni delle logiche sottese alla revisione dei pari nell'avanzamento delle carriere accademiche

L'identificazione delle logiche emergenti che intervengono nella decisione iniziale dell'*editor* ha implicazioni significative riguardo la valutazione dei manoscritti e stimola riflessioni importanti in merito alle condizioni di successo di una carriera accademica, in particolare per i ricercatori all'inizio della carriera (Nicholas *et al.*, 2017; Zhang e Yu, 2020), a causa della crescente associazione tra prestigio accademico e pubblicazione su rivista (Tennant *et al.*, 2017; Thornton, 2004). Ciò può influenzare le strategie di

pubblicazione e, in ultima analisi, modellare la comunicazione accademica, soprattutto nelle discipline non bibliometriche, per le quali l'inclusione delle riviste negli elenchi validi ai fini della valutazione dei candidati all'ASN dipende dalle società scientifiche riconosciute a livello nazionale. I nostri risultati, insieme a studi precedenti (Tennant *et al.*, 2017; Thornton, 2004), rafforzano l'idea che il mercato editoriale internazionale dell'istruzione superiore si sia gradualmente spostato verso logiche di mercato e professionali per definire la qualità della ricerca rappresentata dalle proprie pubblicazioni. Le riviste sono iniziative imprenditoriali che operano esternamente rispetto ai principi e all'assetto della comunità accademica, introducendo così logiche che dipendono dai loro meccanismi di sostenibilità economica (ad esempio conferenze, pubblicità, oppure maggiore attenzione verso determinati argomenti di attualità).

Poiché la ricerca è diventata più complessa e voluminosa, il processo di *peer review*, originariamente ristretto alla cerchia di collaboratori e colleghi, è stato almeno in parte esternalizzato a revisori associati alle varie riviste presenti nel mercato editoriale. Infatti, queste riviste hanno potuto costruire il loro modello di business capitalizzando da un lato sulla necessità degli studiosi di affermare la qualità dei propri lavori di ricerca e dall'altro sull'etica accademica, che prevede la disponibilità alla revisione tra pari come forma di sostegno allo sviluppo della scienza. Sfruttando la necessità degli accademici di legittimare le riviste in cui poter pubblicare i propri manoscritti, le società editoriali competono per una posizione dominante sul mercato attraverso l'IF delle riviste di cui sono proprietarie (Fyfe *et al.*, 2017; Larivière *et al.*, 2015). Il cambiamento di scenario ha portato alla luce logiche come quella del mercato e quella professionale, che hanno lentamente depotenziato le logiche accademiche e hanno cambiato le priorità delle riviste e dei loro *editor*. Di conseguenza, hanno modificato ciò che è rilevante per riuscire a pubblicare i manoscritti. La standardizzazione del processo di *peer review* controllato dall'industria editoriale, altamente competitiva (Larivière *et al.*, 2015), ha spostato l'attenzione sulle caratteristiche estrinseche della qualità (cioè su *proxy*) e lontano da quelle intrinseche (cioè dal contenuto) (Bridges, 2011). I risultati dello studio che abbiamo condotto mostrano come le diverse logiche diano origine a diverse dimensioni della qualità (Aksnes *et al.*, 2019) al di là del contenuto della ricerca stessa. Queste dimensioni si riflettono in termini di argomenti di tendenza, potenziale copertura geografica, collegamento con conferenze (ad esempio, pubblicazione di volumi di atti di conferenze) e la rilevanza dell'articolo nel tempo. Pertanto, dato che il prestigio accademico e conseguentemente l'avanzamento di carriera dipendono in larga misura dalla pubblicazione su riviste, i ricercatori, indipendentemente dal contesto na-

zionale, si trovano di fronte all'obiettivo di produrre ricerca che rifletta le dimensioni legittimate dai processi di pubblicazione, al di là del contenuto. L'esempio dell'ASN dimostra chiaramente come la normativa italiana abbia integrato e legittimato nella carriera accademica le logiche provenienti dal mercato editoriale internazionale, che non scaturiscono direttamente dai principi della scienza. Tuttavia, è opportuno sottolineare che questo fenomeno è diffuso in tutti i contesti accademici, non solo quello italiano, poiché il prestigio accademico è associato, a livello internazionale, alla capacità di pubblicare su riviste di elevato prestigio.

Ma le considerazioni sugli effetti delle logiche dominanti nel mercato delle pubblicazioni accademiche non si esauriscono con le dimensioni della qualità. La predominanza di una logica di mercato implica che solo alcune posizioni ontologiche ed epistemologiche siano considerate più consone e quindi accettabili per produrre ricerca di qualità. Stabilire obiettivi che siano allo stesso tempo impegnativi e raggiungibili, come ottenere una posizione accademica di ruolo grazie al prestigio che deriva dalle pubblicazioni secondo le logiche del mercato editoriale, può ispirare le persone a conformarsi a tali logiche per raggiungere i risultati sperati (Locke e Latham, 2002). Inoltre, la semplice assegnazione di obiettivi che non sono in linea con la logica accademica può avere effetti indesiderati (Welsh e Ordóñez, 2014), come la pressione per raggiungere soglie di pubblicazione sempre più elevate per scopi di sviluppo della carriera. Essa può comportare una focalizzazione tematica ristretta con progressi incrementali della conoscenza piuttosto che ricerche di elevato impatto sociale. Cunliffe (2022, pp. 6-7) parla esplicitamente di una certa forma di cecità ontologica definita come la mancanza di attenzione e di volontà di «comprensione delle ontologie alternative e delle loro implicazioni per la metodologia e la teorizzazione» nell'ambito degli studi organizzativi. Fondamentalmente, ciò evidenzia il fatto che la maggior parte delle comunità accademiche cercano di promuovere un pluralismo di approcci per studiare la nostra realtà (cioè istituzioni, sistemi, meccanismi e processi). Tuttavia, le ontologie soggettiviste vengono raramente affrontate, mentre le ontologie oggettiviste sembrano avere una prevalenza nello studio e nella teorizzazione della nostra realtà. Accordandosi a questa cecità ontologica in un contesto dominato da logiche derivate dal mercato editoriale porta a un'emarginazione degli studiosi che si impegnano in modi alternativi di produzione della conoscenza. Osserviamo quindi come l'associazione del prestigio accademico con il posizionamento della rivista, la valutazione della ricerca basata sulla performance e orientata alla quantità, la priorità di alcune dimensioni della qualità (ad esempio, derivate dalle logiche di mercato dominanti) e l'ege monia di alcune posizioni ontologiche possono allontanare i ricercatori da

una ricerca approfondita, anche se non necessariamente generalizzabile, portando a un'omogeneizzazione dei paradigmi e alla perdita di ricchezza della conoscenza. La deriva paradigmatica si manifesta nella proceduralizzazione dell'immaginazione, dove la creatività e l'intuizione, pur riconosciute come necessarie, vengono incanalate attraverso quadri rigidi e protocolli che limitano l'apertura e la spontaneità nella generazione di idee. Ad esempio, i metodi come l'abduzione e l'analisi narrativa vengono spesso standardizzati attraverso alberi decisionali e modelli strutturati, riducendo la flessibilità interpretativa (Cunliffe, 2022).

Questa prima parte del nostro studio si propone dunque di indagare e comprendere gli effetti delle logiche emergenti che definiscono il processo di *peer review*, modellando in definitiva i discorsi delle comunità scientifiche attraverso le strategie di pubblicazione. L'iniziativa imprenditoriale del mercato editoriale è diventata la condizione attraverso cui è possibile la valutazione dei manoscritti e il riconoscimento del prestigio accademico che consente lo sviluppo della carriera dei ricercatori. La commercializzazione e la standardizzazione del processo di *peer review* hanno influenzato di conseguenza anche la definizione del concetto di qualità della ricerca e orientato le priorità dei ricercatori nel perseguire lo sviluppo della carriera accademica.

Il nostro obiettivo è incoraggiare gli studiosi a meditare sugli effetti di queste logiche, che si basano su presupposti di fondo che differiscono da quelli della scienza, per migliorare progressivamente il sistema sulla base di una crescente consapevolezza dei limiti e delle potenzialità. Il sistema di valutazione della ricerca è fondamentale per accertare la qualità della nuova conoscenza prodotta e responsabilizzare nell'uso delle risorse ma, cosa ancora più importante, non esiste un sistema alternativo a quello attuale. I ricercatori e gli *editor* devono essere consci dei potenziali pregiudizi che il sistema attuale può causare ai processi di generazione di nuova conoscenza e impegnarsi a garantire che la loro ricerca sia inclusiva e rappresentativa. Ciò significa considerare una varietà di prospettive, comprese quelle considerate *non ortodosse* e *non tradizionali*, nello sviluppo della conoscenza: la ricerca prospera in un contesto in cui esiste una diversità di voci che si esprimono all'interno della comunità accademica.

In un mondo ideale, il sistema di valutazione della ricerca dovrebbe integrare criteri di impatto sociale, economico e culturale, valorizzando saggiamente l'interdisciplinarità e l'*open science*. Le interazioni tra publisher, autori, accademia e società civile potrebbero trasformarsi in collaborazioni più trasparenti e partecipative, rompendo il monopolio dei grandi editori. Il concetto di qualità si amplierebbe includendo metriche narrative, impatti pratici e capacità di rispondere ai bisogni sociali. Il reclutamento privilegherebbe carriere non lineari, esperienze diverse e contributi concreti, incenti-

vando percorsi ibridi e collaborativi. Questo approccio promuoverebbe una ricerca più umana, accessibile e connessa alla società, democratizzando il sapere accademico.

4. Logiche istituzionali e finanziamento nella ricerca scientifica²

Il nostro studio ha approfondito l'analisi delle logiche istituzionali in un contesto caratterizzato da finanziamenti pubblici, dove la logica *corporate* gioca un ruolo rilevante perché legata alla sopravvivenza dei centri di produzione della conoscenza scientifica. L'indagine si è orientata verso gli effetti dei modelli di finanziamento *performance-based* sulle pratiche di ricerca, riverberandosi in ultima istanza sull'impatto sociale della ricerca stessa.

Abbiamo considerato la ricerca medica come l'area scientifica più rilevante ai fini di questo studio, poiché i suoi risultati hanno un effetto significativo e immediato sul benessere della società. È fondamentale che la ricerca medica dia priorità alla generazione di un impatto reale sulla società e si impegni attivamente a migliorare la salute e il benessere degli individui e delle comunità. Le ragioni che giustificano la scelta dell'ambito medico sono principalmente due: in primo luogo, la sanità è un settore vitale che riceve investimenti significativi dai governi in tutto il mondo e ha un impatto significativo sulle economie nazionali. In secondo luogo, le organizzazioni sanitarie forniscono un contesto ideale per studiare le logiche istituzionali, dato il coinvolgimento di professionisti, politiche governative, strutture e pratiche che influiscono sul modo in cui le organizzazioni operano e prendono decisioni (Reay *et al.*, 2021).

L'adozione di una prospettiva basata sulle logiche istituzionali ci consente di comprendere come le norme, i valori e le convinzioni all'interno del settore sanitario si riflettono nelle pratiche di ricerca, influenzando in definitiva i risultati conseguiti e l'impatto sulla società.

Oggetto di questo studio sono le logiche emergenti nella gestione della ricerca scientifica in un Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) italiano, che si intersecano con le logiche istituzionali che dominano il mercato delle pubblicazioni scientifiche.

Il contesto in cui gli IRCCS operano è caratterizzato infatti dall'interazione tra molteplici logiche che scaturiscono dai diversi obiettivi e interessi degli stakeholder coinvolti.

2. Elaborazione da Stan, M.R. (2023). *Insights into evaluation of research quality: what role for the researcher?*. Tesi di dottorato. LIUC-Università Cattaneo.

Gli IRCCS possono essere pubblici o privati, specializzarsi in una disciplina (monotematici) oppure più discipline mediche (politematici) e svolgere attività di cura e di ricerca. I professionisti che lavorano negli IRCCS si possono distinguere in due categorie: coloro che lavorano prevalentemente con i pazienti (ricercatore clinico) e conducono ricerca clinica; e coloro che si focalizzano sulla ricerca scientifica, condotta prevalentemente nei laboratori (ricercatore puro). La loro *seniority* viene riconosciuta dal Ministero della Salute³ sulla base dell'esperienza e dei risultati conseguiti e alcuni possono avere duplice affiliazione, ovvero IRCCS e università.

Il finanziamento degli IRCCS avviene per oltre la metà dell'importo sulla base della produzione scientifica, la quale assume due modalità differenti: la ricerca corrente (fonte principale di entrate) e la ricerca finalizzata (contributi o fondi assegnati su base competitiva da enti nazionali o internazionali, oppure dal Ministero della Salute per affrontare bisogni urgenti, come la crisi pandemica). Il punto cruciale è che l'ammontare complessivo dei finanziamenti relativi alla ricerca corrente viene attribuito dal Ministero della Salute sulla base della valutazione delle pubblicazioni scientifiche dell'anno precedente in misura proporzionale rispetto al valore *dell'impact factor* totale conseguito dall'IRCCS, normalizzato⁴ sulla base di una serie di parametri, quali il numero degli autori affiliati all'IRCCS e il loro ruolo nella pubblicazione, nonché il tipo di articolo e la classificazione della rivista.

L'obiettivo di questo studio è comprendere come sistemi e criteri di valutazione basati sulle prestazioni (*performance-based*) possano modificare le pratiche di ricerca all'interno di un ambiente istituzionale, fortemente influenzato da standard disciplinari internazionali e nazionali, e possano avere un impatto sui risultati della ricerca. L'IRCCS rappresenta un caso di studio significativo grazie al suo ambiente istituzionale unico, influenzato da vari enti e fattori, come il governo nazionale e regionale, il *top management* dell'IRCCS e le convenzioni internazionali specifiche relative alle diverse discipline praticate. In questo contesto possono emergere numerose logiche istituzionali che sostengono valori, norme e credenze differenti e potenzialmente contrastanti.

3. www.fpcgil.it/wp-content/uploads/2020/01/D.M.personale_ricerca_sanitaria.pdf
Decreto 20 novembre 2019, n. 164. Regolamento recante valutazione del personale di ricerca sanitaria.

4. La procedura di normalizzazione dell'Impact Factor (IF) per gli IRCCS in Italia è stabilita dal Ministero della Salute per valutare in modo equo la produzione scientifica. Le riviste sono classificate in quartili in base al loro IF grezzo, assegnando un valore specifico a ciascun quartile (Q1: 6, Q2: 4, Q3: 2, Q4: 1); per le riviste con IF ≥ 10 si utilizza il valore reale fino a un massimo di 35. La tipologia di documento (articoli originali, review, lettere con dati scientifici) influenza ulteriormente il punteggio, con valori ridotti per tipologie meno significative. Viene considerata anche la posizione degli autori affiliati all'IRCCS (Fonte: www.salute.gov.it/imgs/C_17_pagineAree_4234_2_file.pdf).

Concentrandosi sul ruolo delle istituzioni, della cultura, delle strutture di ricompensa e del contesto sociale e storico nel modellare il processo di ricerca, questo studio si propone di capire l'influenza di questi fattori sul modo in cui la ricerca viene progettata, condotta e diffusa, con l'obiettivo ultimo di migliorarne l'impatto sociale. Il Ministero della Salute italiano definisce annualmente gli obiettivi di produttività degli IRCCS, a cui corrispondono l'allocazione di risorse e di fondi. Gli obiettivi, come specificato nei decreti ministeriali di riferimento, si basano sulle prassi di valutazione a livello internazionale e sul mercato delle pubblicazioni accademiche. Il Ministero richiede di migliorare il proprio *impact factor* totale incrementando il numero di pubblicazioni sulle riviste internazionali all'interno di uno specifico intervallo di JIF. Le direttive del Ministero si radicano in una concezione di sistemi di valutazione basati sulle performance e altamente competitivi che individua con chiarezza una delle logiche istituzionali dominanti negli IRCCS: la logica aziendale (*corporate logic*). Questa logica si pone in contrasto con la logica alla radice della professione di un ricercatore, ovvero la logica professionale che nel caso specifico consiste nel perseguire l'avanzamento della conoscenza per amore della scienza.

In questo contesto, l'azione dei ricercatori si colloca tra le due logiche contrastanti, quella professionale e quella aziendale, che richiedono uno sforzo di bilanciamento delle opposte istanze e a cui il *top management* è chiamato a fornire un contributo cruciale per il raggiungimento dell'equilibrio.

Il nostro studio ha rilevato come le logiche professionali e quelle aziendali entrino in conflitto nel processo di ricerca. Qui si rivela fondamentale la mediazione di una terza logica (Jia *et al.*, 2017, 2019), individuata come la “logica di comunità”, per compiere una riconciliazione e ottenere risultati soddisfacenti sia sul fronte professionale, ovvero per i ricercatori, sia sul fronte istituzionale, ovvero per l'IRRCS. La “logica di comunità” (che funge anche da meccanismo di riconciliazione tra le due precedenti – professionale e aziendale) rappresenta un processo in cui diverse logiche istituzionali coesistono e vengono armonizzate per soddisfare le molteplici esigenze e pressioni presenti in un contesto organizzativo complesso come quello degli IRCCS. Questa logica permette di tradurre richieste spesso contraddittorie – come quelle dello stato (criteri di valutazione), del mercato (brevettazione e trasferimento tecnologico) e della scienza (rigore metodologico) – in strategie operative attraverso l'uso di meccanismi che creano senso comune tra gli attori coinvolti.

Il ruolo del direttore scientifico emerge quale imprenditore istituzionale nello sforzo di riconciliazione per ottenere i finanziamenti delle ricerche da parte del Ministero e assicurare allo stesso tempo ai ricercatori le condizioni indispensabili per lo svolgimento delle proprie attività di ricerca. In

questa interazione tra logiche, l'azione intenzionale del *top management* diventa cruciale per far convergere gli sforzi dei ricercatori verso il conseguimento degli obiettivi strategici rilevanti per assicurare la continuità della ricerca scientifica intrapresa dall'istituto. Ne emerge una cultura della ricerca attenta all'impatto sociale della ricerca ma anche alla sostenibilità finanziaria dell'Istituto, attraverso un processo di produzione della conoscenza che non può prescindere da un approccio interdisciplinare.

In questo studio, sono esaminate le manifestazioni e le conseguenze di un contrasto tra le logiche che interagiscono nel processo di produzione di nuova conoscenza e nell'impatto sociale della ricerca medica. Un meccanismo di riconciliazione emerge dalla ricerca dinamica di un equilibrio tra istanze contrastanti e si rivela fondamentale per assicurare la continuità dei processi. Esso è visibile, ad esempio, nel ruolo del direttore scientifico che media tra i vincoli normativi e le esigenze dei ricercatori, sfruttando logiche comunitarie per promuovere un ambiente collaborativo e produttivo. La nostra ricerca, a differenza di studi precedenti (Jia *et al.*, 2017, 2019), ha potuto rilevare l'impatto positivo di un'azione del *top management* a sostegno della logica di riconciliazione. Lo studio non solo ha esplorato la dinamica tra le tre logiche coinvolte nella gestione della ricerca (aziendale, professionale, comunitaria) ma anche come queste plasmano la cultura della ricerca coltivata all'interno di questo istituto.

4.1. *Metodologia, raccolta dati e individuazione delle logiche sottese*

Lo studio, di carattere esplorativo, combina il metodo dei casi (Yin, 2018) con la *Grounded Theory* (Gläser e Strauss, 1999) in un orizzonte costruttivista (Hatch, 2018). Da un punto di vista epistemologico, lo studio si allinea all'approccio interpretativista, che sottolinea l'importanza di cogliere una ricca interpretazione soggettiva delle interazioni sociali e del contesto studiato. Per questa ragione la ricerca ha valorizzato alcuni dialoghi informali, ma molto significativi, colti nella fase preliminare di costruzione del disegno della ricerca.

L'approccio *Grounded Theory* si è sviluppato a partire da uno schema iniziale di logiche istituzionali e concetti derivati dalla teoria neostituzionalista ed è stato raffinato con verifiche e validazioni successive per produrre un modello capace di indagare gli effetti della valutazione sul processo di produzione della ricerca scientifica.

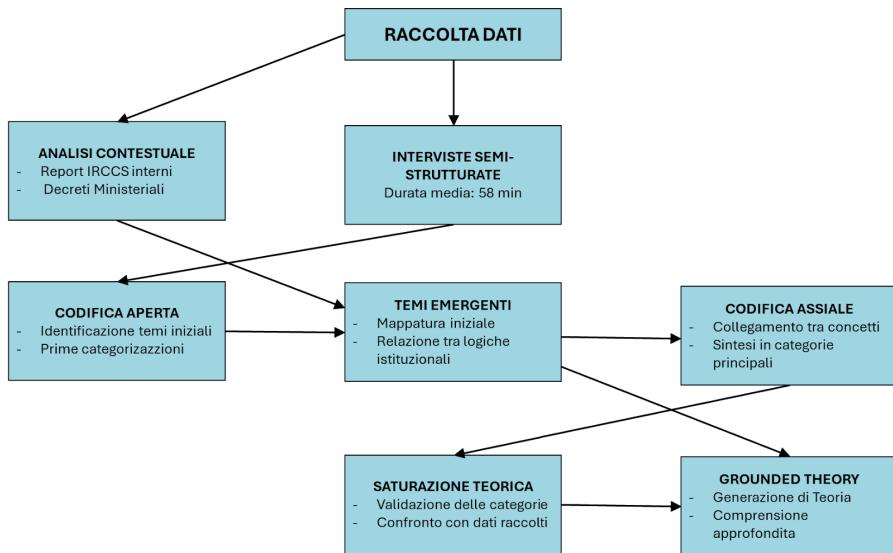
A partire da colloqui informali che hanno permesso di cogliere obiettivi contrastanti tra l'istituto e i ricercatori, i dati raccolti sono stati analizzati induttivamente attraverso un processo iterativo basato su tecni-

che di confronto tra evidenze empiriche e letteratura rilevante (Gläser e Strauss, 1999).

Il metodo dei casi (Yin, 2018) ha permesso di approfondire le ragioni e le modalità delle situazioni analizzate e comprendere le potenziali implicazioni per l'azione (Creswell e Creswell, 2018; Merriam e Tisdell, 2015). Si è rivelato efficace per comprendere gli effetti dei sistemi di valutazione sul modo in cui viene prodotta la ricerca in campo sanitario, per sua natura di rilevante impatto sociale (de Rijcke *et al.*, 2016; Hammarfelt e de Rijcke, 2015; Müller e de Rijcke, 2017; Rushforth e de Rijcke, 2015).

I dati sono stati raccolti attraverso 32 interviste semistrutturate approfondite, proposte a esperti con ruoli rilevanti all'interno dell'IRCCS, disponibili a partecipare al nostro studio.

Fig. 1 - Rappresentazione schematica della metodologia utilizzata nello studio



Sulla base di studi precedenti (Müller e de Rijcke, 2017), le domande delle interviste hanno riguardato la scelta degli argomenti di ricerca, la selezione delle riviste, l'identificazione dei collaboratori (Hammarfelt e de Rijcke, 2015), gli effetti della valutazione della ricerca sul processo di ricerca. Ai dati raccolti è stata applicata la *content analysis* (Krippendorff, 2019), in base a esperienze, percezioni e prospettive dei partecipanti.

Con l'intervista sono state raccolte informazioni non solo sugli idealtipi di logiche che interagiscono nel processo di creazione di nuova conoscenza, ma anche sulle modalità di gestione di istanze in conflitto e sul riallineamento dei diversi obiettivi, a livello istituzionale e individuale, attraverso un meccanismo di riconciliazione basato sul senso di comunità e promosso dal *top management*. In questo modo, è stato possibile evidenziare il ruolo dell'*institutional entrepreneur* nel trovare un bilanciamento tra logiche contrastanti e riallineare gli obiettivi dei diversi stakeholder a livello organizzativo (Jackall, 1988; Stål, 2011; Tetlock, 1985). Il meccanismo di riconciliazione si manifesta come un processo distribuito, in cui ogni attore contribuisce a tradurre esigenze istituzionali e operative in azioni concrete. Questo equilibrio si realizza attraverso pratiche collaborative, negoziazioni costanti e il supporto di una cultura organizzativa che valorizza la fiducia e l'impegno reciproco. Il direttore scientifico funge da catalizzatore, ma il successo dipende dalla capacità di tutti i livelli di bilanciare esigenze conflittuali, mantenendo al centro gli obiettivi strategici dell'istituto. In questo modo, la riconciliazione non è un semplice compromesso, ma un meccanismo attivo per garantire sostenibilità, innovazione e qualità della ricerca in un contesto complesso come quello degli IRCCS.

L'attivazione di una logica di comunità ha inoltre generato interessanti conseguenze sulla cultura della ricerca presente nell'organizzazione.

La letteratura evidenzia come il conflitto tra logiche istituzionali coesistenti all'interno di un'organizzazione porti necessariamente alla prevalenza di una sull'altra (van Gestel e Hillebrand, 2011) e la coesistenza non sia sostenibile a lungo termine (Reay e Hinings, 2005). Tuttavia, diversi studi hanno dimostrato che la riconciliazione è possibile a certe condizioni che necessariamente escludono il processo di *decoupling* (Bromley e Powell, 2012; Conrath-Hargreaves e Wüstemann, 2019; Meyer e Rowan, 1977). Il decoupling si verifica quando una logica è formalmente accettata per ragioni di conformità o reputazione ma non viene effettivamente applicata, mentre un'altra domina nelle pratiche operative. La riconciliazione richiede che entrambe le logiche siano integrate nelle pratiche organizzative attraverso negoziazione e adattamento. Questo processo permette di bilanciare esigenze diverse, evitando che una logica prevalga sull'altra, creando così un equilibrio funzionale che rispecchi la complessità istituzionale.

Lo studio ha individuato aspetti specifici del processo di ricerca e ha approfondito le fonti di tensione e di conflitto in questi ambiti, che sono state mappate all'interno del *framework* teorico delle logiche istituzionali.

4.2. Processo di ricerca e influenza delle logiche istituzionali

Le tre logiche istituzionali che si manifestano nel processo di ricerca all'interno dell'IRCCS sono la professionale, l'aziendalistica e la comunitaria (Brantnell e Baraldi, 2020; Dunn e Jones, 2010; Goodrick e Reay, 2011; Jia *et al.*, 2019). Le prime due sono in conflitto in quanto impongono pressioni contrastanti che gli attori organizzativi devono gestire per poter raggiungere i propri obiettivi (Greenwood *et al.*, 2011). Nella seguente tabella sono riportate le descrizioni degli idealtipi di riferimento.

*Tab. 1 - Dimensioni delle logiche istituzionali che agiscono nel processo di ricerca**

	Professionale	Aziendale	Comunità
Metafora fondamentale	ricerca rilevante nel campo del ricercatore	massimizzazione della produzione di ricerca	scambio reciproco
Fonti di Legittimità	paradigma di ricerca	allineamento con il contesto di valutazione internazionale, decreti ministeriali, rapporti di produttività, decisioni gestionali, gerarchia	fiducia, reciprocità
Fonti di Autorità	comunità scientifica	Ministero della Salute, direttore scientifico, ufficio scientifico	valori comuni
Base di attenzione	stato avanzamento della carriera del ricercatore	aumento dell'indice di impatto degli IRCCS	investimento personale nell'IRCCS
Base strategica	avanzare la conoscenza accademica e nuovi trattamenti	reinvestire i fondi acquisiti nella ricerca socialmente rilevante sfruttando l'interdisciplinarietà	produrre «pubblicazioni di sopravvivenza»

*Elaborazione propria dell'autore dalle dimensioni dei tipi ideali di logica di Thornton *et al.* (2012)

Nell'idealtipo della logica professionale, le pratiche sono definite e orientate dalla mentalità scientifica (Merton, 1973b; Mitroff, 1974) che induce a produrre ricerche di impatto sociale e a progredire nella carriera di ricercatore. Secondo questo idealtipo, i ricercatori hanno autonomia sul contenuto e sull'organizzazione della loro ricerca poiché il loro obiettivo principale è il progresso della conoscenza nel loro rispettivo campo e lo

sviluppo della loro carriera (Fumasoli *et al.*, 2015). La produzione e la pubblicazione di articoli di grande impatto sono essenziali per lo sviluppo della carriera di un ricercatore, come definito dagli standard internazionali, che associano il prestigio scientifico alle pubblicazioni su riviste ad alto impatto (Tennant *et al.*, 2017). Secondo questo idealtipo di logica, i ricercatori hanno la libertà di perseguire i propri interessi professionali, inclusa la partecipazione a borse di studio e la collaborazione individuale con gruppi di ricerca esterni (Agre, 2005).

Nella logica aziendale, le attività dei ricercatori sono strutturate all'interno di schemi stabiliti dall'organizzazione e supervisionati dal *top management* (Canhilal *et al.*, 2016; Grossi *et al.*, 2020; K.-M. Kallio *et al.*, 2017). All'interno dell'IRCCS i ricercatori sono visti come dipendenti e il contenuto e l'organizzazione del loro lavoro sono guidati dai *top manager*. Inoltre, i dirigenti senior, che potrebbero non appartenere alla stessa professione, potrebbero avere obiettivi non in linea con quelli dei ricercatori. Uno degli scopi principali dell'alta dirigenza è quello di tradurre i decreti e gli obiettivi ministeriali di valutazione in obiettivi chiari per i ricercatori a livello individuale. Il Ministero stabilisce gli standard per la qualità della ricerca e supervisiona il processo di valutazione (Timmermans e Oh, 2010) allo scopo di allocare i fondi in base alle prestazioni dell'IRCCS. Questi standard di qualità, stabiliti dal Ministero, sono declinati a livello organizzativo e il *top management* si assume la responsabilità di tradurre queste direttive in obiettivi individuali grazie all'uso strategico del linguaggio e della retorica (Jackall, 1988). La logica aziendale si manifesta quindi attraverso linee guida stabilite all'interno dell'organizzazione e si pone come alternativa alla logica professionale, riferimento naturale dei ricercatori. Queste linee guida consentono ai dipendenti di aderire agli obiettivi definiti dall'IRCCS e alla sua logica sottostante (Tetlock, 1985). Secondo questo tipo di logica ideale, le pratiche dei ricercatori si concentrano sulla produzione, sul valore e/o sull'utilità della ricerca con l'obiettivo primario di ottenere risultati finanziari (Baffy *et al.*, 2020) per consentire il reinvestimento dei fondi nella ricerca ad alto impatto sociale.

Infine, il tipo ideale di logica comunitaria evidenzia il *fil rouge* che unisce un gruppo di individui che si impegnano per gli stessi valori attraverso un forte senso di reciprocità e fiducia (Sperling, 2000). In questo caso, il contenuto e l'organizzazione del lavoro dei ricercatori sono definiti dal loro investimento personale per raggiungere un obiettivo comune a beneficio dell'intero gruppo (Thornton e Ocasio, 2012). Secondo questo idealtipo, i ricercatori tendono a riportare le loro pratiche di ricerca a vantaggio dell'organizzazione (in questo caso, l'IRCCS). Attraverso lo scambio reciproco (Mauss, 1966), anche fornendo risorse e strutture, l'alta dirigenza

allinea gli obiettivi definiti dal Ministero per l'IRCCS con le priorità dei ricercatori, attraverso un approccio *top-down* orientato alla comunità dei ricercatori che agisce come una logica di riconciliazione (si veda Jia *et al.* (2019) per l'approccio *bottom-up*). Questa logica è sostenibile nel tempo e consente all'IRCCS di raggiungere gli obiettivi prefissati e di prosperare reinvestendo in ricerca i fondi ricevuti. Questo meccanismo dissolve quindi il potenziale distruttivo del conflitto tra logiche divergenti e riconduce gli interessi e l'impegno a un obiettivo comune.

Lo studio ha quindi analizzato, sulla base delle narrazioni degli intervistati, le situazioni conflittuali che si manifestano nelle diverse circostanze del processo di ricerca, in particolare nell'organizzazione sociale della ricerca e infine nella scelta delle riviste per la disseminazione.

Nella definizione di nuovi progetti di ricerca si sovrappongono logiche aziendali, che si manifestano sottoforma di obiettivi di massimizzazione della *produzione* della ricerca, e logiche professionali, che sollecitano invece attenzione alla *rilevanza* della ricerca negli ambiti di interesse per i ricercatori in un quadro in cui l'innovazione e l'impatto sociale fanno perno sull'interdisciplinarità. Il *top management* ha la responsabilità di definire gli obiettivi precisi che i ricercatori devono perseguire per massimizzare la produttività dell'IRCCS nelle aree di accreditamento ai fini della valutazione ministeriale, allo scopo di ricevere fondi da reinvestire nella ricerca ad elevato impatto sociale. Si innesca così un meccanismo attraverso cui il perseguimento degli obiettivi stabiliti dai vertici sanitari dell'IRCCS diventa condizione per realizzare innovazione e ricerca ad alto impatto sociale.

Ma ciò non è realizzato con coercizione o accettato dai ricercatori con indifferenza: al contrario, l'allineamento tra interdisciplinarità della ricerca, espressione della diversità degli ambiti scientifici a cui afferiscono i ricercatori, e gli obiettivi organizzativi, fondamentali per assicurare la continuità dell'istituto, si fonda sulla logica di comunità. Quest'ultima è alimentata dal *top management*, in grado di riconciliare le due precedenti logiche grazie al modello di reciprocità attuato proprio dal vertice, che assicura le risorse necessarie per il perseguimento degli interessi specifici di ciascun ricercatore a fronte di un risultato di produttività raggiunto con il contributo di tutto lo staff.

In altre parole, la logica di riconciliazione fornisce ai collaboratori la libertà intellettuale e le risorse necessarie; al contempo mobilita i ricercatori al perseguimento dell'interesse dell'organizzazione (Levitt e Dubner, 2005; Sperling, 2000).

L'organizzazione sociale della ricerca pone dunque in conflitto le aspirazioni di carriera dei singoli ricercatori (sostenute da una logica professionale) e gli obiettivi di produttività (espressione della logica *corporativa*).

te). Anche nell'organizzazione sociale della ricerca la logica dominante è quella aziendale, legata all'ottenimento di risorse fondamentali per la sopravvivenza dell'istituto e per la continuità della ricerca. Essa riesce ad affermarsi proprio grazie alla logica di comunità, radicata in consuetudini preesistenti che permettono di allineare l'attenzione e gli interessi di ricercatori, in particolare clinici e universitari, con quelli dell'istituto, attraverso un meccanismo di scambio reciproco. Da un lato tale scambio comprende la disponibilità delle risorse per i ricercatori, dall'altro il prestigio dell'istituto e l'accesso a progetti differenti, altrimenti inaccessibili, grazie all'affiliazione di alcuni ricercatori all'università.

La logica della comunità può agire in modo più o meno evidente, a seconda della divergenza tra le istanze espresse dalle logiche aziendale e professionale. Ove queste non si allineino naturalmente, la logica della comunità emerge attraverso un approccio *bottom-up* (Jia *et al.*, 2019), come nel caso in cui il *top management* preme sui ricercatori per aumentare il loro tasso di pubblicazione, per soddisfare alcuni requisiti del Ministero, a costo di abbassare il livello della ricerca. La logica riconciliante viene attivata dai ricercatori (da cui l'approccio *bottom-up*) nell'ambito di un meccanismo di reciprocità che assicura un beneficio (il flusso di risorse per finanziare le loro ricerche) a fronte del loro sacrificio (l'imposizione di un tasso di pubblicazioni elevato).

Il *top management* esercita anche un potere coercitivo nella selezione delle riviste su cui pubblicare, in considerazione dei criteri che massimizzano la produttività dell'organizzazione e che permettono di conseguire una valutazione più elevata. Ma la leva della coercizione non è quella che induce i ricercatori ad allineare i propri interessi a quelli dell'IRCCS: ciò che spinge i ricercatori a perseguire l'interesse dell'organizzazione è proprio la logica della comunità, alimentata dalla cultura della ricerca sostenuta dal *top management*. Attraverso la cultura della ricerca, condivisa da tutti i membri dell'organizzazione, ciascun ricercatore si riconosce negli scopi generali dell'IRCCS e si sente parte della comunità di ricerca, nutre fiducia e gratitudine nei confronti dell'organizzazione, che agisce a beneficio dei ricercatori attraverso un meccanismo di reciprocità.

Se, da un lato, il meccanismo di valutazione imposto dal Ministero agli IRCCS per il finanziamento della ricerca richiede il miglioramento del proprio *impact factor* totale attraverso il numero di pubblicazioni sulle riviste internazionali, dall'altro questo sistema genera uno scollamento tra logica professionale e aziendale proprio perché interferisce con il modello di produzione e disseminazione della ricerca scientifica. Il disallineamento, se non gestito in modo adeguato, può vanificare gli scopi stessi della valutazione, improntata al miglioramento della produttività: infatti diminuendo la capa-

cità di innovare e di produrre un impatto sociale rilevante, il contrasto tra le logiche può arrivare a minare la ragione stessa dell'esistenza dell'IRCSS, ovvero la generazione di nuova conoscenza scientifica ad alto valore sociale.

Dunque la logica della comunità, basata su una cultura forte della ricerca e sulla reciprocità, mitiga gli effetti della coercizione realizzando unità di intenti. Nutrita dal *top management* attraverso un meccanismo di reciprocità (Mauss, 1966) che genera fiducia reciproca tra ricercatori e istituto, tale logica offre una mediazione tra logiche in conflitto e potenzialmente in grado di vanificare gli scopi stessi della valutazione della ricerca scientifica e agisce pertanto come meccanismo di riconciliazione.

5. Conclusioni

Obiettivo del capitolo non è proporre un superamento dell'attuale sistema, ma favorire la consapevolezza delle dimensioni che interagiscono nella valutazione della ricerca scientifica per riflettere sulle prospettive di evoluzione del sistema stesso, ovvero sulle conseguenze attese e inattese che si manifestano anche nel lungo periodo. Da un lato, la valutazione deve essere in grado di promuovere la qualità intrinseca della ricerca e l'innovazione, in una prospettiva di rilevanza sociale, dall'altro deve porre le basi per un sistema di assegnazione delle risorse responsabilizzante. Le criticità che il modello di valutazione si trova ad affrontare riguardano l'inclusione di prospettive *non mainstream* nello sviluppo della ricerca accademica, possibile stimolo all'innovazione; l'interdisciplinarità, che permette contaminazioni benefiche di idee, ma richiede un'attenzione specifica alla coerenza delle scelte ontologiche ed epistemologiche. Un ulteriore punto di attenzione è la sostenibilità del sistema di valutazione, sia in termini di costi, sia di disponibilità di risorse cognitive, nonché il rischio di fenomeni inflattivi indotti dai sistemi di carriera basati sulla produttività scientifica, che sovraccaricano i meccanismi di pubblicazione dei lavori accademici. Lo studio delle logiche istituzionali ha dimostrato come la valutazione della ricerca introduca nel sistema logiche estranee alla scienza in senso stretto, quali quelle di mercato, come nel caso del settore dell'editoria scientifica, oppure di tipo aziendale, nell'ambito della ricerca medica.

La riflessione che ne scaturisce si rivolge non solo agli accademici, coinvolti nei ruoli di *editor* e *reviewer*, che rappresentano sempre più il perno di un sistema nel quale intervengono logiche di funzionamento non appartenenti alla scienza, ma riguarda anche (e soprattutto) i decisori politici. L'associazione tra prestigio accademico e JIF delle riviste in cui sono pubblicati i lavori dei ricercatori rappresenta il principio fondante di un

sistema di legittimazione della conoscenza che si autoalimenta per fare fronte alle aspettative della società della conoscenza, nell'era della transizione all'università di massa, caratterizzata da un incremento nel numero di accademici collocati soprattutto nelle prime fasi della carriera. Questo sistema, organizzato e gestito da società editoriali internazionali, è reso possibile dalla disponibilità dei ricercatori a dedicare parte delle proprie energie alla *peer review* per valutare la qualità della ricerca e far crescere la conoscenza scientifica. Ma per funzionare correttamente, minimizzando distorsioni ed effetti collaterali, e per sostenere la produttività scientifica, il sistema richiede maggiore consapevolezza e crescente disponibilità di competenze. E questo riguarda tutti i campi della ricerca scientifica, massimamente la ricerca in ambito sanitario caratterizzata per sua natura da un elevato impatto sociale.

Nella revisione tra pari, che è alla base della valutazione di un manoscritto, la qualità intrinseca del lavoro di ricerca viene valutata dai revisori, sulla base della propria posizione epistemologica, e sancita attraverso la decisione dell'*editor* di pubblicare il manoscritto. In questa circostanza, di fatto l'*editor* contribuisce a produrre una definizione emergente di ‘qualità’ della ricerca. Il ruolo dell'*editor* si è dunque trasformato in custode del futuro della disciplina e *dominus* della carriera degli accademici, in quanto la progressione è legata alla capacità di pubblicare su giornali a elevato IF.

Lo studio condotto sui due campioni di riviste (uno nell'ambito di una disciplina bibliometrica e l'altro in una disciplina non bibliometrica) ha evidenziato in entrambi i casi l'emergere di una logica di mercato, accanto a quella professionale apportata dall'*editor*, che si è sviluppata grazie all'associazione tra *impact factor* della rivista e prestigio accademico e che influisce nel processo di valutazione. La logica di mercato si è introdotta nel contesto accademico modificando le prospettive e le priorità dei ricercatori.

Come già evidenziato dalla letteratura, nello studio si osserva come la standardizzazione del processo di *peer review* nell'ambito dell'editoria scientifica internazionale abbia orientato il *focus* verso caratteristiche estrinseche della qualità, al di là del contenuto della ricerca (argomenti di tendenza, affiliazione a conferenze, perdurante rilevanza di un articolo ecc.).

Ne consegue che il prestigio accademico, che si basa sulle pubblicazioni, è condizionato anche da dimensioni di qualità estranee al contenuto stesso della ricerca, legittimate dalle riviste internazionali. La rilevanza della logica di mercato, privilegiando certi paradigmi di ricerca, può influenzare la percezione dell'accettabilità di alcune posizioni ontologiche ed epistemologiche nella produzione di ricerca di qualità.

L'esempio dell'ASN dimostra come il sistema italiano, similmente ad altri contesti, abbia integrato e legittimato logiche tipiche del mercato

editoriale accademico internazionale che non necessariamente hanno origine nei principi della scienza. L'impatto che queste decisioni hanno sulle carriere invoca la necessità di generare consapevolezza di come la progressione in ambito accademico possa essere influenzata da logiche diverse da quelle proprie delle discipline scientifiche.

L'intento commerciale e la standardizzazione del processo di *peer review* influenzano dunque la multidimensionalità del concetto di qualità della ricerca e le priorità dei ricercatori, sia rispetto alle traiettorie di carriera, sia rispetto all'accesso ai fondi per la ricerca.

Le iniziative imprenditoriali all'interno del mercato delle pubblicazioni scientifiche sono alla base della valutazione dei lavori di ricerca e dello sviluppo della carriera degli accademici. Questa osservazione diventa ancor più rilevante nei contesti in cui agli esiti del processo di pubblicazione è legata l'allocazione di risorse per il finanziamento delle ricerche scientifiche.

La coesistenza di tre logiche istituzionali (aziendale, professionale e comunitaria) nel processo di produzione della ricerca di un IRCCS, dove le prime due definiscono il conflitto, mentre la terza è la leva attraverso cui l'imprenditore istituzionale opera una riconciliazione, dimostra la loro interazione nelle tre fasi della ricerca: concepimento del progetto, organizzazione sociale della ricerca e disseminazione attraverso le riviste scientifiche. Al contempo evidenzia il ruolo della logica di riconciliazione che riallinea le altre due mitigando gli effetti potenzialmente distruttivi per l'organizzazione, rispetto alla quale il *top management* gioca un ruolo centrale.

La massimizzazione della produttività della ricerca in termini di JIF totale dell'IRCCS, richiesta dal Ministero, è condizione per ottenere una valutazione positiva e quindi l'erogazione di fondi che possano alimentare il circuito virtuoso della ricerca, attraverso il reinvestimento nei progetti dei ricercatori. Ma la massimizzazione della produttività non sempre va nella stessa direzione della logica professionale del ricercatore, perché talvolta impone di intraprendere percorsi di ricerca che non coincidono con interessi e obiettivi dei ricercatori: ecco allora che la logica della comunità, costruita su un meccanismo di reciprocità, viene attivata dal *top management* per permettere ai ricercatori un bilanciamento tra risorse disponibili e interessi personali di ricerca basato sulla libertà intellettuale. Questo modello favorisce un senso di fiducia e di gratitudine che alimenta la cultura della ricerca all'interno dell'IRCCS, stemperando le forme di coercizione praticate dal *top management*. Il motore che sostiene l'organizzazione e il processo di produzione di nuova conoscenza è quindi un meccanismo di scambio di benefici che dà vita alla logica della comunità da un lato e, generando fiducia e gratitudine reciproche, alimenta la cultura della ricerca, dall'altro lato, che rende sostenibile nel tempo il processo di ricerca.

Il circuito virtuoso è il frutto di un'attenta navigazione del sistema da parte del *top management* che, per poter garantire la continuità della ricerca in un ambito vitale per il benessere della società, ha trovato e applicato con successo una forma di riconciliazione tra due logiche in conflitto, di cui una di origine totalmente estranea al mondo della ricerca e in contrasto con il modello di produzione e disseminazione della ricerca scientifica.

Il punto centrale, che merita di essere sottolineato, è che il disallineamento tra logica professionale, riferita all'ambito medico, e logica aziendale, se non gestito in modo adeguato, può vanificare gli scopi stessi della valutazione imposta dal Ministero, improntati al miglioramento della produttività, diminuendo la capacità stessa di innovare e di produrre un impatto sociale rilevante e conseguentemente minando la ragione stessa dell'esistenza dell'IRCCS.

Questo studio non intende criticare il sistema esistente per proporne il superamento. Anzi: ne mette in rilievo la ricerca di oggettività e l'intento di responsabilizzare i ricercatori per l'uso delle risorse. Inoltre, si propone di richiamare l'attenzione sugli effetti di logiche estranee alla scienza ma vitali per la sostenibilità del sistema stesso. Le potenziali distorsioni che ne derivano richiedono l'assunzione di nuove responsabilità nella mitigazione degli effetti per favorire lo sviluppo della ricerca scientifica.

7. La risposta delle università italiane all'introduzione della Valutazione della Qualità della Ricerca, VQR 2004-2010

di *Giovanni Abramo, Ciriaco Andrea D'Angelo, Flavia Di Costa*

1. Introduzione

Mai come nel drammatico periodo della pandemia di Covid-19 si è diffusa a livello collettivo una forte consapevolezza dell'importanza della ricerca scientifica. Da decenni i governi più lungimiranti avevano già collocato fra le priorità assolute dell'agenda politica il potenziamento delle infrastrutture tecnico-scientifiche congiuntamente al rafforzamento dell'istruzione superiore.

Un numero crescente di paesi ha scelto di rafforzare i meccanismi competitivi nel settore della ricerca, attraverso un accesso alle risorse basato sul merito. Molti paesi hanno introdotto sistemi di finanziamento della ricerca basati sulla performance (Performance Based Research Funding Systems-PBRFS), volti a stimolare il miglioramento continuo dell'efficacia e dell'efficienza dell'attività di ricerca (Hicks, 2012; Lewis, 2013; Woelert, 2015; Geuna e Piolatto, 2016). Tra i 28 Paesi dell'UE (compreso il Regno Unito), 16 hanno introdotto una qualche forma di PBRFS di cui sono stati analizzati i punti di forza e di debolezza (Zacharewicz *et al.*, 2019).

Gli esercizi nazionali di valutazione della ricerca sono generalmente associati ai PBRFS, poiché in base ai loro risultati vengono assegnate risorse alle organizzazioni pubbliche di ricerca. Oltre a stimolare una maggiore produttività e a convogliare le risorse verso le organizzazioni più produttive, i risultati di questi esercizi consentono analisi strategiche, in grado poi di orientare le politiche nazionali di ricerca e di ridurre l'asimmetria informativa presente nel “mercato della conoscenza”, tra l'offerta (organizzazioni di ricerca) e la domanda (studenti, settore industriale, ecc.). Non da ultimo, gli esercizi di valutazione servono anche a dimostrare ai contribuenti che gli investimenti nella ricerca sono efficaci e producono benefici per la collettività.

Per le organizzazioni di ricerca, l'incentivo a migliorare non deriva solo dalla leva finanziaria associata agli esercizi di valutazione (OECD, 2010; Woelert e McKenzie, 2018), ma anche dalla reputazione (prestigio) che deriva dal buon posizionamento nelle relative classifiche. Il ritorno di immagine che ne consegue costituisce un fattore attrattivo per ricercatori di talento, professori, studenti, finanziamenti privati, donazioni, ecc.

Non tutte le organizzazioni soggette a valutazione reagiscono allo stesso modo agli incentivi dei PBRFS. La risposta dipende da diversi fattori, primo fra tutti l'adozione dei medesimi incentivi nazionali all'interno delle singole organizzazioni di ricerca, ma anche il senso di identificazione e di appartenenza dei ricercatori alla propria organizzazione. Diversi studi hanno cercato di valutare l'impatto degli esercizi di valutazione nazionale sul comportamento e sulle performance di ricerca di individui e organizzazioni. Il presente lavoro si inserisce in questo filone di ricerca. In particolare, ci proponiamo di verificare se e in che misura le università italiane abbiano migliorato la loro performance di ricerca relativa (cioè la posizione in classifica), come risposta al PBRFS basato sui risultati dell'esercizio di valutazione della ricerca VQR 2004-2010, lanciato alla fine del 2011.

È molto diverso valutare la variazione della performance relativa rispetto alla variazione della performance assoluta: un'università può, ad esempio, migliorare la prima anche se la seconda diminuisce (per esempio nel caso in cui le altre università siano peggiorate in media in misura maggiore).

Al contrario della valutazione della performance relativa, l'individuazione di un effetto causale del PBRFS sulla performance assoluta della ricerca è un compito arduo, a causa della difficoltà di controllare tutti i possibili fattori che concorrono a determinarlo. Ad esempio, se non tenuto in debito conto, il costante aumento del numero di riviste, indicizzate dai repertori bibliografici alla base delle valutazioni bibliometriche, può falsare i risultati analitici. Questo fattore è in gran parte irrilevante nel caso della valutazione della variazione relativa di performance, in quanto qualsiasi cambiamento contestuale ex-ante/ex-post, con un potenziale impatto sulla performance, sarebbe problematico solo nel caso in cui avesse effetti sistematici pro/contro determinati gruppi di università. Un altro fattore è la possibile variazione delle risorse destinate alla ricerca nelle singole università. Questo fattore è praticamente impossibile da tracciare, soprattutto a livello individuale o di gruppi di ricerca, come necessario per una corretta valutazione della performance complessiva.

Nonostante una serie di assunti e limiti tipici degli studi bibliometrici, il nostro lavoro si propone di migliorare la letteratura esistente in

diversi modi: i) adottiamo un indicatore di performance derivato dalla teoria microeconomica della produzione, che integra le dimensioni quantitative e qualitative dell'output, tiene conto dei contributi individuali all'output e, soprattutto, inserisce variabili di controllo per gli input, laddove tali dati siano disponibili; ii) per valutare la performance, misureremo tutta la produzione scientifica indicizzata in *Web of Science* (*WoS*) *core collection*, e non solo la piccola percentuale di prodotti presentati dalle università alla VQR; e iii) sfruttiamo una classificazione granulare dei settori scientifici dei professori italiani, che consente di tenere conto della diversa intensità di pubblicazione dei vari settori evitando quindi di fare confronti totalmente fuorvianti (Abramo, Cicero e D'Angelo 2013a); iv) valutiamo la performance, prima e dopo l'introduzione del PBRFS, con lo stesso indicatore.

Nel valutare la risposta dei singoli atenei al PBRFS siamo anche in grado di misurare se la dispersione della performance tra gli atenei è cambiata e se si è ridotto il divario regionale fra gli atenei, un tema delicato in diversi Paesi e molto sentito in Italia, a causa della presenza del “divario nord-sud” (Abramo, D'Angelo e Rosati, 2016b; Checchi *et al.*, 2020; Fondazione RES, 2015; Grisorio e Prota, 2020) che secondo l'attuale dibattito politico potrebbe anche accentuarsi con l'entrata a regime della legge sull'autonomia differenziata, recentemente approvata in Parlamento, almeno stando alle argomentazioni di chi ne propone un referendum abrogativo. Le università, infatti, svolgono un importante ruolo a livello locale, a sostegno della competitività industriale, dello sviluppo socio-economico e della mobilità sociale. In questo contesto le istituzioni accademiche possono infatti contribuire a ridurre le differenze, o viceversa ad aumentarle, sia in termini macroeconomici che sociali. Un'offerta universitaria equilibrata in tutte le regioni garantirebbe quindi pari opportunità di istruzione e formazione, anche per i meno abbienti, che non possono sostenere i maggiori costi di un'istruzione “lontano da casa”. A causa dell'eccezionale divario della performance di ricerca, risultante dalla prima VQR a favore delle università settentrionali, il PBRF diminuendo le risorse a disposizione delle università meridionali, potrebbe infatti causare un ulteriore aumento delle differenze nord-sud nello sviluppo macroeconomico e sociale (Fondazione RES, 2015; Abramo, Cicero e D'Angelo, 2011b).

Lo studio è articolato come segue. Nel paragrafo successivo passiamo in rassegna la letteratura sull'argomento. Nel paragrafo 3 illustriamo la metodologia e il campo di analisi, e nel 4 presentiamo i risultati. Il paragrafo 5 conclude il lavoro con osservazioni e raccomandazioni di politica della ricerca.

2. Analisi della letteratura

La valutazione periodica della ricerca influenza la motivazione, la performance, il comportamento di individui, gruppi e organizzazioni, condizionando direttamente o indirettamente molti aspetti della vita professionale dei ricercatori. Gli esercizi di valutazione incidono sulle politiche di reclutamento e avanzamento di carriera, sulla libertà accademica, nonché sul bilanciamento fra insegnamento e attività di ricerca, e influenzano il trasferimento tecnologico e la localizzazione dei luoghi di potere. Alcuni studiosi dei PBRFS affermano che non ci sono ancora prove sufficienti per valutare gli effetti di questi sistemi e che la comprensione della loro corretta progettazione e attuazione è ancora incompleta (de Boer *et al.*, 2015). La letteratura esistente sull'impatto dei PBRFS può essere classificata in tre filoni di ricerca che si sovrappongono. Il primo riguarda l'impatto sulla gestione strategica e l'organizzazione delle istituzioni di ricerca mentre il secondo analizza gli effetti sul comportamento individuale, in particolare gli effetti perversi. Il nostro studio contribuisce a una terza linea di ricerca, che riguarda gli impatti macro o microeconomici dei PBRFS in termini di i) aumento delle performance di ricerca, ii) convergenza o divergenza di performance all'interno delle istituzioni e tra di esse; iii) convergenza o divergenza fra le regioni.

Per gli studiosi impegnati nella prima linea di ricerca, oggetto di indagine è il comportamento delle organizzazioni in risposta all'introduzione dei PBRFS. Essi indagano in che misura gli incentivi a livello di sistema possano influenzare le pratiche gestionali locali e come queste possano a loro volta migliorare l'efficacia organizzativa, l'efficienza e le pratiche di gestione delle risorse basate sul merito (Espeland e Sauder, 2007; Sauder e Espeland, 2009). Diversi studi hanno analizzato il dispiego degli schemi di incentivazione incorporati nei PBRFS nazionali all'interno delle singole organizzazioni di ricerca. In Australia, è emerso che la maggior parte delle università tende a replicare internamente i principali indicatori di performance nazionali (Woelert e McKenzie, 2018). In Norvegia, sono coesistenti prassi contrastanti: se da un lato è stata riscontrata una sostanziale replica in alcune istituzioni di schemi standardizzati a livello nazionale, dall'altro è stata osservata un'ampia variabilità di approcci tra organizzazioni, discipline e dipartimenti (Aagaard, 2015).

Altri studiosi si sono concentrati sugli effetti dei PBRFS sul comportamento dei singoli, effetti significativi quando le organizzazioni distribuiscono gli incentivi nazionali a livello individuale (Moher *et al.*, 2018). La maggior parte degli studi ha analizzato gli effetti indesiderati (de Rijcke *et al.*, 2016), tra cui: i) la cattiva condotta scientifica, come la moltiplicazio-

ne di articoli simili, il plagio, l'autoplagio, le autocitazioni inappropriate, la frode scientifica (Hazelkorn, 2010; Edwards e Roy, 2017; Seeber *et al.*, 2019); ii) il disincentivo alla diversificazione della ricerca, alla ricerca interdisciplinare e innovativa (Abramo, D'Angelo e Di Costa, 2019a; Hicks, 2012; Rafols *et al.*, 2012; Wilsdon, 2015); e iii) il trade-off tra ricerca e insegnamento (De Philippis, 2015; Enders *et al.*, 2015).

Il terzo filone di ricerca riguarda l'impatto del PBRFS sulla performance complessiva delle istituzioni oggetto di valutazione. Il più delle volte, l'impatto sulla performance viene analizzato secondo due dimensioni distinte: numero di articoli e impatto medio per articolo. I risultati sono piuttosto eterogenei e differenziati, non solo tra i vari paesi, ma anche in studi riferiti allo stesso paese (Adams e Gurney, 2010; Wang e Hicks, 2013). In alcuni casi, i cambiamenti sono analizzati secondo i criteri di performance specificamente incorporati negli esercizi di valutazione della qualità, confrontando gli *score* di performance prima e dopo (ex-ante/ex-post) (Buckle *et al.*, 2020; Checchi *et al.*, 2020).

Sono stati condotti diversi studi per rispondere alle esigenze dei governi di valutare l'impatto dei rispettivi PBRFS sulla produttività di ricerca nazionale. Alcuni di essi hanno valutato le variazioni ex ante-ex post del numero di pubblicazioni e delle citazioni medie per articolo (Regno Unito: McNay, 2015; Australia: Butler, 2003b, 2003a; van den Besselaar *et al.*, 2017; Norvegia: Schneider *et al.*, 2016).

Una recente indagine su 18 Paesi che hanno adottato una qualche forma di PBRFS ha evidenziato un effetto positivo sull'impatto medio per articolo, un effetto positivo ma temporaneo sul numero di pubblicazioni e percentuali quasi invariate di articoli pubblicati in riviste ad alto *impact factor*, IF (Checchi *et al.*, 2019). Cattaneo *et al.* (2016) hanno analizzato la variazione delle pubblicazioni per docente in seguito all'introduzione del primo esercizio di valutazione della ricerca italiana nel 2003. Altri studi hanno valutato le variazioni di performance delle istituzioni derivanti dai criteri di valutazione quantitativi degli esercizi nazionali (Nuova Zelanda: Buckle e Creedy, 2019; Italia: Checchi *et al.*, 2020).

Oltre l'inevitabile difficoltà di controllare tutti i fattori confondenti, il difetto principale degli studi citati risiede nella selezione dell'indicatore di performance. Nessuno dei due indicatori – il numero di pubblicazioni o l'impatto medio per articolo – sono validi indicatori di performance se considerati da soli. Quando entrambi gli indicatori vengono misurati, ma le variazioni sono divergenti, la questione della performance rimane irrisolta. Anche la mancata adozione del *fractional counting* (contributo individuale) delle pubblicazioni con più autori conduce a errori (Waltman e van Eck, 2015). Inoltre, quando gli esercizi di valutazione si basano su un sottoinsieme del portafoglio prodotti dell'organizzazione di ricerca piuttosto che su

tutti, i valori di performance relativa risultano fortemente distorti (Abramo, Cicero e D'Angelo, 2013b), potendo dipendere tra l'altro dalla (in)capacità delle organizzazioni di selezionare i loro prodotti migliori (Abramo, D'Angelo e Di Costa, 2014a). A nostro avviso, non ha molto senso confrontare punteggi di performance così distorti, come fatto da Checchi *et al.* (2020) per il caso italiano. Inoltre, eventuali variazioni di performance potrebbero essere dovute a una migliore o peggiore selezione dei prodotti presentati.

Nell'interpretare o confrontare i risultati di qualsiasi indagine relativa all'impatto dei PBRFS sulle organizzazioni di ricerca, occorre prestare particolare attenzione alle scelte fatte in termini di indicatori di performance e metodi di valutazione.

3. Dati e metodologia

Per verificare quali università italiane abbiano migliorato in termini quantitativi la loro performance relativa di ricerca in risposta alla VQR 2004-2010, adottiamo un'analisi ex-ante/ex-post. Poiché la VQR 2004-2010 è stata lanciata alla fine del 2011, misuriamo la performance bibliometrica media delle università nel quinquennio 2007-2011, assumendo il 2012 come anno di transizione, per poi confrontarla con la performance delle università nel periodo 2013-2017. Per evitare distorsioni nel confronto, adottiamo finestre citazionali identiche¹, contando le citazioni al 31/12/2013 per il primo quinquennio e al 31/12/2019 per il secondo.

3.1. Dati

I dati sui docenti di ciascuna università sono stati estratti dalla banca dati sul personale universitario italiano, gestita dal MUR². Nelle università italiane, tutti i professori sono inquadrati in un solo settore, denominato settore scientifico disciplinare (SSD), in tutto 370. I SSD sono raggruppati in discipline, denominate aree disciplinari universitarie (ADU), 14 in tutto.³ Per motivi di copertura dei repertori bibliografici, il campo di osservazione è limitato alle ADU (10 in tutto, contenenti 199 SSD) che l'ANVUR stessa definisce “bibliometriche” e ha valutato con indicatori bibliometrici.

1. Per finestra citazionale si intende il lasso di tempo che intercorre tra l'istante di pubblicazione di un lavoro scientifico e quello di rilevazione delle citazioni da esso ricevute.

2. <http://cercauniversita.cineca.it/php5/docenti/cerca.php>.

3. L'elenco completo è accessibile su <http://attiministeriali.miur.it/UserFiles/115.htm>.

Il dataset di analisi è costituito da due serie distinte di professori italiani, una per ogni quinquennio:

- per il periodo 2007-2011, 30.922 professori (ricercatori, associati e ordinari) con almeno 3 anni di servizio;
- per il periodo 2013-2017, 28.046 professori con almeno 3 anni di servizio.

La Tab. 1 indica la loro ripartizione per ADU. I due dataset sono abbastanza simili, tuttavia si nota una significativa riduzione del corpo docente nel secondo periodo (-2876 unità pari al 9% del totale) e una ridistribuzione dei professori su un numero maggiore di università (93 nel 2013-2017 contro le 85 nel 2007-2011). La contrazione ha interessato tutte le ADU (in particolare Medicina, -19%), ad eccezione di Ingegneria industriale e dell'informazione.

Tab. 1 - Dataset dell'analisi

ADU	2007-2011		2013-2017			
	SSD	Univ.	Prof.	Univ.	Prof.	
1 – Matematica e informatica	10	64	2.823	73	2.537	
2 – Fisica	8	64	1.931	69	1.783	
3 – Chimica	11	59	2.588	61	2.393	
4 – Scienze della Terra	12	47	936	51	863	
5 – Biologia	19	64	4.256	73	3.885	
6 – Medicina	50	62	8.735	67	7.102	
7 – Agraria e veterinaria	30	55	2.725	55	2.598	
8 – Ingegneria civile	9	53	1.340	59	1.286	
9 – Ing. industriale e dell'informazione	42	70	4.557	74	4.572	
10 – Psicologia	8	61	1.031	67	1.027	
<i>Totale</i>	<i>199</i>	<i>85</i>	<i>30.922</i>	<i>93</i>	<i>28.046</i>	

3.2. Misurare la produttività della ricerca

In linea con la teoria microeconomica della produzione, misuriamo le prestazioni dei professori in base alla loro produttività di ricerca. La produttività del lavoro, comunemente definita come l'output medio per unità di lavoro (a parità di altri fattori produttivi), rappresenta l'indicatore dell'efficienza di qualsiasi sistema produttivo. Poiché nei sistemi di ricerca le pubblicazioni (output) hanno valori diversi (impatto) e le risorse disponibili per la ricerca sono disuguali tra individui e organizzazioni, una definizione

operativa di produttività è quella data dal valore dell'output per euro speso nella ricerca.

Misurare la produttività di ricerca richiede l'adozione di una serie di assunti a causa della mancanza di dati. L'indicatore *FSS* (*fractional scientific strength*) è una *proxy* della produttività della ricerca. Una descrizione approfondita dell'indicatore *FSS* e della teoria di riferimento è contenuta in Abramo e D'Angelo (2014b).

Per misurare la produttività media annuale della ricerca di ogni professore, utilizziamo la seguente formula⁴:

$$FSS_P = \frac{1}{\left(\frac{w_r}{2} + k\right)} \cdot \frac{1}{t} \sum_{i=1}^N c_i f_i \quad [1]$$

dove:

w_r = stipendio medio annuo⁵;

k = capitale medio annuo disponibile per la ricerca;

t = numero di anni di lavoro nel periodo in osservazione;

N = numero di pubblicazioni nel periodo di osservazione;

c_i = combinazione ponderata di citazioni normalizzate per settore scientifico (WoS) ricevute dalla pubblicazione i e *IF* normalizzato per settore scientifico della rivista di pubblicazione⁶;

f_i = contributo dell'autore alla pubblicazione i .

Il contributo è pari all'inverso del numero di autori in quei settori in cui la prassi è quella di inserire gli autori in semplice ordine alfabetico, ma assume pesi diversi negli altri casi. Per Biologia, Ricerca biomedica e Medicina clinica, la prassi diffusa è che gli autori indichino i vari contributi alla ricerca pubblicata attraverso l'ordine dei nomi nella lista autori. Per queste discipline, quindi, attribuiamo pesi diversi a ciascun coautore in base alla sua posizione nella lista degli autori e al carattere della *co-authorship* (*intra-muros* o *extra-muros*)⁷.

4. L'assunto sotteso è che il lavoro e il capitale contribuiscano in egual misura alla produzione.

5. Dimezziamo il costo del lavoro, ipotizzando che il 50% del tempo dei professori sia destinato ad attività diverse dalla ricerca.

6. Le citazioni sono normalizzate alla media della distribuzione delle citazioni ricevute per tutte le pubblicazioni citate da WoS nello stesso anno e settore scientifico della pubblicazione i . L'*IF* si riferisce all'anno di pubblicazione ed è normalizzato alla media della distribuzione degli *IF* delle riviste nello stesso settore scientifico della pubblicazione i .

7. Se la pubblicazione è prodotta da una collaborazione esclusivamente intramurale (una sola affiliazione nell'elenco degli indirizzi), il 40% è attribuito al primo e all'ultimo autore e il restante 20% è suddiviso tra tutti gli altri autori. Al contrario, se l'elenco degli

Il dataset bibliometrico utilizzato per misurare la *FSS* è estratto dall’Osservatorio Italiano della Ricerca Pubblica (ORP), un database sviluppato e aggiornato dai presenti autori e derivato su licenza dalle basi dati di WoS. Applicando ai dati grezzi di WoS un complesso algoritmo di riconciliazione dell’affiliazione dell’autore e di disambiguazione della vera identità degli autori, in ORP ogni pubblicazione è attribuita ai professori universitari che l’hanno realizzata⁸.

I dati relativi ai costi salariali del personale di ricerca sono stati ricavati dalla banca dati DALIA⁹, gestita dal MUR. Gli stipendi dei professori in Italia dipendono solo dal rango accademico e dall’anzianità.

Per quanto riguarda il costo del capitale, k , ci basiamo su dati a livello di disciplina tratti da Abramo, Aksnes e D’Angelo (2020). La Tab. 2 riassume il costo del capitale (k), il costo totale dei fattori di produzione ($\frac{w_r}{2} + k$) e i fattori di normalizzazione del costo totale per rango accademico e disciplina. Il fattore di normalizzazione utilizzato è il valore minimo del costo totale, che si riscontra per i ricercatori di Psicologia (54.064 euro).

Di seguito, utilizziamo i fattori di normalizzazione dei costi totali per ottenere le misure di produttività.

Assumiamo che il capitale sia lo stesso, in entrambi i periodi, per i ricercatori e le istituzioni all’interno delle ADU, per la carenza di informazioni specifiche. Si può obiettare che le università più performanti hanno ricevuto più fondi nel secondo periodo, ma non sappiamo quale quota di questi fondi sia stata destinata alla ricerca e se i governi regionali abbiano compensato i minori finanziamenti nazionali alle loro università con basse performance.

indirizzi della pubblicazione mostra collaborazioni extramurali, il 30% è attribuito sia al primo che all’ultimo autore; il 15% sia al secondo che al penultimo autore; e il restante 10% è suddiviso tra tutti gli altri. I pesi sono stati assegnati seguendo i consigli di un *pool* di professori italiani di scienze della vita. I valori potrebbero essere modificati per adattarsi a pratiche diverse in altri contesti nazionali.

8. La media armonica della *precision* e *recall* (*F-measure*) delle *authorship* disambigue dall’algoritmo è di circa il 97% (margini di errore del 2%, intervallo di confidenza del 98%). Una descrizione approfondita dell’algoritmo si trova in D’Angelo, Giuffrida e Abramo (2011).

9. https://dalia.cineca.it/php4/inizio_access_cnvsu.php.

Tab. 2 - Costo medio annuo dei fattori produttivi, per rango accademico e area disciplinare, per i professori italiani

ADU	K	$\frac{w_r}{2} + k$			Fattore di normalizzazione dei costi totali		
		Ricercatore	Associato	Ordinario	Ricercatore	Associato	Ordinario
1	30.822	58.109	65.079	82.019	1.07	1.20	1.52
2	46.194	73.481	80.451	97.391	1.36	1.49	1.80
3	39.820	67.107	74.077	91.017	1.24	1.37	1.68
4	60.016	87.303	94.273	111.213	1.61	1.74	2.06
5	45.748	73.035	80.005	96.945	1.35	1.48	1.79
6	41.228	68.515	75.485	92.424	1.27	1.40	1.71
7	45.748	73.035	80.005	96.945	1.35	1.48	1.79
8	47.810	75.097	82.067	99.007	1.39	1.52	1.83
9	47.810	75.097	82.067	99.007	1.39	1.52	1.83
10	26.777	54.064	61.034	77.974	1.00	1.13	1.44

1 – Matematica e informatica; 2 – Fisica; 3 – Chimica; 4 – Scienze della Terra; 5 – Biologia; 6 – Medicina; 7 – Scienze agrarie e veterinarie; 8 – Ingegneria civile; 9 – Ingegneria industriale e dell’informazione; 10 – Psicologia.

Data la diversa intensità di pubblicazione nei vari settori di ricerca, il confronto delle performance delle università che operano in SSD eterogenei non può essere ottenuto semplicemente facendo la media delle performance individuali nei vari SSD. È necessaria una procedura che si articola in due fasi: prima occorre misurare la produttività dei singoli professori nel loro SSD (formula 1) e poi normalizzare la produttività individuale per la media del SSD. A livello aggregato di università, ADU e SSD, la produttività annuale *FSS* per l’unità aggregata è:

$$FSS = \frac{1}{RS} \sum_{j=1}^{RS} \frac{FSS_{Pj}}{\overline{FSS}_P} \quad [2]$$

dove:

RS = numero di professori dell’unità, nel periodo osservato;

FSS_{Pj} = produttività del professore *j* nell’unità;

\overline{FSS}_P = produttività media di tutti i professori italiani produttivi sotto osservazione nello stesso SSD del docente *j*.

La *FSS* a livello aggregato è quindi un indicatore di performance relativa. Ad esempio, un valore *FSS* pari a 1,10 significa che la performance dell'unità è del 10% superiore alla media.

A causa della mancanza di dati, l'analisi proposta non tiene conto di eventuali cambiamenti dei “fattori confondenti” della produttività della ricerca, quali la quantità di risorse disponibili per la ricerca, le condizioni di lavoro, ecc.

4. Risultati dell'analisi

Partiamo con i risultati del confronto tra le classifiche generali delle università nei due periodi. Poiché è probabile che per ogni università le variazioni di *ranking* siano diverse per le ADU e i SSD all'interno di ogni singola ADU, successivamente analizzeremo anche le differenze di *ranking* a livello di ADU e SSD. Questa analisi potrebbe servire al management per formulare le proprie strategie di investimento, reclutamento e competitive fino ai livelli dei singoli SSD.

La Fig. 1 mostra la classifica di produttività delle 60 università con almeno 30 professori valutabili in entrambi i quinquenni, in ordine decrescente di variazione del *ranking*¹⁰.

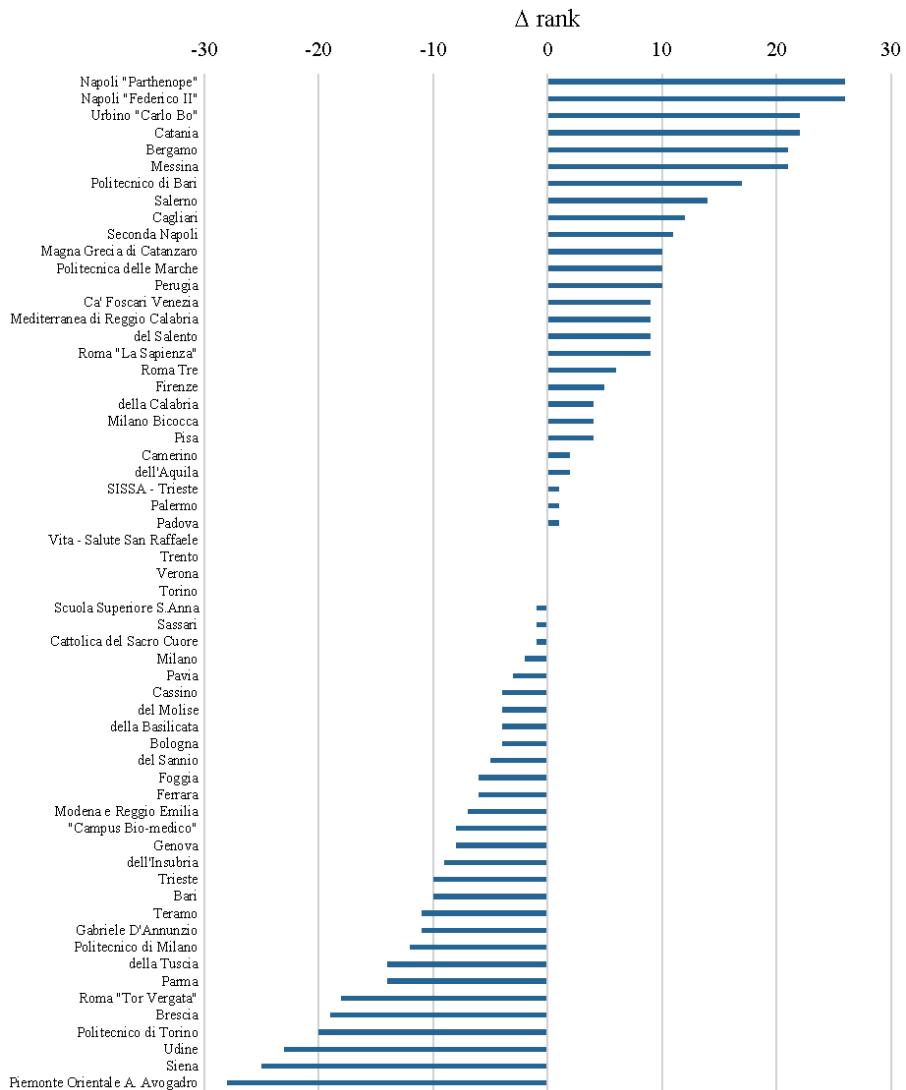
La Tab. 3 mostra i dati analitici di performance e le posizioni in classifica nei due periodi, con l'indicazione della macroregione di appartenenza delle università. L'ultima colonna della tabella riporta la variazione di posizione relativa, calcolata rispetto alla variazione massima possibile per la singola università. La variazione di rank relativa (Δ Rank relativo) è quindi il rapporto tra la variazione di posizione registrata per la singola università e la variazione massima possibile.

Per l'Università Vita-Salute San Raffaele di Milano, ai vertici in entrambi i periodi, un Δ Rank positivo sarebbe stato impossibile. Per l'Università di Salerno, le 14 posizioni guadagnate rappresentano il 74% delle 19 posizioni scalabili tra i due periodi. Seguono, nell'ordine, quattro università: “Federico II” di Napoli, del Salento, “Parthenope” e la SISSA di Trieste. Quest'ultima ha guadagnato una sola posizione, delle due possibili, considerando il suo terzo posto nel primo periodo. La SISSA è anche l'uni-

10. A differenza della VQR, non formuliamo classifiche delle università in base a categorie dimensionali. Il motivo è che nella ricerca accademica non sembrano verificarsi né rendimenti di scala (Abramo, Cicero e D'Angelo, 2012b; Bonaccorsi & Daraio, 2005; Seglen & Aksnes, 2000; Golden & Carstensen, 1992), né di scopo (Abramo, D'Angelo e Di Costa, 2014b).

ca università settentrionale tra le prime dieci della lista: delle restanti nove, due sono situate al centro Italia e sette al sud. Gli atenei con i peggiori ΔRank relativi sono Udine, Siena e l'Università della Basilicata.

Fig. 1 - Variazione del ranking di performance (FSS) 2013-2017 rispetto al 2007-2011 per le 60 università con almeno 30 professori nei settori disciplinari considerati



Tab. 3 - Valori di produttività (FSS), posizione in classifica e relative variazioni tra il 2013-2017 e il 2007-2011 per le 60 università analizzate

Università	Macro regione†	2007–2011			2013–2017			ΔFSS	ΔRank	ΔRank relativo
		Staff	FSS	Rank	Staff	FSS	Rank			
Vita – Salute San Raffaele	N	63	3.27	1	85	2.42	1	-0.85	0	n.a.
Salerno	S	448	0.97	20	417	1.22	6	0.25	14	74%
Napoli “Federico II	S	1.580	0.77	42	1.453	1.04	16	0.27	26	63%
del Salento	S	261	1.00	18	228	1.11	9	0.11	9	53%
Napoli “Parthenope”	S	122	0.65	51	118	1.00	25	0.35	26	52%
SISSA – Trieste	N	50	1.72	3	60	1.85	2	0.13	1	50%
Urbino “Carlo Bo”	C	148	0.65	50	121	0.96	28	0.31	22	45%
Catania	S	823	0.63	52	672	0.95	30	0.32	22	43%
Politecnico di Bari	S	219	0.77	43	179	0.96	26	0.20	17	40%
Messina	S	717	0.62	53	618	0.92	32	0.30	21	40%
Politecnica delle Marche	C	385	0.93	27	365	1.03	17	0.10	10	38%
Milano Bicocca	N	458	1.11	12	436	1.14	8	0.03	4	36%
Bergamo	N	90	0.48	60	86	0.86	39	0.38	21	36%
Firenze	C	1.022	1.02	17	878	1.09	12	0.07	5	31%
Perugia	C	679	0.85	34	596	1.00	24	0.15	10	30%
Roma Tre	C	252	0.94	24	245	1.03	18	0.09	6	26%
Seconda Napoli	S	585	0.73	44	524	0.92	33	0.19	11	26%
Magna Grecia di Catanzaro	S	157	0.78	41	138	0.93	31	0.15	10	25%
Cagliari	S	573	0.61	55	498	0.82	43	0.21	12	22%
della Calabria	S	439	0.98	19	409	1.04	15	0.07	4	22%
Roma “La Sapienza”	C	2.339	0.69	45	1.945	0.87	36	0.18	9	20%
Padova	N	1.335	1.36	6	1.271	1.26	5	-0.10	1	20%
Ca' Foscari Venezia	N	103	0.68	47	85	0.86	38	0.18	9	20%
Medit. di Reggio Calabria	S	139	0.65	49	137	0.85	40	0.20	9	19%
Pisa	C	955	0.95	23	841	1.02	19	0.07	4	18%
Camerino	C	189	0.60	57	192	0.75	55	0.15	2	4%
dell'Aquila	S	425	0.52	59	362	0.75	57	0.23	2	3%
Palermo	S	963	0.68	46	848	0.82	45	0.13	1	2%
Trento	N	239	1.51	4	238	1.40	4	-0.11	0	0%
Verona	N	349	1.33	7	326	1.16	7	-0.17	0	0%
Torino	N	1078	1.14	11	1010	1.10	11	-0.05	0	0%

Tab. 3 - segue

		2007–2011			2013–2017					
Università	Macro regione†	Staff	FSS	Rank	Staff	FSS	Rank	ΔFSS	ΔRank	ΔRank relativo
Scuola Superiore S. Anna	C	43	2.16	2	53	1.80	3	-0.36	-1	-2%
Cattolica del Sacro Cuore	N	740	0.96	22	618	1.00	23	0.04	-1	-3%
Milano	N	1.358	1.22	8	1210	1.10	10	-0.11	-2	-4%
Bologna	N	1.542	1.14	10	1385	1.06	14	-0.08	-4	-8%
Pavia	N	604	0.93	26	512	0.95	29	0.03	-3	-9%
Ferrara	N	389	1.07	14	349	1.02	20	-0.05	-6	-13%
Foggia	S	158	1.06	16	146	1.00	22	-0.06	-6	-14%
Campus Bio-medico	C	78	1.38	5	96	1.09	13	-0.30	-8	-15%
del Sannio	S	108	0.90	29	108	0.91	34	0.01	-5	-16%
Politecnico di Milano	N	819	1.18	9	835	1.01	21	-0.17	-12	-24%
Modena e Reggio Emilia	N	523	0.84	35	470	0.83	42	-0.01	-7	-28%
dell'Insubria	N	248	0.92	28	218	0.86	37	-0.06	-9	-28%
Genova	N	746	0.86	33	699	0.85	41	-0.01	-8	-30%
della Tuscia	C	145	1.09	13	151	0.96	27	-0.13	-14	-30%
Cassino	C	116	0.68	48	104	0.77	52	0.09	-4	-33%
Trieste	N	392	0.84	37	343	0.81	47	-0.03	-10	-43%
Politecnico di Torino	N	583	1.06	15	576	0.90	35	-0.16	-20	-44%
Bari	S	848	0.82	38	748	0.81	48	-0.01	-10	-45%
Sassari	S	375	0.54	58	326	0.67	59	0.13	-1	-50%
Parma	N	625	0.86	32	533	0.81	46	-0.05	-14	-50%
Teramo	S	89	0.79	39	89	0.79	50	0.00	-11	-52%
Brescia	N	380	0.94	25	359	0.82	44	-0.12	-19	-54%
Gabriele D'Annunzio	S	331	0.78	40	307	0.78	51	0.00	-11	-55%
del Molise	S	129	0.62	54	117	0.68	58	0.06	-4	-67%
Piemonte Orientale A. Avogadro	N	186	0.97	21	174	0.81	49	-0.17	-28	-72%
Roma "Tor Vergata"	C	908	0.84	36	775	0.76	54	-0.08	-18	-75%
Udine	N	370	0.89	30	358	0.76	53	-0.13	-23	-77%
Siena	C	451	0.88	31	330	0.75	56	-0.14	-25	-86%
della Basilicata	S	225	0.60	56	219	0.64	60	0.03	-4	-100%

† N, nord; C, centro; S, sud.

Per una comprensione più immediata delle variazioni di performance relativa delle università per i due periodi considerati, la Fig. 2 mostra i grafici di performance (colonne 3 e 6, Tab. 3). Nel secondo grafico, le code si sono inequivocabilmente avvicinate, indicando una riduzione della dispersione delle performance: i 20 atenei in fondo alla classifica in base ai valori di *FSS* 2007-2011 hanno tutti aumentato la loro performance relativa, ma tra i primi 11 atenei classificati, solo la SISSA di Trieste migliora, passando da un valore di *FSS* pari a 1,724 a 1,854.

In termini generali, la dispersione della performance tra i due periodi appare significativamente ridotta (le loro deviazioni standard sono, rispettivamente, 0,425 e 0,290). Per verificare se la convergenza osservata è statisticamente significativa, applichiamo il test di Fisher per l'analisi delle varianze delle distribuzioni prima e dopo, ottenendo un valore *F* pari a 2,140, statisticamente significativo con valori di *p* inferiori allo 0,5%.

La convergenza è in linea con quella riportata da Checchi *et al.*, 2020, che analizzano la variazione della performance misurata secondo i criteri della VQR.

Come si è detto nell'introduzione, il dato negativo di ΔFSS delle università più performanti non deve essere immediatamente interpretato come una riduzione della produttività assoluta dei loro docenti. Infatti, sebbene ciò possa essersi verificato, le riduzioni sono verosimilmente causate anche dall'aumento della produttività media di tutti i docenti italiani. Quelli della coda sinistra hanno contribuito maggiormente ad aumentare il valore di scala delle distribuzioni settoriali e, di conseguenza, ad indurre una riduzione dell'*FSS* normalizzato per i professori della coda destra.

La distribuzione di *FSS* per le università mostra una maggiore variabilità nel primo periodo rispetto al secondo, come evidenziato dal range di variazione max-min (2,791 contro 1,777), dalla deviazione interquartile (Q3-Q1: 0,372 contro 0,228), dalla deviazione standard (0,425 contro 0,290) e dalla *skewness* (3,261 contro 2,874). La media e la mediana di *FSS* del secondo periodo sono leggermente superiori a quelle del primo periodo. Poiché a livello aggregato *FSS* è la media dei valori “riscalati” (dove il fattore di scala è la media delle distribuzioni relative agli SSD, senza gli zeri), la differenza di medie/mediane tra i due periodi può essere dovuta solo alla minore incidenza di professori improduttivi¹¹, cioè di valori *FSS*

11. Poiché, come indicato in [2], le distribuzioni sono riscalate al netto dei valori nulli dovuti agli improduttivi, le differenze tra i valori medi possono essere dovute solo ad una loro diversa incidenza tra i due periodi.

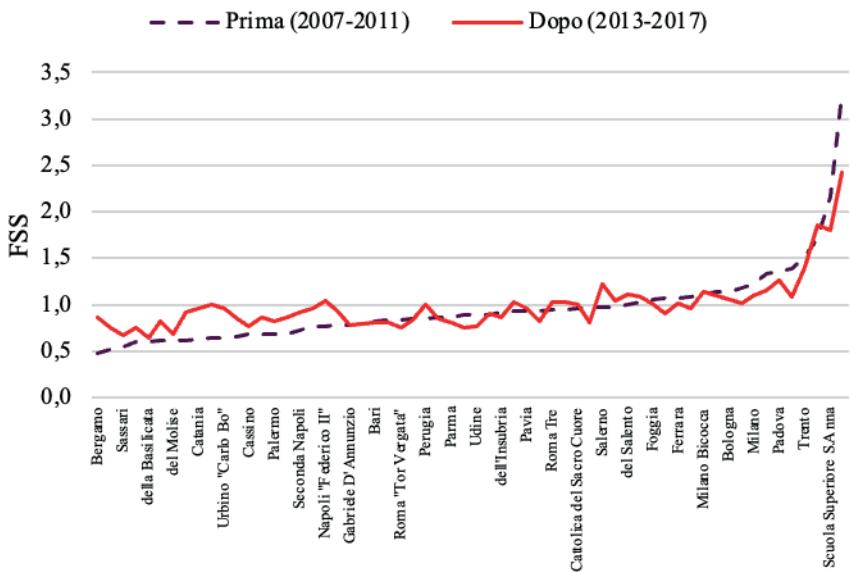
pari a zero, che è in realtà dell'8,3% nel 2007-2011 e del 2,9% nel 2013-2017. La Tab. 4 riporta una sintesi dell'analisi della variazione relativa della performance a livello complessivo degli atenei, suddivisa per macroregione. Delle 24 università del Nord, solo 5 (21%) hanno migliorato la loro posizione nel *ranking* nazionale, rispetto al 64% delle università del Sud (14 su 22). Anche in termini relativi, la performance degli atenei meridionali sembra essere migliorata in modo significativo, visto che il valore medio del loro Δ Rank relativo è positivo e pari a +4%, rispetto al -14% degli atenei del nord e al -2% di quelli del centro.

Per capire meglio in che misura le variazioni di produttività relativa dipendono dalla produttività del primo periodo e dalla localizzazione delle università, stimiamo un modello OLS (Tab. 5), dove la prima variabile è la produttività del primo periodo e la seconda è una dummy, il cui valore è 1 se l'università è situata nel sud e 0 altrimenti.

La generale convergenza delle prestazioni è dimostrata dal segno e dalla significatività del primo regressore (-0,344), ad indicare che l'aumento della produttività è inversamente correlato alla produttività relativa del primo periodo. Inoltre, il coefficiente della dummy "sud" rappresenta una prova significativa del fatto che, in media, le università del sud hanno aumentato la loro produttività più delle università delle altre due macroregioni. Le stime risultanti dai modelli di regressione devono essere interpretate con cautela: osservando le statistiche del coefficiente della dummy "sud" nella Tab. 5, è evidente che il coefficiente è affetto da rumore, l'errore standard è relativamente grande e l'intervallo di confidenza è prossimo allo zero. Alterazioni nell'analisi dei dati e nel campionamento potrebbero cambiare l'entità e/o il segno del coefficiente. Notiamo però che l'analisi si basa su una quasi-popolazione e, al di là di una possibile inferenza quantitativa, i risultati della regressione rivelano un aumento di performance maggiore per le università del Sud, in aggiunta all'effetto di convergenza medio.

Per ogni università analizziamo ora la variazione relativa della performance in ogni ADU. La Tab. 6 presenta il caso dell'Università di Firenze. Tra i due periodi esaminati, su 10 ADU attive, questa università ha migliorato la sua posizione in quattro (Fisica; Biologia; Ingegneria industriale e dell'informazione; Psicologia), è peggiorata in tre (Matematica e informatica; Scienze agrarie e veterinarie; Ingegneria civile) e ha mantenuto invariata la posizione in Chimica (mantenendo il primo posto nazionale), Scienze della Terra (secondo posto) e Medicina (undicesimo).

Fig. 2 - Valori di produttività (FSS) 2013-2017 contro 2007-2011, per le 60 università osservate



Tab. 4 - Variazione di performance (FSS) nei due periodi per le università di ciascuna macroregione

Macroregione	Numero di università	Di cui – migliorate	Di cui – peggiorate	ΔRank relativa medio	Miglioramento massimo	Declino massimo
Nord	24	5	15	-14%	50%	-77%
Centro	14	8	6	-2%	45%	-86%
Sud	22	14	8	4%	74%	-100%

Tab. 5 - Risultati della stima del modello di regressione lineare (variabile dipendente ΔFSS)

	Coeff.	Err. std.	t	P>t	[Intervallo di conf. 95%]
Const.	0.333	0.044	7.630	0.000	0.245 0.420
$FSS_{2007-2011}$	-0.344	0.040	-8.710	0.000	-0.423 -0.265
Sud	0.069	0.031	2.200	0.032	-.006 0.131

Numero di osservazioni = 60; $F(2,57)=51.50$; $Prob>F=0.000$. $R\text{-quadro}=0,675$, $RootMSE=0,114$

Tab. 6 - Valori di produttività (FSS) e relativa posizione in classifica dell'Università di Firenze nei due periodi considerati, per ADU

ADU	2007-2011			2013-2017			ΔRank	Δ FSS
	Personale	FSS	Rank	Personale	FSS	Rank		
1	104	0.87	24	76	0.77	35	-11	-0.10
2	71	0.68	36	58	0.72	33	+3	+0.03
3	109	2.02	1	99	1.87	1	0	-0.16
4	37	1.56	2	37	1.72	2	0	+0.17
5	124	1.20	10	121	1.15	8	+2	-0.05
6	272	1.13	11	220	1.09	11	0	-0.05
7	119	0.62	20	96	0.71	21	-1	+0.09
8	45	0.45	26	43	0.58	30	-4	+0.13
9	99	0.50	44	92	0.95	22	+22	+0.44
10	42	0.67	11	36	1.40	1	+10	+0.74

1 – Matematica e informatica; 2 – Fisica; 3 – Chimica; 4 – Scienze della Terra; 5 – Biologia; 6 – Medicina; 7 – Scienze agrarie e veterinarie; 8 – Ingegneria civile; 9 – Ingegneria industriale e dell'informazione; 10 – Psicologia.

La Tab. 7 mostra, per ciascuna delle 60 università, il numero di ADU valutate, la percentuale di ADU con variazione positiva della posizione tra i due periodi, nonché le specifiche ADU che sono migliorate, peggiorate o rimaste invariate nel posizionamento nazionale.

Tra le università più grandi, valutabili in tutte le 10 ADU considerate, Napoli “Federico II” si colloca al primo posto, avendo migliorato in tutte le discipline tranne Scienze agrarie e veterinarie. Anche Roma “La Sapienza” mostra risultati analoghi, migliorando in 8 ADU su 10. Padova e Bologna sono invece in controtendenza, scendendo nel ranking nazionale in 7 ADU su 10. Anche l'università di Catania, con 9 ADU attive, progredisce in tutte. Sul fronte opposto si collocano Trento e il Politecnico di Milano, che non migliorano in nessuna delle loro 5 ADU attive, anzi peggiorano in quattro di esse.

Ripetiamo l'analisi per ciascuno dei 199 SSD considerati. La Tab. 8 presenta, a titolo di esempio, l'analisi dei SSD di Fisica per l'Università di Firenze, mentre la Tab. 9 riassume i risultati di questo tipo di analisi per tutte le università attive in questa ADU.

Tab. 7 - Variazione della performance (FSS) per ADU delle università nei due periodi

Università	Macro regione [†]	ADU valutate	Di cui in aumento	ADU*									
				1	2	3	4	5	6	7	8	9	10
Napoli "Federico II"	S	10	90%	+	+	+	+	+	+	-	+	+	+
Roma "La Sapienza"	C	10	80%	+	+	+	-	+	+	-	+	+	+
Messina	S	10	60%	-	-	+	=	+	+	+	-	+	+
Palermo	S	10	60%	+	-	+	+	+	-	+	-	+	-
Firenze	C	10	40%	-	+	=	=	+	=	-	-	+	+
Parma	N	10	30%	-	-	-	-	-	+	+	-	-	+
Bologna	N	10	20%	-	+	-	-	+	=	-	-	-	-
Padova	N	10	20%	-	+	-	=	+	-	-	-	-	-
Catania	S	9	100%	+	+	+	+	+	+	+	+	+	+
Pavia	N	9	78%	+	+	+	+	-	-		+	+	+
Cagliari	S	9	67%	+	+	=	+	-	+		+	+	-
Perugia	C	9	56%	+	-	+	-	-	+	+	-	+	
Genova	N	9	44%	+	+	-	+	-	=		-	-	+
Modena e Reggio Emilia	N	9	44%	-	-	+	+	-	+	-		+	-
Milano	N	9	33%	-	-	+	-	-	+	=		+	-
Pisa	C	9	33%	-	+	-	-	+	=	+	-	-	
Trieste	N	9	22%	-	-	-	+	-	-		+	-	-
Bari	S	9	11%	-	-	+	-	-	-	=		-	-
della Calabria	S	8	50%	-	+	-	+	+	-		+	-	
Ferrara	N	8	50%	+	+	+	+	-	-		-	-	
Seconda Napoli	S	8	50%	-	-	+		+	+		+	-	-
Milano Bicocca	N	8	38%	+	+	-	=	-	+			-	-
Politecnica delle Marche	C	8	38%	-	-	=		+	-	+	-	+	
Torino	N	8	38%	+	=	-	+	-	+	-			-
della Basilicata	S	8	25%	-	+	=	-	+		-	=	-	
Udine	N	8	25%	-	-	+		+	-	-	-	-	-
Salerno	S	7	86%	+	+	-		+	+		+	+	
dell'Aquila	S	7	71%	-	+	-		+	+		+	+	
Camerino	C	7	57%	=	-	+	+	+	+	-			
Gabriele D'Annunzio	S	7	29%	-		+	-	=	-		=		+
Roma "Tor Vergata"	C	7	14%	-	+	-		-	-		-	-	
Siena	C	7	14%	=	=	-	-	-	-			+	
Cattolica del Sacro Cuore	N	6	67%	+	+			-	-	+			+

Tab. 7 - segue

Università	Macro regione [†]	ADU valutate	Di cui in aumento	ADU*									
				1	2	3	4	5	6	7	8	9	10
Roma Tre	C	6	67%	-	+		=	+			+	+	
Verona	N	6	33%	-				-	-	+		+	-
Brescia	N	6	17%	+	-			-	-		-	-	
del Salento	S	5	60%	-	-	+		+				+	
Piemonte Orientale A. Avogadro	N	5	40%	+	+	-		-	-				
dell'Insubria	N	5	20%	-	-	-		-	+				
Politecnico di Torino	N	5	20%	-	-	+					-	-	
Politecnico di Milano	N	5	0%	=	-	-					-	-	
Trento	N	5	0%	-	-						-	=	-
Napoli "Parthenope"	S	4	75%	+				=			+	+	
Politecnico di Bari	S	4	75%	+	+						-	+	
Sassari	S	4	75%			+		+	-	+			
Mediterranea di Reggio Calabria	S	4	50%	-						+	-	+	
Urbino "Carlo Bo"	C	4	25%			-	+	-					-
Bergamo	N	3	67%	-								+	+
del Sannio	S	3	67%				-	+				+	
Foggia	S	3	67%				+	-	+				
Magna Grecia di Catanzaro	S	3	67%			-		+	+				
del Molise	S	3	33%					-	+	=			
della Tuscia	C	3	0%			-		-			-		
Ca' Foscari Venezia	N	2	100%	+		+							
Campus Bio-medico	C	2	50%						-			+	
Cassino	C	2	50%								+	-	
SISSA – Trieste	N	2	50%	-	+								
Vita – Salute San Raffaele	N	2	0%					=	=				
Scuola Superiore S.Anna	C	1	0%									=	
Teramo	S	1	0%							-			

[†] N, nord; C, centro; S, sud.

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell'informazione, 10 – Psicologia.

Tab. 8 - Produttività (FSS) e posizione dell'Università di Firenze, nei due periodi considerati, per i 6 SSD di Fisica in cui questo Ateneo è attivo

SSD*	2007-2011			2013-2017			Δ Rank	Δ FSS
	Staff	FSS	Rank	Staff	FSS	Rank		
FIS/01	19	0.62	35	11	0.72	30	+5	+0.09
FIS/02	10	0.34	33	9	0.52	31	+2	+0.18
FIS/03	20	0.94	14	23	0.74	24	-10	-0.21
FIS/04	4	0.94	11	3	1.62	6	+5	+0.68
FIS/05	9	0.60	17	5	0.59	18	-1	-0.01
FIS/07	9	0.46	33	6	0.31	38	-5	-0.15

* FIS/01, Fisica sperimentale; FIS/02, Fisica teorica, modelli e metodi matematici; FIS/03, Fisica della materia; FIS/04, Fisica nucleare e subnucleare; FIS/05, Astronomia e astrofisica; FIS/07, Fisica applicata (beni culturali, ambiente, ecc.).

Tab. 9 - Variazione di posizione per produttività (FSS) degli atenei nei due periodi, per i SSD dell'ADU "Fisica"

Università	Macro regione†	SSD valutati	Di cui in aumento	FIS/01	FIS/02	FIS/03	FIS/04	FIS/05	FIS/06	FIS/07	FIS/08
Bologna	N	8	50%	+	+	+	-	-	=	-	+
Milano	N	8	50%	+	=	-	-	+	-	+	+
Napoli "Federico II"	S	8	50%	+	+	+	-	-	=	+	-
Ferrara	N	7	71%	+	+	-	+	-	+	+	+
Milano Bicocca	N	7	71%	+	-	+	+	=	+	+	+
Roma "La Sapienza"	C	7	71%	+	+	-	+	+	-	+	
Torino	N	7	57%	+	-	+	+	-	-	-	+
Roma Tre	C	7	43%	+	-	+	+	=	=	-	
Padova	N	7	29%	-	+	+	-	-	-	-	-
del Salento	S	7	14%	-	+	-	-	-	-	-	-
Catania	S	6	67%	+	+	-	-	+		+	
dell'Aquila	S	6	67%	+	+	+	+		-	-	
Pavia	N	6	67%	+	-	+	+			+	-
Roma "Tor Vergata"	C	6	67%	+	-	+	-	+		+	
Cagliari	S	6	50%	+	-	-	+	-		+	
Firenze	C	6	50%	+	+	-	+	-		-	
Trento	N	6	50%	-	-	+	-		+		+
Trieste	N	6	50%	-	=	-	+	+		+	
della Calabria	S	6	33%	-	+	+		-	=	-	
Modena e Reggio Emilia	N	6	33%	-	+	-	+		-	=	
Palermo	S	6	33%	-	+	+		-	-	-	-
Pisa	C	6	33%	+	-	-	-	=		+	
dell'Insubria	N	6	17%	+	-	-	-	=		=	
Perugia	C	6	17%	+	-	-	-	-		-	

Tab. 9 - segue

Università	Macro regione [†]	SSD valutati	Di cui in aumento	FIS/01	FIS/02	FIS/03	FIS/04	FIS/05	FIS/06	FIS/07	FIS/08
Bari	S	5	60%	+	-				=	+	+
Genova	N	5	60%	+	+	-	+			-	
Udine	N	5	20%	-	-	-			-	+	
Cattolica del Sacro Cuore	N	4	75%	+	-	+				+	
Piemonte Orientale A. Avogadro	N	4	75%	+	-				+	+	
Salerno	S	4	75%	+	-	+				+	
Scuola Normale Superiore	C	4	75%	+	+	=		+			
Parma	N	4	50%	-	+	+				-	
Politecnico di Torino	N	4	25%	-	-	+	-				
Seconda Napoli	S	4	0%	-	-	-				-	
SISSA – Trieste	N	3	100%		+	+		+			
Camerino	C	3	67%	+	+	-					
Messina	S	3	67%	+		-				+	
Brescia	N	3	33%	-		-				+	
della Basilicata	S	3	33%	+	=				=		
Urbino “Carlo Bo”	C	3	33%	+					-		=
Napoli “Parthenope”	S	2	100%	+				+			
Politecnico di Bari	S	2	100%	+		+					
Siena	C	2	50%	-						+	
Verona	N	2	50%	-						+	
Ca' Foscari Venezia	N	2	0%	-		-					
Politecnica delle Marche	C	2	0%	-						-	
Politecnico di Milano	N	2	0%	-		-					
Campus Bio-medico	C	1	100%		+						
del Sannio	S	1	100%	+							
Guglielmo Marconi (in linea)	C	1	100%	+							
IUSS	N	1	100%					+			
Sassari	S	1	100%							+	
Cassino	C	1	0%	-							
della Tuscia	C	1	0%							-	
Enna – KORE	S	1	0%	-							
Foggia	S	1	0%							-	
Gabriele D'Annunzio	S	1	0%							-	
Mediterranea di Reggio Calabria	S	1	0%	-							
Teramo	S	1	0%							-	
UniNettuno (on line)	C	1	0%	-							

[†] N, nord; C, centro; S, sud

FIS/01, Fisica sperimentale; FIS/02, Fisica teorica, modelli e metodi matematici; FIS/03, Fisica della materia; FIS/04, Fisica nucleare e subnucleare; FIS/05, Astronomia e astrofisica; FIS/06, Fisica per le scienze della terra e dell'atmosfera; FIS/07, Fisica applicata (beni culturali, ambiente, ecc.); FIS/08, Didattica e storia della fisica.

5. Conclusioni

Gli individui e le organizzazioni rispondono in modo diverso agli incentivi economici e reputazionali. Nel caso dei PBRFS nazionali, la risposta dipende dalla misura in cui le organizzazioni a loro volta applicano internamente incentivi a livello individuale, ma anche dal senso di appartenenza e di identificazione dei ricercatori con la missione o la “visione distintiva” della loro organizzazione. In questo lavoro abbiamo verificato la validità di queste premesse, attraverso la valutazione della variazione della performance relativa delle università italiane, a cavallo della prima edizione della VQR.

In diversi contesti nazionali, i PBRFS sono generalmente adottati con l’obiettivo principale di migliorare la produttività scientifica. In questo lavoro, avendo analizzato solo la performance relativa, non è possibile verificare se la VQR abbia raggiunto effettivamente tale obiettivo. Tuttavia, pur non potendo osservare un miglioramento o un peggioramento generalizzato, possiamo certamente affermare che alcune università sono migliorate o peggiorate molto più o molto meno di altre. Probabilmente alcune hanno migliorato la loro performance assoluta al contrario di altre: questa verifica, che è molto più complessa, sarà affrontata in uno studio futuro.

Prima dell’introduzione del PBRFS le università del sud erano molto meno produttive di quelle del nord. Poi, forse spinte dall’incentivo finanziario del PBRFS, di immagine, o forse da questioni di orgoglio hanno reagito alla VQR con maggiore energia delle altre, scalando molte posizioni.

Questo risultato è notevole, in quanto le università meridionali, a causa della loro bassa performance nella VQR, hanno ricevuto in media fondi governativi di entità inferiore rispetto alle altre. Qualora non siano sopravvenuti altri interventi politici a livello locale, in grado di sostenere un simile aumento relativo della produttività di ricerca, si potrebbe concludere che su di esse il PBRFS basato sulla VQR non ha avuto alcun effetto negativo in termini di performance di ricerca. Allo stesso tempo, però, l’effetto sulla reputazione evidenziato dall’aumento della mobilità degli studenti del sud verso le università del nord, come dimostrano le statistiche sulle iscrizioni universitarie (ANVUR, 2014, p. 37), non andrebbe sottovalutato.

Il risultato relativo alla performance è certamente positivo in termini di riduzione del divario nord-sud, un divario molto evidente in Italia in termini di infrastrutture ed attività economiche. Un effetto auspicabile è che il relativo miglioramento di performance delle università meridionali potrebbe rallentare il fenomeno della migrazione degli studenti verso il nord, alla ricerca di una migliore istruzione e attenuare i problemi di discriminazione sociale a livello regionale e nazionale. Infatti, solo le famiglie più ricche

possono sostenere i costi di una istruzione universitaria fuori sede, soprattutto al Nord, dove il costo della vita è molto più alto.

Il nostro studio rivela anche la convergenza delle performance delle università italiane, rappresentata dalla riduzione delle distanze di performance tra quelle più produttive e le altre. Se la VQR è stata concepita come uno stimolo allo sviluppo di università di eccellenza, attraverso finanziamenti selettivi, allora una riflessione potrebbe essere d'obbligo. In ogni caso, nulla si può dire fintantoché non si verificheranno le variazioni nella performance assoluta delle singole università.

L'analisi bibliometrica, come quella svolta, è utile anche a livello del management delle singole organizzazioni di ricerca, per informare, orientare e dunque supportare la loro gestione strategica. Ogni università può acquisire contezza di quale sia il suo posizionamento relativo, quanto sia migliorata o peggiorata nel tempo rispetto ai suoi "concorrenti", in ogni ADU e SSD. Le singole università possono valutare quali politiche e azioni interne hanno funzionato e affinare ulteriormente le loro strategie competitive sulla base di un esame approfondito.

Infine, a nostro avviso, questo lavoro rappresenta un significativo avanzamento negli studi sul tema, essenzialmente per due motivi: i) l'indicatore di performance utilizzato, che non separa la pubblicazione dal suo impatto, tiene conto dei contributi dei singoli autori e, per quanto possibile, delle risorse utilizzate per l'attività di ricerca; ii) l'ampiezza della produzione nazionale considerata, a differenza della piccola frazione utilizzata in molte altre analisi dei PBRFS.

Infine, ultimo ma non meno importante, abbiamo tenuto in conto la diversa intensità di pubblicazione tra i vari settori grazie allo schema di classificazione molto dettagliato dei professori italiani, che consente di limitare notevolmente le distorsioni che si verificano nel confronto diretto delle prestazioni tra ricercatori di ambiti diversi. Questo schema di classificazione, a nostra conoscenza unico al mondo, consente di ottenere diagnosi precise e robuste, e quindi ha possibilità uniche di orientare/influenzare politiche e strategie a livello macro, meso e micro.

Come per altre analisi bibliometriche, lo studio proposto presenta alcuni limiti: i) la disponibilità dei dati, in particolare quelli inerenti alle risorse economiche a livello individuale; ii) la nuova conoscenza prodotta non è sempre incorporata nelle pubblicazioni e i repertori bibliografici (come WoS, qui utilizzato) non indicizzano tutte le pubblicazioni; iii) la misura dell'impatto delle pubblicazioni utilizzando indicatori basati sulle citazioni rimane una previsione, e quindi non è definitiva; iv) le citazioni possono anche essere negative o inappropriate, e in ogni caso certificano solo l'impatto accademico, tralasciando altri tipi di impatto; v) date le limitazioni

sui dati di input (ad esempio, i costi del lavoro e del capitale), si rende necessario il ricorso a diverse ipotesi semplificative per misurare la performance dei professori; vi) i risultati potrebbero essere sensibili agli schemi di classificazione adottati per le pubblicazioni e i professori.

Tali limitazioni dovrebbero indurre cautela nell'interpretazione dei risultati ottenuti. Tuttavia riteniamo che non ci siano effetti sistematici che possano aver determinato una concentrazione di errore su specifiche università o gruppi di università.

8. L'efficacia del reclutamento dei docenti nelle università italiane: gli effetti dell'Abilitazione Scientifica Nazionale

di *Giovanni Abramo, Ciriaco Andrea D'Angelo*

1. Introduzione

La strategia di rafforzamento dei sistemi di istruzione superiore a sostegno della competitività e dello sviluppo socio-economico sta diventando un tema sempre più centrale nelle agende politiche di molti Paesi. L'Unione Europea ritiene che il contributo dell'istruzione superiore sia fondamentale per raggiungere l'obiettivo di diventare l'economia più competitiva e la società maggiormente basata sulla conoscenza del XXI secolo. Tuttavia, il sistema di istruzione superiore dell'UE rimane frammentato e in molti Paesi è ostacolato da una combinazione di eccessivo controllo pubblico e scarsa concorrenza, oltreché da carenti strategie di ricerca e formazione e livelli di spesa insufficienti.

Nei sistemi di istruzione superiore altamente competitivi, come gli Stati Uniti, la ricerca del vantaggio competitivo ha portato allo sviluppo di università d'élite, la cui principale competenza distintiva è la capacità di attrarre, sviluppare e trattenere docenti di grande talento. Diversamente, in diversi Paesi europei, tra cui Germania, Svezia, Danimarca, Italia e Spagna, i meccanismi competitivi sono relativamente più deboli, poiché per molti anni lo Stato ha svolto un ruolo importante nella regolamentazione delle università e i finanziamenti pubblici sono stati solo parzialmente o per nulla regolati su base competitiva. In questi stessi sistemi, il reclutamento e l'avanzamento dei docenti avvengono tipicamente attraverso procedure rigide, regolate da una burocrazia centrale (Auranen e Nieminen, 2010). Di recente sembra però che le cose stiano cambiando e si assiste a una crescente competizione per le posizioni junior e senior in diverse università europee (Bleiklie *et al.*, 2013; Bozzon *et al.*, 2017; Herschberg *et al.*, 2018; Waaijer *et al.*, 2018).

I Paesi con sistemi di istruzione superiore poco competitivi sono in genere caratterizzati da una maggior frequenza di fenomeni di favoritismo nel reclutamento e nella progressione di carriera dei docenti nell'Accademia.

Una delle conseguenze del favoritismo nel reclutamento è la modesta performance complessiva di ricerca, con una scarsa differenziazione tra le università (van der Ploeg e Veugelers, 2008). In Italia, la variabilità della produttività della ricerca tra le università è molto più bassa di quella registrata al loro interno, per cui gli atenei risultano relativamente omogenei in termini di performance di ricerca (Abramo, Cicero e D'Angelo, 2012c).

Il caso italiano è particolarmente interessante, perché mentre il favoritismo nel sistema della ricerca pubblica è generalmente considerato irrimediabilmente cronico (Perotti, 2008; Zagaria, 2007)¹, la performance degli scienziati (misurata sia in termini di contribuzione alla produzione scientifica mondiale nonché di citazioni medie per articolo) è cresciuta costantemente negli ultimi anni, fino a raggiungere i massimi livelli mondiali in termini di impatto scientifico medio per pubblicazione². In teoria, gli effetti cumulativi del favoritismo avrebbero dovuto portare a risultati opposti.

Nel corso dei decenni il governo centrale ha compiuto innumerevoli ma vani tentativi di arginare il fenomeno del favoritismo, in particolare modificando le regole e le procedure per il reclutamento accademico. Con l'istituzione dell'Agenzia Nazionale di Valutazione del sistema Universitario e della Ricerca (ANVUR) cui è stato affidato il processo di Abilitazione Scientifica Nazionale (ASN) per l'accesso al ruolo docente nelle università, sono stati introdotti indicatori di produzione e impatto scientifico e relative soglie il cui superamento costituisce prerequisito necessario (ancorché non sufficiente) per essere abilitati e, dunque, potersi candidare alle procedure valutative bandite dagli atenei per il reclutamento dei propri docenti di prima e seconda fascia.

Sempre ad ANVUR è stata affidata la cosiddetta VQR (Valutazione della Qualità della Ricerca), un esercizio periodico sui cui esiti di base la distribuzione di una percentuale crescente di finanziamenti statali alle università. In particolare, l'indicatore composito che guida la ripartizione della cosiddetta quota premiale del Fondo di Finanziamento Ordinario (FFO) è dato dall'aggregazione di varie misure, tra cui una legata alla qualità della ricerca condotta dai nuovi assunti/reclutati. Evidentemente, l'intento era quello di evitare l'assunzione di persone con performance modeste e di incentivare quella di persone con performance elevate.

1. Vi sono diversi tipi di favoritismo con parziali evidenze empiriche in merito agli esiti.

2. Elaborazioni degli autori su dati Web of Science di *Clarivate Analytics*.

La domanda che ci poniamo in questo capitolo è dunque se, e in che misura, queste politiche, volte a limitare i favoritismi e a promuovere l'efficacia del reclutamento abbiano ottenuto l'effetto sperato. Per rispondere, proponiamo due analisi tra loro correlate. Una basata sulla misurazione e il confronto delle valutazioni delle performance scientifiche dei nuovi assunti (rispetto ai colleghi italiani dello stesso settore e ruolo accademico) in due periodi successivi di cinque anni, prima e dopo l'introduzione della ASN. In caso di successo di tale iniziativa, la seconda coorte dovrebbe mostrare una minore quota di *low performer* e uno *score* medio di performance più elevato. La seconda analisi misura e confronta l'efficacia del reclutamento da parte di tutte le università italiane, nei due quinquenni successivi. Anche in questo caso, se le iniziative intraprese hanno avuto successo, nel secondo periodo si dovrebbe riscontrare un miglioramento dell'efficacia del reclutamento da parte di un'ampia percentuale di università.

Il resto del capitolo è organizzato come segue. Alla rassegna della letteratura sulla discriminazione nei processi di reclutamento, con particolare attenzione al caso italiano, segue la descrizione delle recenti misure volte a promuovere l'efficacia del reclutamento. Presentiamo poi la metodologia adottata e il dataset utilizzato. Seguono la presentazione dei risultati e le considerazioni degli autori nelle conclusioni del lavoro.

2. Discriminazione nel reclutamento dei docenti

In effetti, i problemi di equità e imparzialità nelle nomine accademiche non sono limitati all'Italia. La letteratura internazionale ha dedicato una notevole attenzione allo studio dell'equità nel reclutamento e nella promozione dei docenti (Zinovyeva e Bagues, 2015; Cora-Bramble, 2006; Price *et al.*, 2005; Trotman *et al.*, 2002; Stanley *et al.*, 2007). Una delle conclusioni su cui tali studi convergono è che i fenomeni di discriminazione sembrano svilupparsi soprattutto quando le valutazioni si basano su criteri non trasparenti (Rees, 2004; Allen, 1988). In effetti, il reclutamento dei docenti è spesso descritto come un processo informale in cui pochi accademici influenti selezionano nuovi docenti attraverso meccanismi di cooptazione (Evans, 1995).

Questi fenomeni sono stati esaminati approfonditamente solo in alcune nazioni, in particolare in Turchia (Aydogan, 2012), Australia (Martin, 2009), Spagna (Zinovyeva e Bagues, 2015) e Italia (Abramo, D'Angelo e Rosati, 2015; Perotti, 2008; Zagaria, 2007).

In Italia, di recente, è aumentato l'interesse per lo studio del nepotismo, che è una particolare forma di favoritismo. Mentre Allesina (2011) e Durante *et al.* (2011) segnalano la presenza inconfondibile del fenomeno, Abra-

mo, D'Angelo e Rosati (2014a) sono più cauti: pur non negando la presenza del nepotismo, dimostrano che la probabilità che un "figlio" di un professore ordinario della stessa università non meriti la sua posizione è pari a quella di qualsiasi "non figlio". Questo risultato è in linea con i risultati di molti studi sociologici (Dunn e Holtz-Eakin, 1996; Lentz e Laband, 1989; Simon *et al.*, 1966), secondo i quali i "figli" godono di un vantaggio sostanziale rispetto ai "non figli", poiché possono avere ricevuto dai genitori un bagaglio di conoscenze rilevanti per il loro sviluppo cognitivo e questo faciliterebbe il loro ingresso all'interno dell'accademia.

Qualunque sia la consistenza dei casi di nepotismo, nonché le misure adottate per arginarlo, restano le preoccupazioni per altre forme di favoritismo che distorcono chiaramente il reclutamento dei docenti e la progressione di carriera, in particolare nei Paesi caratterizzati da una scarsa intensità di competizione tra le università. Così Zinovyeva e Bagues (2015), esaminando il fenomeno nel sistema universitario spagnolo, si sono concentrati sul ruolo delle connessioni tra i candidati e i valutatori che compongono le commissioni nominate per stabilire le promozioni accademiche. Essi dimostrano che la performance futura dei candidati promossi che avevano un legame debole con i valutatori era migliore di quella dei loro colleghi "non connessi". Al contrario, i candidati promossi con un forte legame con i valutatori presentavano performance peggiori sia prima sia dopo la promozione.

Per quanto riguarda il sistema accademico italiano, Abramo, D'Angelo e Rosati (2014b) hanno analizzato il processo di selezione per la progressione di carriera prendendo in considerazione tutti i concorsi per professore associato banditi nel 2008 (per un totale di 1.232). Si tratta di un evento molto importante in termini di portata e concentrazione temporale, in gran parte riconducibile al passaggio dal ruolo di ricercatore a quello di professore associato. L'analisi ha mostrato che, nei tre anni successivi ai concorsi, i nuovi professori associati sono stati in media più produttivi dal punto di vista scientifico rispetto ai colleghi in servizio. Tuttavia, sono emerse diverse criticità, in particolare il caso di tanti candidati "perdenti" nelle selezioni che però mostrano una performance di ricerca superiore a quella dei vincitori del concorso e di alcuni vincitori che sono risultati totalmente improduttivi nel periodo successivo al concorso. In particolare, è emerso che nei tre anni precedenti il concorso, il 29% dei vincitori aveva una produttività inferiore alla mediana della distribuzione delle prestazioni dei loro pari nello stesso settore e che il 5,5% dei vincitori non aveva contribuito ad alcun avanzamento scientifico nel proprio settore. Un'analisi dei singoli concorsi ha evidenziato che nel triennio successivo, quasi la metà dei candidati vincitori si collocavano al di sotto della media nel loro settore per produttività di ricerca.

Un lavoro successivo (Abramo, D'Angelo e Rosati, 2015) ha dimostrato che il successo di un candidato accademico in Italia non è dovuto tanto al merito scientifico, quanto piuttosto al numero di anni di appartenenza alla stessa università del presidente del comitato di selezione.

3. Il *framework* di riferimento determinato dalle *policy* recenti

Come già anticipato nel capitolo 3 (cfr. paragrafo 2 del capitolo 3), gli innumerevoli tentativi dei governi italiani di contrastare il favoritismo nei concorsi accademici si sono rivelati per lo più infruttuosi. Qui illustreremo alcuni tratti salienti delle iniziative introdotte attraverso la Legge di riforma 240 del 2010 per meglio definire il contesto teoretico che ha ispirato lo studio i cui risultati sono illustrati in questo capitolo. La Legge 240 ha introdotto un doppio livello di valutazione per la selezione dei professori associati e ordinari. Il primo livello è nazionale, gestito direttamente dal MUR, ed è finalizzato ad abilitare solo candidati in possesso di una sufficiente qualificazione scientifica. Il secondo è gestito dalle singole università, per selezionare i candidati più adatti alle esigenze specifiche dell'ateneo tra quelli prima abilitati a livello nazionale.

L'abilitazione non è un meccanismo nuovo nel panorama dei processi di reclutamento universitario a livello europeo. Da alcuni anni viene applicata in forme diverse in Germania (Enders, 2001), in Francia (Musselin, 2004) e in Spagna (MEC, 2007; Sanz-Menéndez e Cruz-Castro, 2019).

Una caratteristica nuova e unica dell'ASN italiana è l'utilizzo di indicatori bibliometrici a supporto delle valutazioni delle commissioni settoriali. Infatti, nello schema iniziale del Ministero, l'abilitazione era subordinata al superamento di valori soglia in tutti e tre gli indicatori bibliometrici di performance di ricerca (numero di pubblicazioni, citazioni, h-index), creando così un vero e proprio blocco nei confronti dei candidati mediocri.

Tuttavia, la pubblicazione del decreto che definiva i criteri e i parametri di valutazione, seguita dalla diffusione dei valori soglia, ha scatenato un dibattito divenuto via via sempre più acceso. Questo ha indotto il ministro a rimuovere i vincoli di cui sopra, e a chiarire che sarebbe stata le commissioni settoriali a definire i parametri e i criteri di valutazione dei candidati e che il decreto ministeriale si limitava a fissare delle "linee guida": la posizione dei candidati rispetto alle soglie degli indicatori ministeriali non era condizione (né garanzia) per ottenere l'abilitazione. Di conseguenza, alcune commissioni di valutazione hanno semplicemente ignorato tali soglie, mentre altre hanno richiesto che i candidati ne superassero solo una. Pochi studiosi hanno indagato l'equità dei giudizi nel primo bando ASN. Marzolla (2015)

ha analizzato la correlazione tra esito e valori degli indicatori bibliometrici. Pautasso (2015) ha invece esplorato le questioni di genere sia nella partecipazione che negli esiti. Marini (2017) ha dimostrato che i migliori predittori del conseguimento dell'abilitazione sono: (1) la qualità della ricerca; (2) il ruolo accademico attuale; (3) la giovane età, soprattutto tra candidati dello stesso ruolo. Fassari e Spanò (2019) si chiedono se artefatti valutativi come l'ASN rimuovano o accentuino le differenze di performance legate al genere.

Abramo e D'Angelo (2015a) hanno analizzato la relazione tra gli esiti dell'abilitazione, il merito scientifico dei candidati (misurato dai valori degli indicatori bibliometrici) e altre variabili esplicative quali proxy di possibili pratiche di favoritismo e discriminazione. Dalle analisi emerge molto chiaramente, tra l'altro, che per i candidati già in ruolo come ricercatore o professore associato, ai fini della progressione di carriera è sufficiente avere una performance scientifica inferiore a quella dei candidati esterni (provenienti da enti di ricerca, industria, ecc.) al mondo accademico, per avere le stesse probabilità di successo.

D'altra parte un anno dopo l'introduzione dell'abilitazione scientifica, il MUR ha lanciato il secondo esercizio di valutazione della ricerca nazionale, la VQR 2004-2010, con lo scopo (come già accennato in precedenza) di disporre di un cruscotto informativo per assegnare una quota di finanziamenti pubblici alle università e agli enti pubblici di ricerca, sulla base della loro performance di ricerca.

La valutazione complessiva delle istituzioni era affidata alla media ponderata di una serie di indicatori e, in particolare: per il 50% alla qualità dei prodotti di ricerca presentati e per il restante 50% ad un indicatore composto ottenuto dalla combinazione di altri 6 indicatori: 10% per, rispettivamente, la capacità di attrarre fondi di ricerca, la qualità dei prodotti presentati dal cosiddetto personale in mobilità (reclutati e avanzati di ruolo), l'internazionalizzazione e i programmi di dottorato; e un ulteriore 5% per il finanziamento indipendente della ricerca e per il miglioramento complessivo.

Poiché i risultati della ricerca dei nuovi assunti nel periodo in esame contribuivano allo *score* complessivo delle istituzioni e dei dipartimenti più di quelli dei colleghi in servizio, e una quota dei finanziamenti governativi si basava su tale *score*, l'incentivo per le università era quello di essere efficaci nel processo di reclutamento. Va tuttavia notato che questo può avere introdotto una distorsione a favore delle istituzioni più prestigiose, che sono più abili delle altre nell'attrarre scienziati di alto livello.

Analogamente ad altri sistemi di finanziamento della ricerca basati sulle performance, la “ricompensa” finanziaria va all'università ed è quindi collettiva, mentre i risultati dell'esercizio dipendono dalla qualità delle singole pubblicazioni e, dunque, dalle prestazioni individuali dei professori.

Inoltre, le università non sono obbligate a spendere i fondi per promuovere la ricerca degli accademici, né sono obbligate ad assegnarli internamente in base al merito.

L'ASN, dal canto suo, rappresenta un vero e proprio blocco per l'ingresso nel sistema accademico di soggetti con performance di ricerca scarso. Nelle aspettative dei *policy maker*, gli effetti sinergici di ASN e VQR avrebbero dovuto limitare i favoritismi e aumentare l'efficacia del reclutamento. Questo è ciò che valuteremo in seguito, sia a livello complessivo (intero sistema accademico) che di singole università.

4. Specifiche metodologiche

In questa sezione presentiamo le motivazioni alla base del nostro approccio per valutare il successo delle iniziative intraprese per promuovere il reclutamento basato sul merito nel mondo accademico, e le metodologie specifiche applicate per misurare e confrontare le performance di ricerca dei nuovi assunti e l'efficacia del reclutamento da parte delle università. Nel fare ciò, identifichiamo anche i limiti e le ipotesi che gli approcci quantitativi inevitabilmente comportano.

Sebbene i criteri di selezione siano diversi tra i settori e tra le commissioni valutatrici dello stesso settore, il criterio più importante nella selezione dei candidati al ruolo di professore in Italia rimane l'eccellenza nella ricerca. Quando il numero di osservazioni è elevato, in assenza di altre informazioni, affidarsi alle sole misure di performance della ricerca per valutare l'efficacia del reclutamento potrebbe essere accettabile.

Tuttavia, i cambiamenti di contesto tra i due periodi potrebbero distorcere il confronto. Ad esempio, è probabile che le distribuzioni degli indicatori bibliometrici usati per valutare la performance di ricerca, in due periodi diversi, siano diverse perché il numero di pubblicazioni indicizzate nei database bibliografici tende ad aumentare nel tempo, o perché le risorse messe a disposizione della ricerca possono variare. Per evitare distorsioni, non confrontiamo direttamente gli *score* di performance di ogni neoassunto delle due coorti. Per ogni singola coorte calcoliamo lo *score* di performance di ciascun soggetto reclutato rispetto a tutti i colleghi italiani appartenenti allo stesso ruolo accademico e allo stesso settore. A questo punto confrontiamo gli *score* di performance delle due coorti.

Confrontare le valutazioni di performance comparata, piuttosto che gli *score*, ci permette di controllare altri possibili fattori che possono influenzare nel tempo l'ambiente competitivo. È stato dimostrato che, quando in un dato momento storico sono disponibili più posizioni che candidati validi, la qua-

lità dei selezionati è relativamente più bassa (Cruz-Castro e Sanz-Menéndez, 2010; Lissoni *et al.*, 2011; Hargens, 2012)³. Poiché la valutazione della performance che calcoliamo è riferita a tutti i colleghi dello stesso settore e ruolo accademico, coloro che sono stati assunti nella prima tornata (che si suppone abbiano un profilo scientifico elevato) appariterebbero a un ruolo accademico superiore nella seconda tornata, quindi non falserebbero il confronto. Ad esempio, il secondo miglior ricercatore del primo turno risulterebbe il migliore del secondo turno una volta che il migliore del primo turno è stato promosso al ruolo successivo, a parità di altre condizioni. In realtà, come mostreremo e come ci si aspetterebbe, poiché le riforme sono state introdotte proprio perché il favoritismo, anziché il merito, caratterizzava il reclutamento accademico, raramente sono stati selezionati i candidati migliori nel primo periodo (come da rassegna della letteratura precedentemente illustrata).

Un'altra possibilità è che nel secondo periodo una quota relativamente maggiore di persone con risultati elevati non si sia candidata. In questo caso, il confronto con il primo periodo introdurebbe una distorsione nella valutazione dell'impatto reale delle riforme introdotte. A parte il fatto che sarebbe difficile trovare una spiegazione ragionevole, ma in ogni caso le università non sono comunque obbligate a reclutare candidati con performance mediocri. Come mostreremo, infatti, la maggior parte delle assunzioni nel secondo periodo sono state promozioni interne (83%), che non hanno contribuito a un aumento del personale, rivelando che non c'era un bisogno urgente di reclutamento.

Per questo motivo, il confronto tra le valutazioni delle performance complessive dei neoassunti e l'efficacia media del reclutamento in due periodi successivi, prima e dopo l'introduzione di queste misure di *policy*, dovrebbe consentirci di valutare se tali misure sono riuscite a migliorare l'efficacia del reclutamento a livello complessivo e di singola università.

Di seguito spiegheremo prima come misuriamo la performance di ricerca e successivamente l'efficacia del processo di reclutamento.

4.1. Misurare la performance della ricerca

Negli ultimi anni si è assistito a una proliferazione di indicatori bibliometrici per la misurazione della performance di ricerca. Alcuni studiosi hanno proposto indicatori compositi piuttosto complicati (Murias *et al.*,

3. Lo stesso fenomeno si verifica con le allocazioni competitive dei fondi (Gush *et al.*, 2018).

2008; Hardeman *et al.*, 2013), mentre altri preferiscono indicatori multidimensionali, in grado di cogliere le diverse sfaccettature dell'attività di ricerca (Gorraiz e Gumpenberger, 2015; Moed e Halevi, 2015). Altri ancora hanno criticato aspramente i più diffusi indicatori bibliometrici di performance, suggerendo che l'adozione o la combinazione di più indicatori non idonei rende inaffidabile l'intero processo di valutazione della ricerca (Abramo e D'Angelo, 2016a).

In ogni caso, la performance e i suoi indicatori non possono essere disgiunti dagli obiettivi da perseguire e dal contesto di riferimento. In questo caso specifico, partiamo dal presupposto che le università sono desiderose di assumere gli scienziati più produttivi, cioè quelli che si distinguono in termini di impatto scientifico totale della loro attività di ricerca.⁴ Di conseguenza, cerchiamo di misurare la produttività della ricerca.

La maggior parte degli studiosi di scientometria/bibliometria definisce la produttività di ricerca come il numero di pubblicazioni realizzato nel periodo di osservazione. Poiché le pubblicazioni hanno valore diverso (impatto), preferiamo adottare una definizione più significativa di produttività: il valore della produzione per unità di valore del lavoro necessario a realizzarla, a parità di altri fattori produttivi. Questa definizione riconosce che le pubblicazioni che incorporano nuova conoscenza hanno un impatto diverso sul progresso scientifico, impatto che i bibliometristi approssimano attraverso il conteggio citazionale.

In particolare, data la dipendenza del comportamento citazionale dal settore, per ogni pubblicazione si standardizzano le citazioni rispetto ad un valore atteso di riferimento, nel nostro caso pari alla media della distribuzione delle citazioni ricevute da tutte le pubblicazioni italiane indicizzate nello stesso anno e categoria tematica (*subject category – SC*).

Inoltre, teniamo conto del contributo del singolo autore al prodotto di ricerca cui egli ha partecipato, contributo inversamente proporzionale al numero di co-autori e che, in alcune discipline, è ulteriormente “segnalato” dalla posizione dei nominativi nella lista degli autori. Purtroppo, non sono generalmente noti dati dettagliati sui fattori di produzione a disposizione di ciascun professore, né in Italia né in altri paesi, almeno a conoscenza degli autori. Partiamo quindi dal presupposto che tutti i professori dello stesso settore, per fare ricerca dispongano delle stesse risorse e che le ore dedicate ad attività istituzionali complementari (didattica, terza missione, ecc.) siano più o meno le stesse per tutti.

4. Un target complementare potrebbe essere costituito da scienziati in grado di realizzare progressi scientifici o tecnologici (Abramo & D'Angelo, 2015b).

A livello individuale, possiamo misurare la produttività media annuale, mediante una *proxy* denominata *fractional scientific strength* (*FSS*), già definita operativamente nel capitolo precedente.

I dati sui docenti di ciascuna università sono stati estratti dal database del personale universitario italiano, gestito dal MUR⁵. Il dataset bibliometrico utilizzato per misurare la *FSS* è stato estratto dall’Osservatorio Italiano della Ricerca Pubblica (ORP), una banca dati sviluppata e gestita dagli autori e derivata su licenza dalla Core collection del Web of Science (WoS). Partendo dai dati grezzi di WoS e applicando un complesso algoritmo di riconciliazione dell’affiliazione dell’autore e di disambiguazione della sua reale identità, ogni pubblicazione (articolo, *review*, *letter* o *conference proceeding*) viene attribuita agli accademici che l’hanno realizzata⁶. Grazie a questo algoritmo possiamo produrre classifiche di produttività di ricerca a livello individuale, su scala nazionale. In base al valore di *FSS* si ottiene, per ogni SSD e ruolo accademico, una graduatoria espressa sia su una scala di percentili da 0 a 100 (dal peggiore al migliore) o come rapporto *FSS* rispetto alla media, da confrontare con la performance di tutti gli accademici italiani dello stesso SSD e ruolo accademico.

Poiché la copertura di WoS è piuttosto limitata nelle aree umanistiche e parte di quelle sociali, il campo di osservazione è limitato alle discipline STEMM⁷, al fine di garantire la robustezza delle analisi condotte.

4.2. Misurare l’efficacia del reclutamento

Coerentemente con Abramo, D’Angelo e Rosati (2016a), nel misurare l’efficacia del reclutamento adottiamo due prospettive. La prima consiste nell’indagare la capacità di un’università di attrarre i professori con il più elevato profilo scientifico rispetto al corpo docente già in servizio presso l’Ateneo ovvero in relazione ai risultati scientifici o alla “reputazione” dell’università stessa. In base a questa prospettiva, definita “interna”, confrontiamo la performance media dei professori assunti con quella del SSD della organizzazione in cui vengono inseriti. La seconda prospettiva, che definiamo “esterna”, ignora la reputazione scientifica dei docenti in servi-

5. <http://cercauniversita.cineca.it/php5/docenti/cerca.php>, ultimo accesso 31 luglio 2023.

6. La media armonica della *precision* e della *recall* (F-measure) delle *authorship* disambigue dall’algoritmo è di circa il 97% (margine di errore del 2%, intervallo di confidenza del 98%).

7. STEMM è l’acronimo inglese di Science, Technology, Engineering, Math, and Medicine.

zio e confronta la performance media dei professori assunti con quella di tutti i professori italiani dello stesso SSD e dello stesso ruolo accademico.

Ciascuna delle due prospettive viene analizzata utilizzando due indicatori, per un totale di quattro indicatori.

Per la prospettiva interna utilizziamo i due indicatori:

R 1.1: Per ogni nuovo assunto si calcola il rapporto tra la sua FSS e la FSS media dei docenti della stessa università, appartenenti allo stesso SSD e con lo stesso ruolo accademico, $\overline{FSS^I}$ (produttività interna)^{8,9}. La media di questi rapporti, calcolata sul totale dei nuovi assunti (N), ci dà il valore dell'indicatore R 1.1 per ogni università. In formula:

$$R\ 1.1 = \frac{1}{N} \sum_{j=1}^N \frac{FSS_j}{\overline{FSS^I}} \quad [2]$$

Questo indicatore esprime la capacità dell'università di attrarre professori di talento, rispetto al proprio potenziale. Un'università con solo professori di alto livello può al massimo raggiungere un rapporto pari a 1. Le università meno produttive possono invece avere rapporti molto alti. Un rapporto superiore a 1 indica che i nuovi assunti hanno aumentato la performance media dell'università.

R 1.2: Questo indicatore è dato dalla percentuale di assunti in un'università che hanno FSS superiori alla media di FSS per il corpo docente, di pari ruolo, dello stesso SSD, sul totale dei nuovi assunti. In formula:

$$R\ 1.2 = \frac{1}{N} \sum_{j=1}^N x_j \quad [3]$$

Con $x_j = 1$ se $FSS_j > \overline{FSS^I}$; altrimenti 0.

Questo secondo indicatore compensa i potenziali effetti degli *outlier*, che possono influenzare R 1.1.

Per misurare l'efficacia del reclutamento secondo la prospettiva “esterna”, utilizziamo due indicatori:

8. Poiché il numero di professori in servizio nello stesso SSD e con lo stesso ruolo accademico potrebbe essere molto basso (o nullo), in una determinata università, confrontiamo la performance di tutti i gradi, dopo aver normalizzato l'FSS in base allo stipendio medio per ruolo.

9. In teoria, è possibile che tutto il personale di ricerca in servizio sia improduttivo ($\overline{FSS^I} = 0$), nel qual caso si avrebbe ($R\ 1.1 = \infty$). Per aggirare il problema matematico, assegniamo al denominatore il valore più basso della distribuzione nazionale della produttività FSS^I e aumentiamo il numeratore dello stesso valore. Infine, in caso di assenza di personale di ricerca in servizio, l'efficacia interna non è misurabile.

R 2.1: Per ogni nuovo assunto si calcola innanzitutto il rapporto tra la sua FSS e la FSS media di tutti i professori dell'intero sistema universitario italiano nello stesso SSD e ruolo accademico ($\overline{FSS^E}$ = produttività esterna). L'indicatore R 2.1 per ogni università è dato dalla media di questi rapporti calcolati sul totale dei nuovi assunti (N). In formula:

$$R\ 2.1 = \frac{1}{N} \sum_{j=1}^N \frac{FSS_j}{\overline{FSS^E}} \quad [4]$$

R 2.2: Questo indicatore è dato dalla percentuale di assunti in un'università, con FSS superiore alla media delle FSS misurata per i colleghi italiani di pari ruolo ed SSD, sul totale di tutti gli assunti (N). In formula:

$$R\ 2.2 = \frac{1}{N} \sum_{j=1}^N y_j \quad [5]$$

Con $y_j = 1$ se $FSSE_j > \overline{FSS^E}$; altrimenti 0.

Si utilizzano le consuete cautele nell'applicazione della metodologia bibliometrica e nell'interpretazione dei relativi risultati. Limiti e assunzioni sono presenti sia per l'indicatore FSS sia per la metodologia utilizzata nel valutare la capacità di reclutamento e il turnover. Il principale limite della metodologia risiede nell'ipotesi che la performance di ricerca dei professori sia in grado di spiegare il loro “merito” accademico.

4.3. *Il dataset di analisi*

La VQR 2004-2010 è stata lanciata nel novembre 2011 e i risultati sono stati pubblicati nel 2013. Poiché l'iter amministrativo che le università devono seguire per bandire i concorsi per il reclutamento del personale docente, istituire le commissioni di valutazione e giudicare i candidati è lungo, ci aspettiamo che gli effetti positivi sull'efficacia del reclutamento derivanti dalla combinazione sinergica di abilitazione scientifica e VQR possano essere osservabili a partire dal 2013. Per questo motivo confrontiamo l'efficacia del reclutamento nel quinquennio 2013-2017 con quella del precedente periodo 2008-2012¹⁰.

10. Un periodo di pubblicazione di cinque anni è adeguato per ridurre il problema della scarsità di pubblicazioni a livello di settore e delle fluttuazioni dipendenti dall'anno con effetti sistematici sui risultati (Abramo, D'Angelo e Cicero, 2012a).

Dal database sul personale universitario italiano estraiamo tutti i professori assunti tra il 2008 e il 2017. Per la misurazione della loro produttività e il calcolo degli indicatori descritti nella Sezione 4.2, consideriamo:

- la finestra di pubblicazione 2007-2011, per i professori assunti nel periodo 2008-2012;
- la finestra di pubblicazione 2012-2016, per i professori assunti nel periodo 2013-2017¹¹.

In entrambe le finestre, consideriamo il ruolo accademico dei professori alla fine di ciascun periodo.

Per garantire la solidità della valutazione della produttività della ricerca, escludiamo i professori con meno di tre anni di servizio in ciascun periodo di osservazione. Per lo stesso motivo, escludiamo dall'analisi le assunzioni nelle scienze sociali e nelle discipline umanistiche, a causa della limitata copertura di WoS in queste discipline.

I reclutamenti sono classificabili secondo le seguenti tipologie:

- promozioni interne (IP), cioè promozioni a un ruolo accademico superiore all'interno della stessa università;
- promozioni esterne (EP), cioè reclutamenti per un ruolo accademico superiore rispetto a quello di provenienza posseduto in altra università;
- trasferimenti (TR), cioè reclutamenti da un'altra università ma con lo stesso ruolo accademico precedente;
- reclutamenti esterni (OR), cioè reclutamenti al di fuori del sistema accademico e dai ruoli universitari esistenti (includendo dunque le assunzioni di tutti i “ricercatori”).

La Tab. 1 presenta il dataset utilizzato per l'analisi, suddiviso per ADU¹² e tipo di assunzione. Si osserva che il numero totale di reclutamenti è quasi raddoppiato nel secondo periodo. Allo stesso tempo, il numero di professori assunti per promozione interna è quasi triplicato. Questo fenomeno sembra particolarmente legato alla possibilità di una minore efficacia del reclutamento. Un altro aspetto significativo è che nel secondo periodo la maggior parte dei reclutamenti avviene per gradi superiori a quello di ricercatore.

11. Gli autori sono consapevoli che si tratta di una scelta non ottimale. Un modo più appropriato di procedere sarebbe quello di misurare e confrontare la performance di ogni neoassunto nel quinquennio precedente il bando di concorso. Tuttavia, ciò richiederebbe un processo operativo molto oneroso e, in particolare, l'impostazione di 10 diversi *benchmark*, ciascuno dei quali includerebbe l'FSS quinquennale di ogni professore italiano (circa 40.000 in totale) nei SSD osservati.

12. La denominazione completa dell'ADU 8 è Ingegneria Civile e Architettura, e dell'ADU 11 Scienze storiche, filosofiche, pedagogiche e psicologiche. Nel testo riportiamo solo le sotto-aree valutabili con il metodo bibliometrico, omettendo le altre che non abbiamo considerato.

Tab. 1 - Dataset analisi, per ADU e tipo di reclutamento

ADU*	2008-2012					2013-2017				
	Totali reclutati	EP (%)	IP (%)	TR (%)	OR (%)	Totali reclutati	EP (%)	IP (%)	TR (%)	OR (%)
1	347	4.0	44.7	18.4	32.9	651	6.0	82.3	4.5	7.2
2	230	1.3	56.1	4.3	38.3	631	2.9	84.5	2.5	10.1
3	276	0.0	59.8	4.7	35.5	653	1.1	91.3	0.6	7.0
4	120	2.5	49.2	10.0	38.3	224	2.2	87.1	3.1	7.6
5	460	0.9	47.0	6.5	45.7	884	2.7	81.9	3.2	12.2
6	958	3.1	48.3	9.7	38.8	1655	6.0	77.6	5.7	10.6
7	287	1.7	57.8	4.5	35.9	591	1.7	88.5	1.9	8.0
8	179	2.2	46.9	7.3	43.6	318	5.3	81.4	5.3	7.9
9	677	2.1	56.7	5.0	36.2	1314	3.3	86.5	2.7	7.5
11	207	2.9	40.6	12.1	44.4	303	5.0	78.9	6.3	9.9
<i>Totale</i>	<i>3.741</i>	<i>2.2</i>	<i>50.9</i>	<i>8.2</i>	<i>38.7</i>	<i>7224</i>	<i>3.9</i>	<i>83.4</i>	<i>3.6</i>	<i>9.1</i>

EP: Promozioni esterne; IP: Promozioni interne; TR: Trasferimenti da altra università; OR: Reclutamenti dall'esterno del sistema accademico.

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell'informazione, 11 – Psicologia.

5. Risultati

Questa sezione confronta innanzitutto la valutazione della performance di ricerca del personale reclutato prima e dopo l'introduzione di ASN e VQR. Quindi, presenta e confronta le misure dell'efficacia universitaria del reclutamento nei due periodi, sia dal punto di vista interno che esterno, applicando i quattro indicatori descritti nella sezione 4.2. Il confronto dei risultati viene realizzato prima a livello complessivo e successivamente a livello di singola università.

5.1. Performance di ricerca dei reclutati

Abbiamo misurato l'FSS di tutti i professori e valutato la performance di ogni neoassunto rispetto a tutti i professori dello stesso SSD con pari ruolo accademico. In Tab. 2 riportiamo le statistiche descrittive dei *rating*

di performance nei due periodi, per un confronto diretto. Osserviamo che dopo l'introduzione delle due misure, la percentuale di assunzioni di soggetti improduttivi è diminuita, probabilmente soprattutto grazie alle soglie previste dall'ASN. Questo effetto positivo è compensato dalla percentuale di assunti al di sopra della mediana, che è anch'essa diminuita. Di conseguenza, la performance media complessiva non è aumentata, come dimostrano i percentili medi e mediani.

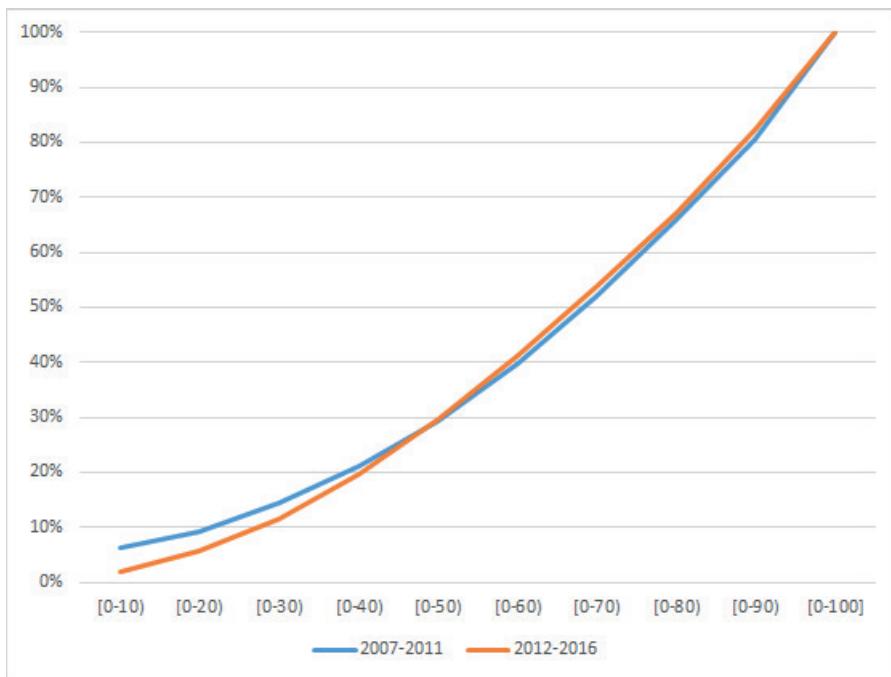
Tab. 2 - Statistiche descrittive della performance di ricerca (percentile di FSS) dei reclutati nei due periodi

	2007-2011	2012-2016
Numero di reclutati	3.741	7.224
Media	63	64
Mediana	69	67
Deviazione std.	27.8	25.1
100 percentile (top)	3.0%	2.9%
[75;100)	38.2%	37.1%
[50;75)	29.5%	30.5%
[25;50)	17.5%	21.3%
(0;25)	6.6%	7.4%
Improduttivi (FSS=0)	5.3%	0.8%

Le distribuzioni di frequenza cumulata delle performance nei due periodi sono quasi totalmente sovrapponibili, tranne che per i valori estremi più bassi (Fig. 1).

In Tab. 3 sono riportati, per i due periodi, i percentili medi dei *rating* di performance per tipo di reclutamento. Osserviamo che nel secondo periodo le promozioni interne, che rappresentano oltre l'83% del totale dei reclutati, registrano uno score medio di performance più basso rispetto a quello del primo periodo, in cui le promozioni interne si attestano al 50% del totale dei reclutamenti. Lo score medio di performance dei primi è il più basso tra tutti i tipi di reclutamento, ad eccezione di quelli trasferiti da altre università.

Fig. 1 - Distribuzioni cumulative di frequenza degli score di performance di ricerca (percentile di FSS) dei reclutati nei due periodi



Tab. 3 - Statistiche descrittive degli score di performance di ricerca (percentile di FSS) dei reclutati nei due periodi

Tipo*	2007-2011		2012-2016	
	%	FSS percentile medio	%	FSS percentile medio
EP	2.2	73.2	3.9	66.7
IP	50.9	65.6	83.4	63.5
TR	8.2	61.4	3.6	62.4
OR	38.7	59.3	9.1	65.9
Media		63.0		63.8

EP: Promozioni esterne; IP: Promozioni interne; TR: Trasferimenti da altra università; OR: Reclutamenti dall'esterno del sistema accademico.

5.2. Efficacia del reclutamento

Di seguito, misuriamo e confrontiamo l'efficacia del reclutamento nei due periodi, sia dal punto di vista interno che da quello esterno, applicando i quattro indicatori descritti nella sezione 4.2. Il confronto avviene prima a livello complessivo e poi a livello di singola università.

5.2.1. Prospettiva interna

La Tab. 4 mostra che i valori di R1.1 sono maggiori di 1 in tutte le ADU in entrambi i periodi, il che significa che i reclutati hanno mediamente aumentato la performance degli atenei che li hanno reclutati. Tuttavia, il confronto mostra che il valore medio di R1.1 diminuisce dal primo al secondo quinquennio, complessivamente e in tutte le ADU, ad eccezione dell'ADU 1 (Matematica). Il calo è particolarmente evidente in Psicologia (ADU 11) e Scienze della Terra (ADU 4). Al contrario, i valori di R1.2 aumentano complessivamente e per tutte le ADU, ad eccezione di Fisica (ADU 2). Ciò significa che nel secondo periodo, nonostante cresca la percentuale dei neoassunti con performance superiori rispetto ai colleghi in servizio, il loro contributo al miglioramento delle performance dell'università

Tab. 4 - Valori medi dei due indicatori secondo la prospettiva interna, per ADU

ADU*	R1.1		R1.2	
	08-12	13-17	08-12	13-17
1	4.79	5.76	58.7%	66.4%
2	2.62	2.58	61.7%	57.9%
3	4.76	0.98	54.6%	64.4%
4	9.76	1.07	60.6%	70.1%
5	2.15	1.11	53.5%	63.1%
6	4.59	2.76	53.0%	63.5%
7	6.16	1.41	59.3%	65.5%
8	4.06	1.12	61.4%	68.1%
9	3.43	2.97	60.4%	65.0%
11	35.87	8.01	51.7%	61.4%
<i>Totale</i>	<i>5.87</i>	<i>2.66</i>	<i>56.7%</i>	<i>64.0%</i>

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell'informazione, 11 – Psicologia.

in cui lavorano diminuisce rispetto al primo periodo. Nel caso di Fisica, in particolare, si osserva un peggioramento rispetto al primo periodo, sia per quanto riguarda la percentuale di assunti con prestazioni superiori a quelle dei colleghi già in servizio, sia per il contributo alla performance dell'università. Il contrario avviene in Matematica, dove entrambi gli indicatori migliorano rispetto al primo periodo.

5.2.2. Prospettiva esterna

La Tab. 5 mostra che i valori di R2.1 sono maggiori di 1 in tutte le ADU in entrambi i periodi, il che significa che i neoassunti ottengono performance medie superiori a quelle di tutti i professori di pari livello accademico nello stesso SSD a livello nazionale. Tuttavia, il confronto rivela che il valore medio di R2.1 diminuisce complessivamente e per tutte le ADU passando dal primo al secondo periodo, con l'eccezione dell'ADU 11 (Psicologia). Al contrario, i valori di R2.2 aumentano complessivamente e per tutte le ADU, ad eccezione di Fisica (ADU 2)¹³. In Fisica la situazione

Tab. 5 - Valori medi dei due indicatori per la prospettiva esterna, per ADU

ADU*	R2.1		R2.2	
	08-12	13-17	08-12	13-17
1	1.754	1.581	52.3%	53.3%
2	1.564	1.302	58.3%	47.9%
3	1.833	1.388	46.7%	54.4%
4	1.787	1.474	59.2%	60.7%
5	1.503	1.462	44.3%	50.5%
6	1.652	1.508	44.7%	49.3%
7	1.571	1.509	48.4%	52.7%
8	1.857	1.454	49.2%	55.7%
9	1.504	1.384	49.6%	51.9%
11	1.466	1.474	37.7%	48.4%
<i>Totalle</i>	<i>1.622</i>	<i>1.453</i>	<i>47.8%</i>	<i>51.5%</i>

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell'informazione, 11 – Psicologia.

13. Da notare che, al di là delle differenze tra i due periodi osservati, i valori molto prossimi al 50% per R2.2 per quasi tutte le ADU attestano che mediamente metà dei nuovi assunti ha un profilo scientifico migliore rispetto ai colleghi nazionali già in ruolo e metà no.

peggiora non solo in termini di percentuale di nuovi assunti con performance superiori a tutti gli accademici italiani della disciplina, ma anche a livello di performance rispetto alla media di tutti gli accademici italiani della disciplina. Il contrario avviene in Psicologia, dove entrambi gli indicatori migliorano rispetto al primo periodo.

Fisica si distingue quindi come l'unica disciplina in cui l'efficacia del reclutamento diminuisce sia dal punto di vista interno che esterno.

5.2.3. Promozioni interne e reclutamento esterno

Per valutare le differenze nell'efficacia del reclutamento per tipologia, distinguiamo i nuovi assunti tramite promozione interna (IP) da quelli tramite tutte le tipologie “esterne”, cioè i nuovi assunti nel sistema accademico e i promossi e trasferiti da altre università (rispettivamente OR, OP, TR). La Tab. 6 presenta i risultati e, in particolare, il confronto tra neoassunti interni ed esterni in ciascun periodo e tra i periodi, rispetto al personale strutturato in servizio nelle singole università e all'intera popolazione accademica italiana. Ricordiamo che nella sezione precedente abbiamo già osservato che la percentuale di promozioni interne sul totale dei reclutati è aumentata notevolmente nel secondo periodo, aprendo così alla possibilità di una minore efficacia nel reclutamento.

In entrambi i periodi osserviamo che le promozioni interne contribuiscono meno di quelle esterne alla performance universitaria assoluta nei SSD relativi (valori di R1.1). L'entità del fenomeno aumenta nel secondo periodo, quando il forte aumento della percentuale di promozioni interne sembra ridurre l'efficacia complessiva del reclutamento. Infatti, nel secondo periodo si osserva che le assunzioni esterne, rispetto alle promozioni interne, hanno un impatto positivo più che doppio sulla performance media. Nel primo periodo la percentuale di promozioni interne con performance superiori alla media del personale in servizio era maggiore di quella degli assunti dall'esterno (valori di R1.2), ma nel secondo periodo si verifica il contrario, a ulteriore conferma del peggioramento dell'efficacia del reclutamento.

Secondo invece la prospettiva esterna, confrontando i promossi interni con i reclutati esterni nel primo periodo, entrambi ottengono performance superiori al resto della popolazione accademica italiana nello stesso SSD; tuttavia i promossi interni fanno registrare risultati migliori per questo indicatore (valori di R2.1). Nel secondo periodo si verifica invece il contrario. Inoltre, nel primo periodo, sia gli assunti interni che quelli esterni hanno una performance media intorno alla media dell'intero sistema universitario italiano (valore di R2.2), ma nel secondo periodo la percentuale

di promossi interni con performance superiori alla media scende a circa il 45% mentre quella degli assunti esterni sale a oltre il 55%, confermando ulteriormente una diminuzione dell'efficacia delle politiche di reclutamento complessive.

Tab. 6 - Valori medi degli indicatori per la valutazione delle assunzioni in periodi successivi

	Promozioni interne		Altri	
	2008-12	2013-17	2008-12	2013-17
R1.1	5.35	2.21	6.44	5.15
R1.2	58.9%	63.7%	54.3%	66.1%
R2.1	1.708	1.393	1.532	1.743
R2.2	50.7%	50.7%	44.7%	55.1%

Altri: Promozioni esterne + Trasferimenti + Assunzioni dall'esterno del sistema accademico.

5.2.4. Efficacia del reclutamento a livello di singola università

Confrontiamo ora l'efficacia del reclutamento nei due periodi a livello di singole università, misurando l'entità del miglioramento o del peggioramento dei quattro indicatori.

Dal punto di vista interno, due terzi delle università hanno mostrato un'efficacia di reclutamento decrescente per l'indicatore R1.1 tra il primo e il secondo periodo, mentre un terzo ha mostrato una diminuzione dell'indicatore R2.1. In sostanza due terzi delle università sono state più brave nel primo periodo che nel secondo a reclutare professori in grado di migliorare la performance complessiva di ricerca dell'università. Allo stesso tempo, due terzi delle università hanno avuto più successo nel secondo periodo rispetto al primo nel reclutare professori con prestazioni superiori a quelle del personale di ricerca in forza all'istituzione.

Secondo la prospettiva esterna, invece, il 60% delle università è stato meno efficace nel secondo periodo rispetto al primo nel reclutare professori con prestazioni medie superiori alla media dei colleghi nazionali di pari ruolo e SSD. Tuttavia, sempre nel secondo periodo, il 54,7% delle università è risultato più efficace in termini di percentuale di professori reclutati con prestazioni di ricerca superiori alla media nazionale di FSS.

Le Tabb. 7 e 8 mostrano rispettivamente le università che hanno migliorato la loro posizione e quelle che sono peggiorate in tutti e quattro gli indicatori considerati. In particolare, il gruppo di università che ha ottenuto

i migliori risultati comprende solo quattro università, contro le dieci che sono peggiorate in tutti gli indicatori. Non si notano caratteristiche comuni tra le università di entrambi i gruppi.

Tab. 7 - Università con variazioni positive in tutti gli indicatori, nei due periodi

Università	Delta R1.1	Delta R1.2	Delta R2.1	Delta R2.2
Basilicata	23.3%	1.4%	4.4%	6.4%
Modena e Reggio Emilia	96.6%	29.2%	16.8%	13.3%
SISSA – Trieste	80.3%	33.3%	78.1%	45.0%
“Mediterranea” – Reggio Calabria	21.3%	21.4%	130.1%	36.3%

Tab. 8 - Università con variazione negativa in tutti gli indicatori, nei due periodi

Università	Delta R1.1	Delta R1.2	Delta R2.1	Delta R2.2
Cagliari	-76.5%	-8.0%	-15.9%	-5.2%
Pavia	-4.5%	-0.8%	-28.8%	-7.6%
Trento	-37.3%	-6.9%	-21.1%	-5.4%
Verona	-81.5%	-0.1%	-21.9%	-2.9%
Salento	-73.1%	-24.3%	-28.3%	-14.7%
Milano-Bicocca	-91.6%	-4.0%	-26.1%	-14.0%
Roma Tre	-91.9%	-0.8%	-26.3%	-1.2%
Piemonte Orientale A. Avogadro	-90.5%	-33.6%	-52.9%	-22.0%
Ca' Foscari – Venezia	-57.0%	-12.5%	-3.1%	-6.0%
Sannio – Benevento	-79.4%	-25.1%	-46.4%	-6.7%

6. Conclusioni

In una economia sempre più basata sulla conoscenza, l’istruzione superiore svolge un ruolo fondamentale a sostegno della competitività industriale e dello sviluppo socio-economico.

Nei sistemi di istruzione superiore altamente competitivi, le università di livello mondiale competono tra loro per attrarre i migliori ricercatori e professori, sia nazionali che stranieri. In un ambiente meno competitivo, tipico di molti sistemi d’istruzione europei, questa esigenza è meno sentita. Inoltre, se la scarsa competizione è associata a fenomeni diffusi di favori-

tismo, come avviene in Italia, ne consegue che il merito non è sempre il criterio principale utilizzato per la selezione dei docenti.

Uno degli strumenti per limitare il favoritismo nel reclutamento accademico potrebbe essere quello di monitorare l'efficacia dei relativi processi nelle singole università e di renderlo parte integrante dei sistemi di valutazione della performance degli atenei. In Italia, due iniziative di *policy* sono state recentemente introdotte con la legge 240/2010: l'abilitazione scientifica nazionale, con la funzione di sbarrare l'ingresso ai ruoli accademici ai candidati privi di un adeguato profilo scientifico, e soprattutto l'esercizio di valutazione della ricerca nazionale 2004-2010. Quest'ultimo ha introdotto un incentivo diretto basato sul “reclutamento efficace”: un “bonus” aggiuntivo allo *score* finale della performance universitaria, per i prodotti di ricerca realizzati da professori assunti dall'università nel periodo di osservazione e che hanno ottenuto *score* “eccellenti”. Nonostante le debolezze metodologiche della VQR (Abramo e D'Angelo, 2015c) e la marginalità dell'impatto di tale bonus, un'attenzione supplementare al reclutamento sembrava essere promettente.

In questo capitolo abbiamo confrontato la valutazione della performance dei neoassunti e l'efficacia del reclutamento universitario italiano in corrispondenza all'introduzione della VQR e della nuova “abilitazione scientifica”. I risultati mostrano che la percentuale di neoassunti improduttivi è diminuita, grazie alle soglie introdotte con l'ASN riferite ad indicatori bibliometrici misurati sulla produzione scientifica degli aspiranti docenti accademici. Questo effetto positivo è compensato dalla percentuale di quelli che si collocano al di sopra della mediana, che è significativamente diminuita. Di conseguenza, il rendimento medio complessivo non è aumentato, come dimostrano i valori medi e mediani dei *rating* di produttività delle due coorti osservate.

Per misurare l'efficacia del reclutamento abbiamo utilizzato quattro indicatori, offrendo due prospettive di valutazione: una interna, su scala universitaria; l'altra esterna, su scala nazionale. La prima valuta il reclutamento rispetto alla capacità potenziale dell'università di attrarre e trattenerne risorse umane, in termini di qualità della ricerca del corpo docente già in servizio al suo interno. La seconda valuta l'efficacia in termini assoluti, confrontando la performance di ricerca dei professori assunti con l'intera popolazione dei professori italiani nel medesimo settore e ruolo.

I risultati del confronto non mostrano un miglioramento generale dell'efficacia complessiva del reclutamento. Grazie alle soglie ASN, la percentuale dei neoassunti con performance mediocri è diminuita. Parallelamente, come già osservato in precedenza, è diminuita però anche la percentuale di *performer* superiori alla media, il che ha determinato valori medi e mediani di *rating* delle performance praticamente invariati.

Solo quattro università su sessanta hanno migliorato l'efficacia lungo tutte e quattro le dimensioni considerate, contro ben dieci università che l'hanno peggiorata lungo le stesse dimensioni. Due terzi delle università sono diventati meno efficaci nel reclutare professori le cui performance hanno aumentato i risultati complessivi dell'istituzione nello specifico settore di ricerca, anche se hanno aumentato la percentuale di coloro che registrano prestazioni superiori a quelle dei docenti in servizio. Ciò implica che il reclutamento delle università interessate deve aver portato una percentuale minore di *top performer*, o una maggiore di *very low performer*, o entrambe le cose. Fisica è l'area disciplinare che presenta la maggiore perdita di efficacia del reclutamento. L'inefficacia del reclutamento, nonostante le azioni di *policy* messe in atto, potrebbe essere spiegata dalla percentuale prevalente delle promozioni interne, nel secondo periodo, rispetto alle diverse forme di reclutamento esterno. È difficile credere che nell'83% dei concorsi i candidati interni, tra tutti i partecipanti, abbiano ottenuto i migliori risultati. Il fenomeno può essere in parte spiegato dai tagli ai finanziamenti che il sistema di istruzione superiore italiano ha subito negli ultimi anni (Pruvot, Estermann e Lisi, 2019), per cui le promozioni interne incidono sul bilancio annuale dell'università molto meno delle assunzioni dall'esterno, anche se non aumentano la "capacità" lavorativa.

Un'altra possibile spiegazione è la pratica comune (e diffusa) in Italia che vede i candidati interni essere sostenuti dai gruppi di ricerca a cui appartengono, mentre le posizioni aperte nel concorso sono il frutto di una negoziazione interna che decide quale gruppo avrà la possibilità, in una determinata occasione, di far promuovere il proprio candidato. Questo cosiddetto "candidato interno" sarà in genere sostenuto dai membri interni del comitato di selezione e diventerà quindi il probabile vincitore del concorso. Questo "modus operandi", particolarmente comune per i professori associati e ordinari, ha evidentemente portato all'esclusione di nuovi candidati esterni di alto livello che avrebbero potuto contribuire al miglioramento della performance complessiva di ricerca dell'università reclutante.

La vittoria del "proprio" candidato in un concorso per una qualsiasi posizione accademica diventa una questione d'onore per i cosiddetti "baroni", i professori influenti, che hanno accumulato potere di indirizzo delle decisioni sulla destinazione delle risorse, comprese quelle connesse alla selezione del personale docente. I baroni competono, offrendo favori e intrecciando alleanze, in cambio del sostegno ai propri candidati. Più alto è il potere contrattuale maggiore sarà la probabilità di successo nelle progressioni di carriera; allo stesso tempo si osserva una maggiore richiesta di chi è gerarchicamente collocato nei ruoli inferiori di entrare nel gruppo di

ricerca di un cattedratico influente, contribuendo alla sua visibilità accademica, in cambio di protezione e vantaggi futuri.

Questa struttura ‘feudale’, che caratterizza le dinamiche comportamentali nel contesto dei concorsi nazionali, mette evidentemente a rischio l’efficacia del reclutamento. Un potenziale strumento di contrasto potrebbe essere la creazione di un sistema di incentivi che preveda la misurazione diretta dell’efficacia del reclutamento (Abramo, D’Angelo e Rosati, 2016a) e quindi la premialità (o la penalizzazione) delle parti interessate in funzione della virtuosità (o della vizirosità) delle loro scelte.

Nei sistemi universitari a bassa competitività, dove il terreno per far germogliare buone pratiche di selezione è meno fertile, anche le politiche contro il favoritismo dovrebbero essere più decise e incisive. Una possibile soluzione potrebbe essere quella di aumentare i valori delle soglie di performance per l’abilitazione e di renderne cogente il loro superamento. Questa misura ridurrebbe ulteriormente la percentuale di personale con scarso rendimento tra i nuovi assunti. Parallelamente, si dovrebbe aumentare il peso della performance dei neoassunti sullo *score* finale delle università nella prossima VQR.

Purtroppo, le edizioni successive alla prima VQR non sembrano aver colto questa possibilità. In particolare, la terza edizione, quella 2015-2019 richiedeva alle università di presentare un numero di prodotti pari al triplo del proprio corpo docente, con un limite massimo di quattro per professore. Implicitamente questo requisito ammette la possibilità per un’università di avere fino ad un quarto di professori totalmente improduttivi, senza conseguenze sugli *score* dell’esercizio e, dunque, sui fondi governativi assegnati, a parità di altre condizioni.

Il rigore che l’approccio scientifico al tema richiede comporta l’obbligo di una serie di avvertenze riguardo ai limiti dell’analisi proposta e suggerisce cautela nell’interpretazione dei risultati e delle relative implicazioni e raccomandazioni. Nel misurare la performance di ricerca attraverso l’analisi citazionale si applicano tutti i limiti, i *caveat*, le assunzioni e le riserve consuete della bibliometria, in particolare: i) le pubblicazioni non sono rappresentative di tutta la conoscenza prodotta; ii) i repertori bibliometrici utilizzati (WoS) non censiscono tutte le pubblicazioni; e iii) le citazioni non sono sempre una attestazione dell’uso reale e rappresentativo di ogni tipo di uso/impatto.

Riconosciamo inoltre che il profilo scientifico di un candidato, per come rappresentato dalla sua produzione scientifica, non è l’unica dimensione presa in considerazione da una commissione in un concorso per il reclutamento di un docente universitario. La valutazione dei candidati dovrebbe considerare anche le dimensioni rappresentative delle altre due missioni

istituzionali delle università, ovvero l'insegnamento e la cosiddetta "terza missione"¹⁴. Inoltre, la valutazione delle performance di ricerca in base alla quantità e alla qualità dei risultati trascura altri aspetti dell'attività degli accademici, quali ad esempio la capacità di gestire gruppi di ricerca e i programmi di dottorato, di attrarre fondi, le attività di consulenza, il lavoro editoriale e così via. Anche se ci si aspetterebbe un certo livello di correlazione tra la produttività della ricerca e il livello di queste altre variabili, si raccomanda cautela nell'interpretazione dei risultati qui proposti, che sono limitati alle sole dimensioni relative alla performance di ricerca.

14. Si osserva un crescente impegno verso una "quarta missione", che consiste nell'autovalutazione e nella certificazione delle azioni intraprese. Pur apparendo come un'operazione eticamente corretta e poco onerosa, la vera sfida risiede nel peso crescente delle attività amministrative e di rendicontazione, alimentato da normative nazionali sempre più stringenti.

9. Gli effetti dell'Abilitazione Scientifica Nazionale sulla propensione degli accademici ad auto-citarsi

di *Giovanni Abramo, Ciriaco Andrea D'Angelo, Leonardo Grilli*

1. Introduzione

Un numero crescente di governi nazionali ha introdotto di recente sistemi incentivanti finalizzati ad aumentare la produttività di ricerca nelle università e nelle organizzazioni pubbliche di ricerca. Tra gli strumenti adottati un ruolo sempre più importante riveste la valutazione comparativa della qualità della ricerca, supportata più o meno estensivamente dalla scientometria.

Il crescente ricorso alla valutazione bibliometrica ha portato studiosi e *policy maker* a interrogarsi su eventuali comportamenti opportunistici e sui possibili effetti indesiderati sulla gestione strategica delle istituzioni di ricerca e sul comportamento dei singoli ricercatori.

Nel 2017, *Journal of Informetrics* ha dedicato una sezione speciale agli effetti dei sistemi di valutazione, basati sulle metriche, sul comportamento dei ricercatori (volume 11, numero 3). Il dibattito si è aperto con un documento di discussione di van den Besselaar *et al.* (2017) che contesta le conclusioni dei lavori pionieristici di Linda Butler sugli effetti comportamentali del primo esercizio nazionale di valutazione della ricerca universitaria australiana (Butler, 2003a, 2003b). Secondo la studiosa, la scarsa attenzione alla valutazione dell'impatto della produzione scientifica, ha indotto i ricercatori a pubblicare di più a scapito della qualità delle pubblicazioni. Diversi studiosi hanno contribuito al dibattito (Aagaard e Schneider, 2017; Gläser, 2017; Hicks, 2017; B.R. Martin, 2017).

Secondo il modello dell’“opportunista razionale”, i dipendenti delle organizzazioni anticipano le conseguenze delle loro azioni e adottano comportamenti opportunistici quando i benefici marginali che ne traggono superano i costi marginali (Nagin *et al.*, 2002). Partendo da tale teoria, diversi studiosi hanno ipotizzato e indagato gli effetti perversi sul comportamento di ricercatori e istituzioni, causati dal crescente ricorso alle metriche

nei sistemi di valutazione (Fang *et al.*, 2012; Haustein e Larivière, 2015; de Rijcke *et al.*, 2016). Tra gli effetti indesiderati, si riscontrano la moltiplicazione di pubblicazioni irrilevanti, il plagio, l'auto-plagio e la frode scientifica (Edwards e Roy, 2017; Hazelkorn, 2010; Honig e Bedi, 2012), lo scoraggiamento della diversificazione della ricerca e della ricerca interdisciplinare o innovativa (Hicks, 2012; Rafols *et al.*, 2012; Wilsdon, 2015).

Negli ultimi anni, la sindrome “publish or perish” (Fanelli, 2010; Neill, 2008; van Dalen e Henkens, 2012) si è evoluta nella più sofisticata sindrome “top cited or perish” (Chandler *et al.*, 2002; Gill, 2009; Sa *et al.*, 2013). Gli esercizi nazionali di valutazione della ricerca utilizzano metriche citazionali, come pure le agenzie di finanziamento della ricerca e le commissioni di concorso per il reclutamento di ricercatori e docenti accademici. Alcuni Paesi hanno introdotto perfino bonus economici associati ad indicatori bibliometrici e, in particolare, alla presenza nel proprio curriculum di pubblicazioni ad altro impatto ossia ad elevato numero di citazioni (Abritis e McCook, 2017; Stephan *et al.*, 2017; Tonta e Akbulut, 2020). Ciò ha indotto alcuni studiosi ad indagare su possibili comportamenti opportunistici dei ricercatori volti ad accrescere il numero di citazioni ricevute, attraverso le auto-citazioni (Biagioli, 2016; Scarpa *et al.*, 2018; Baccini *et al.*, 2019; Seeber *et al.*, 2019; D’Antuono e Ciavarella, 2019).

Rispetto alla letteratura, riteniamo che il punto di forza di questo studio sia il metodo di misura delle auto-citazioni. La quasi totalità degli studi sul tema adottano infatti un approccio semplificato, che considera la pubblicazione come l’unità di analisi di base. Secondo tale approccio, l’auto-citazione rappresenta un legame tra due pubblicazioni, citata e citante, che condividono almeno un co-autore. Il nostro approccio, invece, considera l’autore come unità di analisi di base, e definisce una citazione come “auto-citazione” se l’autore della pubblicazione citata è anche autore di quella citante. Non lo fa se è un coautore a citarla. La differenza tra i due approcci è sostanziale, perché il primo considera auto-citazioni anche quelle dei co-autori, inflazionando così il numero di auto-citazioni reali dei singoli autori.

L’applicazione della nostra metodologia comporta il superamento di diversi ostacoli di operazionalizzazione bibliometrica, e inevitabili assunti, tanto maggiori quanto più larga è la scala di analisi. Nel presentare l’approccio metodologico presteremo particolare attenzione a motivare le scelte operative fatte e a esplicitare i limiti nell’operazionalizzazione delle misure, al fine di fornire al lettore tutti gli elementi per valutare la robustezza e l’affidabilità dei risultati, per compararli con quelli di studi precedenti, nonché per facilitarne l’eventuale applicabilità in altri contesti. Come è noto, infatti, spesso i risultati degli studi sono difficilmente comparabili a causa delle diverse metodologie e indicatori utilizzati.

Va detto fin da subito che in genere la risposta opportunistica a sistemi incentivanti è più o meno significativa, concentrata in pochi soggetti o diffusa tra molti, a seconda i) dell'entità dei benefici in gioco, ii) dei costi di attuazione del comportamento opportunistico, iii) del rischio di essere scoperti, iv) delle relative penalizzazioni, e v) del senso etico della comunità scientifica sotto osservazione. Qualsiasi indagine sul comportamento auto-citazionale, come pure le interpretazioni dei risultati che ne scaturiscono, non può prescindere quindi dal contesto di riferimento. Di conseguenza, qualsiasi generalizzazione dei risultati e confronto internazionale richiedono la massima cautela.

Va anche detto che un incremento del tasso di auto-citazione (rapporto tra auto-citazioni e citazioni totali) non sottende necessariamente un comportamento opportunistico. Può anche essere il frutto di una maggiore attenzione dell'autore a citare i propri lavori quando legittimo.

Il presente studio s'inserisce in questo filone di ricerca con l'obiettivo di verificare se e fino a che punto i ricercatori aumentano il tasso di auto-citazione, quando fortemente tentati da schemi incentivanti. Nel caso in cui lo facciano, ci chiediamo anche quanto il fenomeno sia diffuso o concentrato. Un sottoprodotto altrettanto importante di questo studio è l'analisi delle differenze per genere, rango accademico e area geografica.

Per raggiungere il nostro obiettivo, abbiamo confrontato il comportamento auto-citazionale degli accademici in Italia, prima e dopo una significativa riforma del sistema di reclutamento universitario, volta, tra l'altro, a porre un freno al reclutamento di professori carenti sul piano scientifico. La riforma ha imposto che per questo ruolo, tutti i potenziali candidati dovessero superare una fase preliminare di Abilitazione Scientifica Nazionale (ASN). Nelle scienze dure in generale, questo richiede il superamento di valori soglia in un massimo di tre indicatori bibliometrici di performance di ricerca, due dei quali basati sulle citazioni.

Nel paragrafo successivo approfondiremo il tema del comportamento auto-citazionale e le difficoltà che sottendono la misurazione delle auto-citazioni, con l'intento di circoscrivere il campo di quelle auto-citazioni che possono essere considerate solo frutto di un comportamento esclusivamente opportunistico. Nel paragrafo 3, condurremo un'analisi della letteratura sul comportamento auto-citazionale. In particolare, analizzeremo gli approcci metodologici degli studi precedenti e l'affidabilità dei loro risultati. Il paragrafo 4 tratterà dell'abilitazione scientifica nazionale e degli "incentivi" nei settori bibliometrici. Nel paragrafo 5 illustreremo il nostro schema investigativo, presentando il campo di osservazione e le metodologie utilizzate. Il paragrafo 6 riporta i risultati dell'indagine e il 7 le nostre considerazioni e raccomandazioni.

2. Comportamento auto-citazionale e problemi di misurazione

L'approccio della scientometria applicata alla valutazione fonda la sua legittimazione concettuale su fenomeni teorizzati e studiati dai sociologi della scienza. Attraverso una *review* di studi sul comportamento citazionale degli scienziati, Bornmann e Daniel (2008) e successivamente Tahamtan e Bornmann (2019) analizzano le motivazioni che spingono gli scienziati a citare il lavoro di altri. I risultati mostrano che il comportamento citazionale non è motivato solo dal desiderio di riconoscere le influenze intellettuali e cognitive dei colleghi scienziati, ma anche da altri fattori, in parte non scientifici, che giocano un ruolo nella decisione di citare altri lavori. Tuttavia, vi sono evidenze che le diverse motivazioni dei citanti non sono così diverse o casuali da far perdere alla citazione il suo ruolo di misura affidabile dell'impatto scientifico dei lavori citati.

Nella sociologia della scienza sono state sviluppate due diverse teorie del comportamento citazionale: quella normativa e quella costruttivista sociale. La prima, basata su un lavoro fondamentale di Robert Merton (1973a), afferma che gli scienziati, attraverso la citazione di un lavoro scientifico, riconoscono un credito nei confronti di un collega di cui hanno utilizzato i risultati, il che significa che le citazioni rappresentano un'influenza intellettuale o cognitiva sul loro lavoro scientifico.

La seconda trae le sue origini nei lavori di Latour e Woolgar (1979) e Knorr-Cetina (1981). L'approccio costruttivista sociale contesta gli assunti alla base della teoria normativa e quindi la validità del ricorso alle citazioni per la misura dell'impatto scientifico. I costruttivisti sostengono che "la conoscenza scientifica è costruita socialmente attraverso la manipolazione di risorse politiche e finanziarie e l'uso di dispositivi retorici" (Knorr-Cetina, 1981), il che significa che le citazioni non sarebbero necessariamente collegate in modo consequenziale ai contenuti scientifici dell'articolo citato.

Secondo Tahamtan *et al.* (2016) e Tahamtan e Bornmann (2018) esistono diverse ragioni, scientifiche e non scientifiche, per cui una pubblicazione può essere citata. La rilevanza delle caratteristiche pertinenti al documento, all'autore o alla rivista gioca un ruolo fondamentale nel processo di citazione. Le auto-citazioni rappresentano una particolare sottoclasse di citazioni. Il ricorso all'auto-citazione è legittimo quando ci sono buone ragioni scientifiche per auto-citarsi, prima fra tutte quella di riconoscere un merito al proprio lavoro, come si farebbe con gli altri. Gli autori tendono a citare propri lavori precedenti anche per evitare accuse di auto-plagio (Garfield, 1979; Pichappan e Sarasvady, 2002). L'auto-citazione è invece illegittima quando le pubblicazioni citate non rappresentano un vero contributo al contenuto scientifico della pubblicazione citante. Il ricorso a tali auto-cita-

zioni si rivelerebbe funzionale ad aumentare il numero totale delle proprie citazioni, e il loro conteggio rappresenterebbe in modo errato l'importanza del lavoro specifico e anche della rivista ospitante, falsando in ultima analisi qualsiasi valutazione.

Si possono individuare tre ragioni principali per cui i ricercatori ricorrono all'auto-citazione con il fine di aumentare le loro citazioni totali, rafforzando indirettamente la posizione dei socio-costruttivisti sulla manipolazione. Una ragione per gonfiare le auto-citazioni potrebbe essere la tendenza all'autogratificazione personale, nonché all'autoesaltazione e all'autopromozione, nella competizione per la visibilità e l'autorevolezza scientifica (Brysbbaert e Smyth, 2011; Hyland, 2003).

Una seconda ragione ha carattere strategico, in quanto l'auto-citazione può rivelarsi un efficace strumento pubblicitario e promozionale, soprattutto per le pubblicazioni recenti, per aumentare le citazioni future da parte di altri (van Raan, 2008). È stato stimato che "ogni auto-citazione aggiuntiva aumenta il numero di citazioni da parte di altri di un'unità dopo un anno e di circa tre dopo cinque anni. Inoltre, non c'è una penalizzazione significativa per chi si auto-cita più spesso: l'effetto dell'auto-citazione rimane positivo anche per tassi di auto-citazione molto elevati (Fowler e Aksnes, 2007).

Una terza ragione è una risposta a meccanismi incentivanti che premiano alti livelli citazionali. Trattasi di comportamento opportunistico.

Quando si studiano eventuali effetti perversi associati a sistemi di valutazione bibliometrica della ricerca, onde evitare risultati e interpretazioni distorti, occorre distinguere le auto-citazioni che scaturiscono da una risposta ai meccanismi incentivanti (comportamento opportunistico), da altri tipi di auto-citazione illegittima e dalle auto-citazioni perfettamente legittime.

Qualunque sia il motivo che la sottende, l'auto-citazione illegittima può essere diretta o indiretta. Le auto-citazioni si dicono "dirette" quando l'autore di un articolo cita propri articoli precedenti. Le auto-citazioni "indirette" sono a loro volta di due tipi: i) "indotte in modo coercitivo", quando un manipolatore, come un consulente scientifico, un revisore di manoscritti o di domande di finanziamento, induce altri ricercatori a citare i propri lavori, anche se questi non contribuiscono al contenuto scientifico degli articoli citanti (Thombs *et al.*, 2015; Ioannidis, 2015); oppure ii) frutto di comportamento collusivo, per cui due o più ricercatori citano reciprocamente i propri lavori (Sala e Brooks, 2008). Sia l'auto-citazione coercitiva indotta sia quella "collusiva", detta anche "citazione incrociata", sono molto difficili da identificare in particolare nelle analisi bibliometriche su larga scala. Questo problema rappresenta un limite nella maggior parte degli studi sul comportamento auto-citazionale.

Nello studio del comportamento auto-citazionale diretto, una questione critica è la definizione di auto-citazione e la conseguente operazionalizzazione della misurazione. La letteratura fornisce due definizioni. La prima applica il senso comune: una citazione è definita come auto-citazione ogni volta che l'autore di un articolo citato (in cotitolarità o meno) è anche (co) autore dell'articolo citante. Operativamente, questa definizione implica l'adozione dell'autore come unità di base dell'analisi. Per comodità, chiameremo questo approccio “basato sull'autore”.

Un secondo approccio parte invece dalla pubblicazione: una citazione viene definita auto-citazione ogni volta che l'insieme dei coautori dell'articolo citato e di quello che viene citato non sono disgiunti, ossia che i due insiemi condividono almeno una persona (Snyder e Bonzi, 1998). Questa definizione, tuttavia, va oltre il reale significato di auto-citazione, poiché un singolo autore potrebbe risultare un auto-citatore seriale anche se non ha mai auto-citato i propri lavori. Ciò si verificherebbe quando, tra i gruppi di coautori di pubblicazioni, vi sono auto-citatori seriali: in tal caso, un autore “importerebbe” le auto-citazioni dalle azioni degli altri. In termini operativi questa definizione implica l'adozione della pubblicazione e della relativa lista autori come unità di base dell'analisi, per questo chiamiamo questo approccio “basato sulla pubblicazione”.

Naturalmente questo secondo approccio porta a un conteggio delle auto-citazioni superiore rispetto al primo. In effetti, in un macrostudio su oltre mezzo milione di citazioni di articoli di ricercatori norvegesi, indicizzati nel Science Citation Index, Fowler e Aksnes (2007) hanno riscontrato che l'approccio basato sulla pubblicazione attribuisce tassi di auto-citazione notevolmente più elevati (21%) rispetto a quello che considera l'autore come unità analitica di base (11%).

Gli studiosi che adottano l'approccio basato sulla pubblicazione spesso lo giustificano con l'intento di svelare le citazioni incrociate indirette collusive. Nella maggior parte dei casi, però, sospettiamo che il motivo vero sia la maggiore facilità operativa nell'individuare e misurare le auto-citazioni. L'adozione della definizione basata sull'autore, infatti, presuppone la disambiguazione iniziale di tutti i nomi degli autori, che richiede un lungo lavoro manuale o lo sviluppo di algoritmi di disambiguazione. La definizione basata sulla pubblicazione offre invece un'operatività molto più semplice, in quanto richiede solo l'identificazione dell'intersezione dei due insiemi di autori delle pubblicazioni citate e citanti, come riportati nella lista autori. Questo spiegherebbe perché pochissimi studi hanno utilizzato l'approccio basato sull'autore (Fowler e Aksnes, 2007; King *et al.*, 2017; Seeber *et al.*, 2019).

La nostra personale convinzione è che l'approccio basato sull'autore sia quello appropriato, a meno che non si pensi che la comunità scientifica

globale comprenda in qualche modo cartelli di coautori ampiamente interconnessi, impegnati a coltivare citazioni incrociate (Fister *et al.*, 2016). In realtà, l'obiettivo di indagare le citazioni incrociate collusive sarebbe meglio perseguito utilizzando tecniche di analisi delle reti di co-titolarità (González-Teruel *et al.*, 2015; Zhao e Strotmann, 2015).

Quando si analizza la letteratura sul tema e si interpretano e confrontano i risultati è fondamentale altresì distinguere tra due diverse metodologie operative di conteggio delle auto-citazioni. Le auto-citazioni possono infatti essere contate in due modi diversi, in modo sincrono (guardando indietro) o diacronico (guardando avanti) (Costas *et al.*, 2010; Glänzel *et al.*, 2004; Lawani, 1982). Prendendo un dato articolo, l'approccio sincrono conta le auto-citazioni dell'autore estraendole dall'elenco degli articoli citati in bibliografia. In teoria, il conteggio totale osservato è deterministico. Questo approccio, tuttavia, comporta la disambigua-zione della paternità di tutte le pubblicazioni precedenti dell'autore, che potrebbe comportare analisi ultradecennali, il che rappresenta un compito tutt'altro che semplice. Nella pratica operativa la ricerca delle pubblicazio-ni citabili viene spesso interrotta a una certa data, il che rende il conteg-gio delle auto-citazioni una previsione.

L'approccio diacronico viceversa, conta le auto-citazioni ricevute da un dato articolo. Guardando in avanti, per definizione, il conteggio delle auto-citazioni dipende dalla scelta della finestra temporale. Anche questo metodo offre quindi anche una previsione, tanto più accurata quanto più ampia è la finestra citazionale.

3. La letteratura sul comportamento auto-citazionale

Dato il ruolo centrale delle citazioni nelle valutazioni bibliometriche e l'effetto distorcente che le auto-citazioni illegittime possono avere, i bi-bliometristi avvertono una crescente pressione a formulare degli indici che permettano di individuare gli auto-citanti estremi. Singoli indici sono stati proposti o utilizzati, per lo più denominati tasso di auto-citazione, ma ove numeratore o denominatore possono essere diversi a seconda della formu-lazione. Il numeratore comune a tutte può cambiare a seconda della defi-nizione di auto-citazione adottata. Al variare del denominatore possiamo avere poi tassi di auto-citazione definiti come il rapporto del numero totale di auto-citazioni sul numero totale di citazioni ricevute o del numero totale di pubblicazioni dell'autore; il rapporto medio del numero di auto-citazioni sul numero totale di citazioni nella bibliografia di ciascuna pubblicazione dell'autore; il numero medio di citazioni per pubblicazione.

Con grande creatività, è stato proposto addirittura un *h*-index per gli auto-citanti, denominato *s*-index, ove *s* è il numero di articoli con almeno *s* auto-citazioni dell'autore (Flatt *et al.*, 2017).

Le prime classifiche di scienziati per tasso di auto-citazione hanno cominciato ad apparire di recente. Ioannidis *et al.* (2019) hanno misurato il rapporto tra auto-citazioni e citazioni totali ricevute dei 100.000 ricercatori più citati al mondo. Sfortunatamente il conteggio delle auto-citazioni include impropriamente anche quelle da parte dei coautori, con potenziali significative distorsioni delle classifiche. Ad esempio, scorrendo i dati messi a disposizione dagli autori delle classifiche, emerge il caso di Sonia Del Prete, ricercatrice presso l'Istituto di Bioscienze e Biorisorse del Consiglio Nazionale delle Ricerche. I suoi articoli pubblicati nel 2017 hanno ricevuto nel medesimo anno 456 citazioni totali di cui il 92% sarebbero auto-citazioni secondo Ioannidis *et al.* (2019). Dei 106.000 ricercatori più citati al mondo, solo 8 mostrano un tasso di auto-citazione maggiore. Un'attenta verifica manuale dei dati rivela che in realtà le auto-citazioni reali di Sonia Del Prete sono meno del 25%, e che nel 97% delle sue pubblicazioni compare tra i coautori un suo collega dell'Università di Firenze, molto più prolifico, le cui auto-citazioni reali rappresentano oltre il 56% delle citazioni totali ricevute. Trattasi di un emblematico caso di auto-citazioni “importate”, dovuto all'approccio improprio di misura delle auto-citazioni, che esporrebbe alla gogna della comunità scientifica uno scienziato “innocente”.

Va detto che gli stessi autori di tale classifica raccomandano di non utilizzarla per esporre al pubblico ludibrio quei ricercatori con tassi di auto-citazione particolarmente elevati, né per inferire soglie per distinguere i ricercatori disonesti dagli onesti. Lo studio però potrebbe aiutare a identificare, dopo accurata ulteriore verifica, gli auto-citanti estremi e, possibilmente, i “cartelli” di citanti incrociati.

La nostra posizione sull'argomento è che l'adozione di singoli indicatori di tassi di auto-citazione, per studiare il comportamento auto-citazionale, può essere molto rischiosa in quanto induce a trarre conclusioni fuorvianti. Infatti sono molti i fattori che potrebbero spiegare perché uno studioso legittimamente si cita più di altri, il che richiede l'adozione di modelli inferenziali di analisi per poterli controllare. La letteratura ci fornisce utili indicazioni su quali possano essere tali fattori.

Ci sono diversi studi sulle auto-citazioni che hanno preso in esame un numero limitato di discipline e un numero ristretto di articoli, ma pochissimi studi su larga scala e ancora meno che hanno preso in considerazione i comportamenti sottostanti. Di seguito passiamo in rassegna quanto disponibile, alla ricerca di informazioni utili per formulare il modello statistico

necessario a rispondere alla nostra specifica domanda di ricerca: gli scienziati aumentano i tassi di auto-citazione quando sono fortemente tentati da schemi di incentivazione e, se sì, in che misura?

3.1. Fattori che influenzano i tassi di auto-citazione

I ricercatori tendono ad auto-citare le pubblicazioni recenti più di quelle meno recenti e la percentuale di auto-citazioni è più bassa per le pubblicazioni molto citate (Aksnes, 2003). L'autore riscontra anche una forte correlazione positiva tra il numero di auto-citazioni e il numero di autori delle pubblicazioni. È anche noto che le pubblicazioni con coautori internazionali sono più citate di quelle nazionali (Adams, 2013; Kumar *et al.*, 2014), e questo vale anche per le pubblicazioni con coautori italiani (Abramo, D'Angelo e Murgia, 2017) anche quando si includono nel conteggio le auto-citazioni italiane (Abramo, D'Angelo e Di Costa, 2021).

I ricercatori si differenziano per il grado di diversificazione della loro attività di ricerca (Abramo, D'Angelo e Di Costa, 2017; Stirling, 2007) ed è lecito attendersi che, a parità di altre condizioni, coloro che svolgono ricerca più specializzata tendano ad auto-citarsi di più.

Le diverse discipline mostrano differenze significative nei tassi di auto-citazione: alti nelle STEM, più bassi nelle scienze sociali e ancora più bassi nelle scienze umane (Aksnes, 2003; Snyder e Bonzi, 1998). Ciò suggerisce la necessità di condurre analisi che tengano conto del campo di ricerca, per valutare la propensione all'auto-citazione opportunistica.

Sembra che gli uomini tendano a citare i propri articoli il 70% in più delle donne (King *et al.*, 2017; Mishra *et al.*, 2018), in parte perché producendo più articoli delle donne hanno più materiale da citare (Larivière *et al.*, 2013; De Kleijn, M., Jayabalasingham, B., Falk-Krzesinski, H.J., Collins, T., Kuiper-Hoyng, L., Cingolani, I., Zhang, J., Roberge, G., 2020), e questo vale sicuramente anche per l'Italia (Abramo, Aksnes e D'Angelo, 2020). Per la Danimarca, invece, Nielsen (2016) non ha rilevato differenze statisticamente significative nei tassi di auto-citazione tra uomini e donne, tranne che nelle scienze mediche. Ciò suggerisce che, nell'esaminare la risposta comportamentale all'ASN, vanno considerate anche le differenze di genere.

Sembra che le citazioni per articolo varino a seconda del rango accademico degli autori (Ventura e Mombrú, 2006). In Italia, la media delle citazioni per articolo dei professori ordinari è leggermente superiore a quella dei professori associati e notevolmente superiore a quella dei ricercatori (Abramo, D'Angelo e Di Costa, 2011). Alla luce di ciò, è interessante chie-

dersi in che misura il comportamento auto-citazionale vari a seconda del rango accademico.

Le organizzazioni sono caratterizzate da diverse identità, sistemi di valori e codici etici distinti che servono a guidare il comportamento dei dipendenti (Collins e Porras, 1991). Potremmo quindi aspettarci che tra le università ci siano propensioni diverse ad aumentare artificialmente le auto-citazioni. L'Italia poi è caratterizzata da sensibili differenze culturali, sociali ed economiche tra aree geografiche (Cafagna, 1989; Daniele, 2015), e queste differenze sono accompagnate anche da un notevole divario nei risultati della ricerca (Abramo, D'Angelo e Rosati, 2016b).

Questo suggerisce un'analisi anche per area geografica.

3.2. Effetti della valutazione bibliometrica sulle auto-citazioni

La letteratura sugli effetti perversi della valutazione bibliometrica sul comportamento degli scienziati in generale è piuttosto ricca, ma è piuttosto scarsa quella che riguarda gli effetti di specifici sistemi di incentivazione sulle auto-citazioni. Abbiamo trovato solo cinque studi strettamente incentrati sull'argomento e, sorprendentemente, tutti riguardano il campo di osservazione italiano. Scarpa *et al.* (2018) hanno analizzato gli effetti sui tassi di pubblicazione e di auto-citazione dell'esercizio nazionale di valutazione della qualità della ricerca (VQR) e dell'ASN. Lo studio è limitato ai soli professori di ingegneria (7151 in totale). Per ogni professore, nel periodo 2006-2016, gli autori hanno misurato il tasso annuale di auto-citazione, definito come il totale delle auto-citazioni diviso per il totale delle citazioni ricevute. Sulla base di questo indicatore, l'analisi descrittiva ha mostrato un notevole aumento del tasso medio di auto-citazione. Gli autori concludono che “l'influenza dannosa dell'introduzione degli esercizi di valutazione è chiaramente visibile analizzando la risposta auto-citazionale degli accademici. Questo risultato suggerisce il ricorso a indici di impatto che siano scarsamente sensibili a tale fenomeno”. Lo studio presenta a nostro avviso importanti limiti, tra cui: i) la mancanza di un'analisi inferenziale che consideri altri fattori, oltre al desiderio di gonfiare le proprie citazioni, che potrebbero giustificare l'aumento delle auto-citazioni; ii) l'osservazione di una sola disciplina.

Seeber *et al.* (2019) hanno condotto un'indagine simile su un totale di 866 professori appartenenti a quattro settori scientifici disciplinari: Ingegneria economico-gestionale, Genetica, Psichiatria ed Economia applicata (quest'ultimo come settore di controllo). Combinando statistiche descrittive e inferenziali l'analisi mostra l'evoluzione nel tempo delle auto-citazioni tra il 2002 e il 2014. Gli autori hanno concluso che l'introduzione dell'ASN nel

2012, che lega la possibilità di avanzamento di carriera al numero di citazioni ricevute è correlata a un forte e significativo aumento delle auto-citazioni tra gli accademici che possono beneficiare maggiormente dell'aumento delle citazioni, ossia i ricercatori e i professori associati e, in particolare, quelli relativamente meno citati. A nostro avviso, sebbene l'indagine abbia applicato un approccio empirico rigoroso, il campo di applicazione (tre settori scientifici su 109 individuati, per l'Italia) è forse troppo ristretto per giungere a conclusioni generalizzabili.

Anche Baccini *et al.* (2019) hanno studiato l'impatto dell'ASN sul comportamento auto-citazionale. A tal fine hanno ideato un indicatore a livello Paese, “appositamente progettato per essere sensibile agli effetti sia dell'uso opportunistico dell'auto-citazione diretta da parte degli autori sia della creazione di reti citazionali”. L'indicatore, denominato “inwardness”, è definito come il rapporto tra il numero totale di auto-citazioni di un Paese e il totale di tutte le citazioni ricevute dallo stesso Paese. Gli autori hanno quindi confrontato l'evoluzione 2000-2016 dell'inwardness in Italia con quella degli altri Paesi del G10. Osservando un aumento significativo dell'inwardness dopo il lancio dell'ASN, rispetto all'evoluzione degli altri Paesi, hanno concluso che la comunità scientifica italiana ha adottato strategie opportunistiche come risposta all'ASN.

Questo studio ha attirato un'ampia attenzione (Guglielmi, 2019; Huet, 2019; Singh Chawla, 2019; Van Noorden, 2019; Magnani, 2019), tuttavia soffre di limiti nella metodologia e nell'indicatore stesso. Tra gli altri: i) le auto-citazioni nazionali sono definite come citazioni provenienti dall'interno del Paese, ma anche da tutti i Paesi di eventuali coautori, indipendentemente dal fatto che gli autori stranieri dei lavori citanti fossero davvero i coautori dell'articolo citato¹. Questo implicherebbe che la comunità scientifica mondiale è organizzata come cartelli a livello di Paesi per realizzare comportamenti collusivi attraverso citazioni incrociate. ii) Anche quando le auto-citazioni diminuiscono, il punteggio di inwardness del Paese potrebbe aumentare, nel semplice caso in cui il denominatore, il totale delle citazioni ricevute, diminuisca di più. iii) L'ASN utilizza indicatori basati sulle citazioni per la valutazione dei candidati accademici solo nelle scienze *dure*, e per questi potrebbe essere un incentivo a manipolare le citazioni. Tuttavia, lo studio applica l'indicatore di *inwardness* anche alle scienze sociali e alle discipline umanistiche, ossia a pubblicazioni di un'ampia comunità molto poco interessata all'incremento dei propri indicatori bibliometrici, compresi

1. A titolo esemplificativo, tutte le citazioni da pubblicazioni statunitensi a una pubblicazione di cui l'Italia è coautrice insieme agli USA, sono considerate come auto-citazioni italiane.

i ricercatori del settore privato e quelli della sfera pubblica non accademica.
iv) Lo studio manca di un'analisi inferenziale che controlli tutti quei fattori che potrebbero concorrere a un incremento delle auto-citazioni.

D'Antuono e Ciavarella (2019) hanno utilizzato il suindicato dataset dei 100.000 ricercatori più citati al mondo (Ioannidis *et al.*, 2019) per confrontare i tassi di auto-citazione dei *top scientist* italiani con quelli degli altri scienziati nazionali. Sulla base dei risultati, gli autori hanno concluso che “il problema delle auto-citazioni, in relazione a questa élite scientifica, non è significativo in Italia”.

Infine, Peroni *et al.*, (2020) hanno analizzato oltre 35.000 articoli pubblicati tra il 1957 e il 2016 dai candidati all'ASN 2012, evidenziando un aumento complessivo delle auto-citazioni e una forte relazione causale con l'ASN in 10 delle 24 discipline considerate. Il metodo adottato però per il conteggio delle citazioni è quello basato sulla pubblicazione come unità di base per le analisi.

4. Abilitazione Scientifica Nazionale e “incentivi” nei settori bibliometrici

Come già detto in precedenza la Legge 240 del 2010, ha definito la politica e i criteri per un'ampia riforma del sistema universitario stabilendo, nell'ambito del reclutamento, che i concorsi per le posizioni di professore siano accessibili solo ai candidati precedentemente qualificati secondo lo schema ASN. Tale qualifica è affidata ad una commissione di esperti settoriali che si avvale, tra l'altro, di tre indicatori bibliometrici, di cui due basati sulle citazioni. A seconda del settore concorsuale il superamento di una, due o tre soglie bibliometriche è un prerequisito per l'accreditamento, rappresentando pertanto un forte incentivo all'adozione di un comportamento opportunistico al fine di superare le soglie.

Nel dettaglio, secondo il quadro normativo, ogni due anni il Ministero dell'Università e della Ricerca avrebbe dovuto emanare un bando per la nomina di 184 Commissioni di Accreditamento, ossia una commissione per ogni Settore Concorsuale (SC). I SC derivano dall'aggregazione dei cosiddetti Settori Scientifico Disciplinari (SSD, 370 in tutto), che servono in generale per la classificazione e l'amministrazione di tutti gli accademici italiani. Secondo questo sistema, ogni professore è classificato in modo univoco, in funzione del suo campo di ricerca, in uno specifico SSD. I SSD, e quindi anche le loro aggregazioni in SC, sono raggruppati in 14 Aree Disciplinari Universitarie (ADU). Dei 184 SC, l'ANVUR (l'agenzia nazionale di valutazione italiana) ne ha classificati originariamente 109

come “bibliometrici” e i restanti 75 come “non bibliometrici”. Per i SC bibliometrici, i candidati membri delle prime commissioni di accreditamento sono stati scelti a caso tra tutti i professori ordinari che superavano le tre soglie bibliometriche.

I bandi per l'accreditamento sono stati invece originariamente emessi su base annuale, e solo più recentemente su base trimestrale. I candidati devono presentare domanda per specifici SC e fascia accademica, ma non c'è limite al numero di SC richiesti, e i candidati possono richiedere contemporaneamente l'accreditamento a professore associato e ordinario. Le candidature devono essere accompagnate da un curriculum vitae e da un elenco dei titoli e delle pubblicazioni scientifiche. Per i SC bibliometrici l'elenco delle pubblicazioni è servito, tra l'altro, per il calcolo dei tre indicatori bibliometrici suddetti². Per i candidati all'ASN 2012, essi sono stati³:

- numero di articoli su riviste nel periodo 2002-2012, con normalizzazione in caso di anzianità accademica inferiore a 10 anni (calcolata dal primo anno di pubblicazione);
- numero di citazioni ricevute per la produzione scientifica complessiva, normalizzato per l'anzianità accademica del candidato;
- h-index contemporaneo⁴ (Sidiropoulos *et al.*, 2007) della produzione scientifica complessiva.

I suddetti indicatori sono stati calcolati sulla base delle pubblicazioni indicizzate nei due maggiori archivi bibliografici: Scopus e Web of Science (WoS). Nel caso di pubblicazioni indicizzate in entrambe le fonti, il valore delle citazioni assegnate era il più alto tra i due. Sulla base di un database di tutte le pubblicazioni inserite dai professori italiani nel corso del 2012, l'ANVUR ha proceduto a calcolare i diversi indicatori per i professori ordinari e associati e ha poi pubblicato i valori mediani, utilizzati sia per la selezione dei membri della commissione sia per le loro successive valutazioni dei candidati.

La Tab. 1 riporta le statistiche descrittive del bando di accreditamento 2012. La risposta è stata di oltre 59.148 domande (55,7% già in organico), di cui più di un sesto per Medicina (9.987). Le richieste di accreditamento per professore ordinario sono state il 30,5% del totale.

2. Per i SC “non bibliometrici”, gli indicatori non erano basati sulle citazioni.

3. Per i candidati a posizioni di membro di commissione, gli indicatori erano il numero e l'impatto della produzione scientifica complessiva, senza normalizzazione per gli anni di anzianità accademica.

4. Di fatto normalizza l'h-index, pesando maggiormente i lavori più recenti.

*Tab. 1 - Statistiche dell'Abilitazione Scientifica Nazionale 2012 (elaborate dagli autori su dati pubblicati sulle pagine ASN del sito MUR <http://abilitazione.miur.it/public/candidati.php?sersel=50&>)**

ADU	SC	Di cui “bibliometrici”	Domande	Di cui per I fascia
01 – Matematica e informatica	7	7	2,492	911 (36.6%)
02 – Fisica	6	6	4,372	1,451 (33.2%)
03 – Chimica	8	8	2,344	695 (29.7%)
04 – Scienze della Terra	4	4	1,231	400 (32.5%)
05 – Biologia	13	13	6,244	1,690 (27.1%)
06 – Medicina	26	26	9,987	3,298 (33.0%)
07 – Scienze agrarie e veterinarie	14	14	2,093	650 (31.1%)
08 – Ingegneria civile e architettura	12	7	3,599	1,027 (28.5%)
09 – Ingegneria industriale e dell’informazione	20	20	4,535	1,573 (34.7%)
10 – Storia antica, filologia, letteratura e arte	19	0	6,324	1,718 (27.2%)
11 – Storia, filosofia, pedagogia, psicologia	17	4	5,909	1,491 (25.2%)
12 – Scienze giuridiche	16	0	3,037	887 (29.2%)
13 – Economia e statistica	15	0	4,853	1,755 (36.2%)
14 – Scienze politiche e sociali	7	0	2,128	515 (24.2%)
<i>Totale</i>	<i>184</i>	<i>109</i>	<i>59,148</i>	<i>18,061 (30.5%)</i>

* Il MUR ha ritirato la maggior parte dei dati 60 giorni dopo le decisioni delle commissioni di concorso, lasciando solo gli elenchi dei candidati selezionati.

5. Ipotesi di ricerca e specifiche metodologiche

Per i ricercatori e i professori associati che aspirano ad avanzamenti di carriera nei SC bibliometrici e, in misura minore, per i professori ordinari che ambiscono ad essere selezionati quali membri delle commissioni giudicatrici, due dei tre indicatori bibliometrici, basandosi sulle citazioni, rappresentano un potenziale incentivo a “gonfiare” le proprie citazioni. Gli incentivi agiscono ovviamente anche sui candidati non universitari che ambiscono a diventarlo. Tuttavia limitiamo la nostra analisi ai soli accademici, poiché per gli altri non è disponibile un analogo database pubblico che elenca i loro nomi, i settori disciplinari, la fascia accademica, l'affiliazione, ecc.

Dalla nostra domanda di ricerca deriviamo due ipotesi:

Ipotesi 1. I ricercatori e i professori associati nei SC bibliometrici, che ambiscono alla progressione di carriera, aumentano i loro tassi di auto-citazione dopo l'introduzione dell'ASN.

Ipotesi 2. I professori ordinari e i professori associati abilitati al primo bando ASN, che aspirano ad essere inclusi nelle commissioni di accreditamento, essendo soggetti a un incentivo più debole, aumentano in misura minore i loro tassi di auto-citazione.

Per verificare le ipotesi, analizziamo il comportamento auto-citazionale dei professori nei cinque anni precedenti l'introduzione dell'ASN e nei cinque anni successivi. Poiché i criteri e gli indicatori di accreditamento sono stati pubblicati per la prima volta nel giugno 2012, è ragionevole attendersi che le prime eventuali evidenze di un'alterazione dei tassi di auto-citazione occorrono nel 2013⁵. Analizziamo quindi il numero annuale di auto-citazioni indicizzate in WoS, da parte di professori italiani in SC bibliometrici, nei 5 anni precedenti e successivi a tale data, ossia nei periodi 2008-2012 e 2013-2017. Eventuali aumenti repentini delle auto-citazioni dopo il 2012, a parità di altre condizioni, segnalerebbero possibili comportamenti opportunistici da parte dei soggetti interessati.

5.1. *Il dataset di analisi*

Secondo la base dati del personale di ricerca accademico curata dal MUR⁶, alla fine del 2012 i ricercatori e professori nelle università italiane erano 57.400, di cui 35.242 appartenenti a SC bibliometrici.

Per la costruzione del dataset bibliometrico di analisi abbiamo utilizzato l'algoritmo di disambiguazione del nome dell'autore sviluppato da D'Angelo, Giuffrida e Abramo (2011), basato sull'accoppiamento delle pubblicazioni estratte dall'Italian National Citation Report (I-NCR) di WoS e dalla base dati del MUR. Questo algoritmo⁷ assegna una pubblicazione I-NCR a un dato professore⁸ se quest'ultimo:

5. È altamente improbabile che si possa scrivere un articolo, superare il processo di valutazione da parte della rivista, e ottenere la pubblicazione in meno di sei mesi.

6. <http://cercauniversita.cineca.it/php5/docenti/cerca.php>.

7. La media armonica di precisione e recupero (misura F) delle *authorship* disambigue dall'algoritmo è di circa il 97% (margine di errore del 2%, intervallo di confidenza del 98%).

8. Ove non sia necessario discriminare, nel seguito utilizzeremo il termine generico professore per indicare sia i professori sia i ricercatori.

- ha un nome compatibile con uno degli autori della pubblicazione;
- appartiene a una delle università riconosciute nell'elenco degli indirizzi indicati dagli autori della pubblicazione;
- appartiene a una SC compatibile con la categoria tematica (*subject category* WoS) della pubblicazione;
- ricopre una posizione di ruolo al 31 dicembre dell'anno precedente a quello della pubblicazione.

L'I-NCR a nostra disposizione è composto da tutte le pubblicazioni 2001-2017⁹ indicate nella *core collection* WoS, con “Italia” come paese di appartenenza di almeno un autore. Il dataset di analisi è composto dal personale di ricerca universitario in ruolo nel periodo 2008-2017, che afferisce a un SC bibliometrico e che è stato in ruolo anche nel periodo 2001-2007¹⁰ (15.037 unità in tutto). La Tab. 2 mostra la ripartizione di questo database per area disciplinare.

In questo lavoro definiamo il tasso di auto-citazione di un autore come il numero di auto-citazioni per pubblicazione. I dati bibliometrici (pubblicazioni, auto-citazioni e loro rapporto) di ciascun professore si riferiscono a un panel di 10 anni (Tab. 3). La struttura del dataset è sbilanciata, poiché il tasso di auto-citazioni non può essere definito negli anni in cui un professore ha zero pubblicazioni. Ne consegue che 6.224 professori hanno osservazioni su tutti i 10 anni e il resto ha almeno un anno con un valore mancante. Complessivamente, sono disponibili 120.615 anni-professore.

I 15.037 professori hanno prodotto un totale di 563.457 pubblicazioni, distribuite per ADU e anno come indicato nelle colonne 2 e 5 della Tab. 3. Le pubblicazioni contengono un totale di 1.062.559 auto-citazioni, con una media di 1,74 auto-citazioni per pubblicazione nel periodo 2008-2012 e di 2,00 nel periodo 2013-2017 (colonne 4 e 7 della Tab. 3). Nel primo periodo, l'ADU con il tasso medio di auto-citazione più alto è stata Fisica (3,18), seguita da Chimica (2,80) e Biologia (2,24). In tutte le altre ADU, il tasso medio è stato inferiore a 2, con il minimo in Ingegneria industriale e dell'informazione (0,86). Nel secondo periodo il tasso medio di auto-citazione è più alto in tutte le discipline. Le ADU con i tassi più alti e più bassi rimangono invariate (Fisica a 3,87; Ingegneria industriale e dell'informazione a 1,07). L'aumento maggiore si registra in Ingegneria

9. Sono inclusi articoli, articoli di rassegna, lettere e atti di conferenze; sono esclusi editoriali, *abstract* di riunioni, ristampe e qualsiasi altro tipo di documento che non contenga risultati di ricerca.

10. Questa ulteriore condizione si rende necessaria per la misurazione del tasso di auto-citazione, che richiede la conoscenza della produzione scientifica negli anni precedenti le auto-citazioni. Ricordiamo che la nostra disambiguation delle *authorship* riguarda solo il personale universitario.

civile (da 1,07 auto-citazioni per pubblicazione nel primo periodo a 1,61 nel secondo, ossia +50%), seguita da Scienze agrarie e veterinarie (1,25 contro 1,67, ossia +34%).

Tab. 2 - Dataset di analisi

ADU*	N. di SC	Numero di SSD	N. di Professori**			
			Ordinari	Associati	Ricercatori	Totale
1	2	10	649	598	287	1.534
2	4	8	347	521	202	1.070
3	4	11	445	660	350	1.455
4	1	12	153	226	112	491
5	9	19	760	814	508	2.082
6	12	50	1.312	1.355	965	3.632
7	9	30	543	626	290	1.459
8	2	9	310	281	69	660
9	8	42	1.123	994	250	2.367
11	1	8	159	108	20	287
<i>Totale</i>	<i>52</i>	<i>199</i>	<i>5.801</i>	<i>6.183</i>	<i>3.053</i>	<i>15.037</i>

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell’informazione, 11 – Psicologia.

** I conteggi si riferiscono agli inquadramenti e alla disciplina al 31/12/2012.

Tab. 3 - Pubblicazioni e auto-citazioni dei professori del dataset

ADU*	2008-2012			2013-2017		
	Pubblicazioni	Auto-citazioni	Auto-citazioni/ pubblicazioni	Pubblicazioni	Auto-citazioni	Auto-citazioni/ pubblicazioni
1	13.798	13.902	1,01	15.513	18.511	1,19
2	32.474	103.192	3,18	38.475	148.897	3,87
3	29.450	82.383	2,80	33.098	101.377	3,06
4	4.955	9.231	1,86	6.628	15.216	2,30
5	30.552	68.437	2,24	34.853	85.655	2,46

Tab. 3 - segue

ADU*	2008-2012			2013-2017		
	Pubblicazioni	Auto-citazioni	Auto-citazioni/ pubblicazioni	Pubblicazioni	Auto-citazioni	Auto-citazioni/ pubblicazioni
6	75.795	97.092	1,28	92.640	135.941	1,47
7	15.615	19.551	1,25	20.582	34.354	1,67
8	6.095	6.549	1,07	9.995	16.059	1,61
9	40.724	34.916	0,86	55.187	59.119	1,07
11	2.985	4.767	1,60	4.043	7.410	1,83
<i>Totale</i>	<i>252.443</i>	<i>440.020</i>	<i>1,74</i>	<i>311.014</i>	<i>622.539</i>	<i>2,00</i>

* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell’informazione, 11 – Psicologia.

5.2. Modello statistico

Il semplice aumento del numero di auto-citazioni negli anni successivi all’introduzione dell’ASN non è una prova sufficiente del fatto che ci sia stato un cambiamento nel comportamento auto-citazionale illegittimo da parte degli accademici dovuto agli incentivi sottesi dall’ASN. La verifica delle ipotesi di ricerca richiede l’identificazione e il controllo di altre variabili che potrebbero influenzare i tassi di auto-citazione, come evidenziato dalla rassegna della letteratura, tra cui:

1. *Il numero di pubblicazioni del ricercatore in un determinato anno.* A parità di propensione all’auto-citazione, il numero di auto-citazioni per anno sarà maggiore con l’aumentare della produzione di pubblicazioni del ricercatore in quell’anno.
2. *Il numero totale di pubblicazioni del ricercatore.* La relazione prevista tra il numero di auto-citazioni in un determinato anno e la dimensione del portafoglio di pubblicazioni fino a quell’anno è positiva ma non lineare, con incrementi decrescenti e tendenti a zero oltre un certo numero di pubblicazioni nel portafoglio.
3. *Il campo di ricerca.* L’intensità delle citazioni varia a seconda dei campi di ricerca, pertanto le analisi devono essere condotte a livello di settore scientifico, classificando i professori nei loro SC o, meglio, SSD.

4. *La prossimità cognitiva delle pubblicazioni nel portafoglio di pubblicazioni.* Maggiore è la prossimità cognitiva (cioè la somiglianza delle basi di conoscenza che sottendono le pubblicazioni), maggiore è la probabilità che le pubblicazioni precedenti abbiano un'influenza intellettuale o cognitiva su quelle future, legittimando così le auto-citazioni. Quindi, a parità di altre condizioni, più uno scienziato si specializza (meno diversifica) più alto è il numero atteso di auto-citazioni, e viceversa.
5. *L'impatto citazionale del portafoglio di pubblicazioni dell'autore.* Tanto meno i propri lavori sono citati da altri, tanto maggiore sarà l'incen-tivo del ricercatore a gonfiare le auto-citazioni, per cercare di superare le relative soglie ASN.
6. *L'impatto delle pubblicazioni auto-citate.* Se i lavori sono *highly cited* dalla comunità scientifica, altrettanto ci si aspetta che legittimamente lo siano dai suoi autori.

Ci sono poi fattori personali e sociali che potrebbero influenzare i benefici derivanti dall'aumento del tasso di auto-citazione:

7. *Il rango accademico.* L'intensità delle citazioni varia a seconda del ran-go accademico e i benefici associati all'aumento del numero di auto-ci-tazioni sono minori per i professori ordinari rispetto ai ranghi inferiori.
 8. *Il genere.* È stato rilevato che gli uomini hanno un tasso di auto-citazione superiore alle donne, anche se in Italia ci si aspetterebbe un maggiore incentivo per le donne ad aumentare le auto-citazioni, dal momento che hanno una produttività di ricerca inferiore.
 9. *L'area geografica.* La performance di ricerca dei professori italiani va-ria a seconda dell'area geografica, e da questo ci si aspetta un incentivo diverso a gonfiare le auto-citazioni.
 10. *Il numero di coautori delle pubblicazioni citanti.* È lecito attendersi che i coautori esercitino una sorta di controllo sociale come deterrente al comportamento opportunistico; quindi, più alto è il numero di co-autori, minore è l'opportunità per il ricercatore di auto-citarsi in modo inappropriato.
 11. *Il numero di coautori internazionali delle pubblicazioni citate.* La prossimità sociale potrebbe favorire un comportamento collusivo. La distanza sociale a livello internazionale è presumibilmente maggiore rispetto a quella a livello nazionale e *intra-muros*, e quindi dovrebbe offrire minori margini per auto-citazioni opportunistiche.
- Le variabili 6-11 differiscono da 1-5 perché agiscono sull'entità dei benefici, dei rischi e dei controlli associati alle auto-citazioni opportunistiche, e solo indirettamente sul loro numero. Le abbiamo selezionate partendo dal presupposto che esista una tendenza al comportamento opportunistico da parte della comunità scientifica. Se così non fosse, non ci sarebbe bisogno

di considerarle: uno scienziato deontologicamente corretto, che non cita mai i propri lavori illegittimamente, non lo farebbe comunque, indipendentemente dai benefici, rischi e deterrenti sociali connessi. La funzione di tali variabili nel modello inferenziale utilizzato è quella di osservare eventuali cambiamenti auto-citazionali, a parità di livello di benefici o rischi associati al comportamento opportunistico.

Abbiamo quindi applicato un modello *panel* a effetti casuali binomiali negativi (Hilbe, 2011). In particolare, la variabile di risposta y_{it} è il numero di auto-citazioni del professore i nell'anno t (considerando gli articoli pubblicati negli ultimi 8 anni)¹¹, assunto come variabile casuale binomiale negativa con media μ_{it} e parametro di sovra-dispersione α , in modo che la varianza sia $(1+\alpha\mu_{it})\mu_{it}$. La binomiale negativa (o distribuzione di Pascal)¹² è un modello più flessibile per i dati di conteggio rispetto a quello di Poisson, in quanto consente una varianza maggiore della media. La variabile di esposizione n_{it} è il numero di pubblicazioni del professore i nell'anno t . Quindi $\theta_{it} = \mu_{it}/n_{it}$ è il tasso di auto-citazione, per il quale specifichiamo un modello di regressione log-lineare:

$$\log\theta_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \dots + \beta_p x_{pit} + v_i \sim N(0, \sigma^2) \quad [1]$$

Il modello prevede p covariate (x), che possono essere variabili nel tempo, più un effetto casuale v_i , che raccoglie fattori non osservati costanti nel tempo del professore i . I coefficienti di regressione sono solitamente riportati in forma esponenziale, poiché $\exp(\beta_k)$ è il rapporto dei tassi che confronta due unità statistiche con una differenza di 1 nella k -esima covariata.

I parametri del modello sono i coefficienti di regressione p più il parametro di iperdispersione α e la varianza dell'effetto casuale σ^2 . La stima viene eseguita con il comando “menbreg” di Stata, che esegue la massima verosimiglianza con quadratura Gauss-Hermite adattiva per l'effetto casuale.

11. Come detto in precedenza il dataset bibliometrico si riferisce alla produzione scientifica 2001-2017 di tutti i professori italiani in ruolo in tale periodo. Pertanto, per il 2008 (estremo inferiore del primo periodo di analisi) è possibile osservare la produzione scientifica pregressa relativa solo agli 8 anni precedenti. Tale limite, tuttavia, non risulta particolarmente critico poiché i ricercatori tendono ad auto-citare di più i lavori più recenti (Aksnes, 2003). Ovviamente, assumeremo tale finestra (di 8 anni appunto) come riferimento per tutti gli anni di ambo i periodi.

12. La distribuzione binomiale negativa è una distribuzione di probabilità discreta che descrive il numero di prove necessarie per ottenere un certo numero di successi in una serie di prove indipendenti.

Le covariate obiettivo, ossia le covariate relative alle nostre ipotesi di ricerca, sono:

- *PostASN*: variabile *dummy* che assume valore 0 fino al 2012 e valore 1 dal 2013.
- *StatusASN*: una variabile costante nel tempo con quattro categorie¹³.
- *Professori ordinari*: non partecipano all’ASN, poiché sono già inquadrati nella massima fascia.
- *Professori associati e ricercatori non abilitati nell’ASN 2012*: professori che non hanno partecipato o che hanno partecipato ma non sono stati abilitati, e che quindi dovrebbero essere interessati a partecipare ai futuri bandi ASN.
- *Ricercatori abilitati nel 2012*: avendo ottenuto l’abilitazione per il ruolo di professore associato, sarebbero motivati a partecipare ai futuri bandi di ASN, con l’intenzione di ottenere l’abilitazione per professore ordinario.
- *Professori associati abilitati nel 2012*: avendo ottenuto l’abilitazione per professore ordinario, non hanno bisogno di partecipare ulteriormente all’ASN.

Le categorie di *StatusASN* hanno incentivi diversi ad aumentare artificialmente le auto-citazioni. Nel modello, *PostASN* è interagita con *StatusASN* in modo che la variazione delle auto-citazioni dopo il 2012 sia stimata separatamente per ciascuna categoria di *StatusASN*.

Inoltre il modello include le seguenti variabili di controllo, tutte mutabili nel tempo ad eccezione del sesso:

- *Articoli citabili*: numero di articoli pubblicati negli ultimi otto anni.
- *Autori*: numero medio degli autori degli articoli pubblicati nell’anno in corso.
- *Internazionale*: variabile *dummy* che assume valore 1 se ci sono coautori internazionali degli articoli pubblicati nell’anno in corso, 0 in caso contrario.
- *Diversificazione*: numero di settori scientifici (*WoS subject category*) associati agli articoli pubblicati negli ultimi otto anni.
- *ADU*: area disciplinare universitaria del professore nell’anno in corso (10 variabili *dummy*, senza categoria di riferimento, poiché l’intercetta è omessa).

13. Per costruire la variabile *StatusASN*, ci riferiamo al rango accademico nel 2012, appena prima dell’ASN. Se il professore non ha alcun record nel 2012 a causa di zero auto-citazioni, ci riferiamo al rango accademico nell’ultimo anno osservato. Si noti che il dataset non considera se un ricercatore stia cercando di ottenere l’accreditamento per l’avanzamento a professore associato o ordinario, ma quest’ultimo caso è particolarmente raro da poter essere ignorato.

- *SSD*: settore scientifico disciplinare del professore nell’anno in corso (196 variabili *dummy*).
- *Rango*: rango accademico del professore nell’anno in corso (variabili *dummy* per ricercatore, professore associato, e professore ordinario).
- *Area geografica*: area geografica di afferenza dell’università del professore nell’anno in corso (variabili *dummy* per le macroregioni italiane del Nord, del Centro, categoria di riferimento, e del Sud).
- *Femmina*: variabile *dummy* che assume valore 1 per i professori di genere femminile, 0 in caso contrario.

Il tasso di auto-citazione di base può variare per ogni combinazione di rango e ADU, quindi ci sono 30 tassi base.

Come per altre covariate evidenziate nella rassegna della letteratura (in particolare quelle che vanno da 5 a 11 nell’elenco precedente), abbiamo preso in considerazione la possibilità di utilizzare anche l’impatto medio degli articoli auto-citati nell’anno t . Tuttavia questa variabile è determinata contemporaneamente all’esito del modello (il numero di auto-citazioni nell’anno t), quindi il suo utilizzo come covariata è discutibile. Poiché l’inserimento di questa variabile nel modello finale non comporta cambiamenti rilevanti negli altri coefficienti, la omettiamo.

L’analisi esplorativa ha suggerito l’idoneità della trasformazione logaritmica delle variabili di conteggio (*Articoli citabili*, *Autori*, *Diversificazione*). Per migliorare l’interpretabilità dei tassi di auto-citazione di base, le variabili log-trasformate sono state centrate rispetto alle medie del campione. L’anno solare è rappresentato da una variabile di conteggio discreta con valori da 0 (2008) a 9 (2017). Questa variabile è stata inserita come polinomio di quarto grado, per una maggiore flessibilità nella modellazione dell’andamento temporale.

6. Risultati

Presentiamo nell’ordine le statistiche descrittive dell’analisi empirica, quindi le statistiche inferenziali, e infine il livello di diffusione del fenomeno dell’auto-citazione.

6.1. Statistiche descrittive

Le statistiche descrittive delle variabili mutabili nel tempo sono riportate in Tab. 4, mentre quelle delle variabili costanti nel tempo in Tab. 5.

Tab. 4 - Statistiche descrittive per le variabili mutabili nel tempo (120.615 anni-professore)

Variabile	Min	Max	Media	Dev. std.
Articoli citabili	1	614	31.009	34.433
Autori	1	1.009	13.255	58.137
Diversificazione	1	70	11.557	6.771
Internazionale	0	1	0.560	—
PostASN	0	1	0.512	—
<i>Rango</i>				
Ricercatore	0	1	0.177	—
Associato	0	1	0.413	—
Ordinario	0	1	0.410	—
<i>Macroregione</i>				
Nord	0	1	0.452	—
Centro	0	1	0.279	—
Sud	0	1	0.270	—

Tab. 5 - Statistiche descrittive per le variabili costanti nel tempo (15.037 professori)

Variabile	Minimo	Massimo	Media
Donna	0	1	0.287
Status ASN			
Professore ordinario	0	1	0.377
Ricercatore o professore associato non abilitato nell'ASN 2012	0	1	0.439
Ricercatore abilitato nell'ASN 2012	0	1	0.062
Professore associato abilitato nell'ASN 2012	0	1	0.122

La Tab. 4 mostra che le variabili continue (*Articoli citabili*, *Autori* e *Diversificazione*) sono asimmetriche. Invero, entrano nel modello dopo una trasformazione logaritmica. Per le variabili binarie la media è equivalente alla proporzione, quindi per *International* risulta che nel 56% delle osservazioni c'è almeno un coautore internazionale. Si noti che la media di *PostASN* è 0,512, più grande di 0,50 poiché le osservazioni dopo ASN sono leggermente più numerose. I valori di *Rango* rivelano che i professori associati e ordinari hanno quasi lo stesso peso, mentre i ricercatori sono il

gruppo più piccolo. La variabile *Macroregione* è definita sulla base dell'affiliazione nell'anno, ad esempio il 45,2% delle osservazioni si riferisce a professori strutturati in un'università del Nord Italia.

La Tab. 5 mostra che il 28,7% dei professori è di sesso femminile e illustra la distribuzione della variabile chiave *StatusASN*, che combina il rango del professore prima dell'ASN con l'esito dell'ASN.

6.2. Inferenze dal modello statistico

La Tab. 6 riporta le stime dei parametri del modello binomiale negativo a effetti casuali (equazione [1]) in termini di rapporti di tasso, omettendo quelli meno rilevanti, ossia i tassi di base (per ogni combinazione di rango accademico e ADU), i coefficienti delle 196 *dummy* per i SSD e i 4 coefficienti polinomiali del modello temporale. La tabella omette anche il parametro di sovradispersione e la varianza dell'effetto casuale, che commentiamo di seguito.

Il parametro di sovradispersione α della binomiale negativa è stimato pari a 0,221 (s.e. 0,002), il che indica una forte sovradispersione, per cui il modello di Poisson sarebbe inappropriate. La varianza dell'effetto casuale σ^2 , stimata a 0,353 (s.e. 0,006), dovuta alle caratteristiche non osservate dei professori, è alta, come è tipico dei dati *panel*. Significa che un aumento di una deviazione standard nella distribuzione degli effetti causali implica un rapporto dei tassi di auto-citazione pari a $\exp(\sqrt{0,353})=1,812$ vale a dire che considerando due professori con le stesse covariate, il professore con effetto casuale maggiore di una deviazione standard avrebbe un tasso di autocitazione superiore dell'81,2%.

I coefficienti esponenziati $\exp(\beta)$ sono rapporti di tassi. Ad esempio, il rapporto di tassi di 1,077 per la variabile *Donna* significa che, considerando due professori con valori identici delle covariate, ad eccezione del genere, una professoressa ha un tasso di auto-citazione superiore del 7,7% rispetto a un professore. Poiché la statistica z è pari a 5,3, questo dato è statisticamente significativo. Considerando l'area geografica, il tasso di autocitazione è leggermente più basso per i professori delle università del Nord e più alto per quelli del Sud.

Tab. 6 - Stime dei parametri principali del modello panel a effetti casuali binomiali negativi per l'analisi del tasso di auto-citazione degli accademici italiani

Variabile esplicativa	ORR	Errore stand.	z	Valore p	Intervallo di confidenza 95%	
<i>StatoASN</i>						
Professore ordinario (categoria di riferimento)	1.000	–	–	–	–	–
Ricercatore o professore associato non abilitato in ASN 2012	0.884	0.019	-5.8	0.000	0.848	0.921
Ricercatore abilitato in ASN 2012	1.131	0.036	3.9	0.000	1.062	1.203
Professore associato abilitato in ASN 2012	1.101	0.026	4.1	0.000	1.051	1.154
<i>PostASN</i>						
Professore ordinario	1.099	0.014	7.4	0.000	1.072	1.127
Ricercatore o professore associato non abilitato in ASN 2012	1.150	0.016	10.2	0.000	1.120	1.182
Ricercatore abilitato in ASN 2012	1.042	0.020	2.2	0.029	1.004	1.082
Professore associato abilitato in ASN 2012	1.033	0.015	2.2	0.030	1.003	1.064
<i>ln(Articoli citabili)</i>	2.304	0.019	100.2	0.000	2.267	2.342
<i>ln(Autori)</i>	0.928	0.005	-13.3	0.000	0.918	0.939
<i>ln(Diversificazione)</i>	0.769	0.007	-27.2	0.000	0.755	0.784
<i>Internazionale</i>	0.977	0.006	-3.9	0.000	0.966	0.989
<i>Nord</i>	0.986	0.014	-1.0	0.322	0.960	1.013
Centro (categoria di riferimento)	1.000	–	–	–	–	–
<i>Sud</i>	1.022	0.016	1.4	0.168	0.991	1.053
<i>Donna</i>	1.077	0.015	5.3	0.000	1.048	1.107

Nota: l'errore standard e l'intervallo di confidenza si riferiscono al rapporto di tassi $\exp(\beta)$, mentre la statistica z e il corrispondente valore p si riferiscono al parametro originale β .

Le variabili di controllo del comportamento auto-citazionale agiscono nelle direzioni previste: il tasso di auto-citazione aumenta con il numero di articoli citabili e diminuisce con il numero di autori, la presenza di autori internazionali e il grado di diversificazione. Poiché le variabili di controllo

sono state sottoposte a una trasformazione logaritmica, i rapporti di tassi si riferiscono a un aumento di 1 nella scala logaritmica, quindi è difficile apprezzare l'entità degli effetti. Per facilitare l'interpretazione, calcoliamo i rapporti di tassi quando la covariante aumenta dal primo al terzo quartile: i rapporti di tassi sono 2,617 per gli *articoli citabili*, 0,944 per gli *autori* e 0,819 per la *diversificazione*. Come previsto, il numero di articoli citabili è di gran lunga la variabile di controllo più rilevante, ma anche il grado di diversificazione ha un effetto importante.

Le variabili target, *StatusASN* (variabile costante nel tempo con quattro categorie di status rispetto all'ASN) e *PostASN* (variabile *dummy* variabile nel tempo per gli anni successivi all'ASN), sono incluse con un'interazione. L'interpretazione delle stime è la seguente:

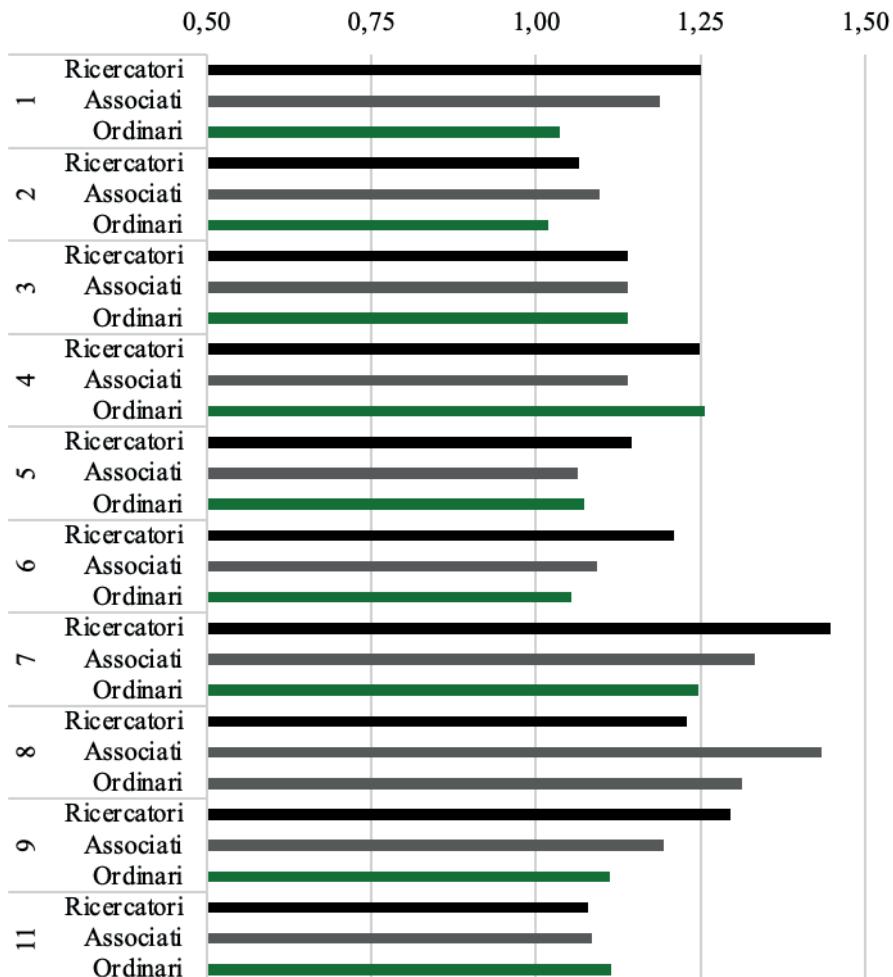
- i coefficienti di *StatusASN* riguardano i tassi di auto-citazione delle quattro categorie di status antecedenti la prima ASN, cioè nel periodo 2008-2012. Rispetto ai professori ordinari (categoria di riferimento), i ricercatori e i professori associati che non sono stati abilitati hanno un tasso di auto-citazione pari a 0,884, vale a dire che, considerando le variabili di controllo, il loro tasso di auto-citazione è stato inferiore dell'11,6%; i candidati abilitati invece presentano un tasso di auto-citazione più alto, in particolare del 13,1% i ricercatori e del 10,1% i professori associati;
- i coefficienti di *PostASN* misurano il cambiamento del comportamento auto-citazionale a seguito dell'introduzione dell'ASN per ogni categoria di *StatusASN*, attraverso il confronto tra i tassi di auto-citazione del 2013-2017 e quelli del 2008-2012. Tutte le categorie mostrano un aumento, soprattutto i ricercatori e i professori associati non abilitati nell'ASN 2012 (1,150), seguiti dai professori ordinari (1,099), dai ricercatori che hanno ottenuto l'abilitazione nel 2012 (1,042) e dai professori associati abilitati sempre nel 2012 (1,033).

Pertanto, controllando le covariate disponibili, tutti i professori hanno aumentato i tassi di auto-citazione dopo l'esperienza dell'ASN 2012. L'entità dell'aumento tra le categorie di *StatusASN* sembra essere coerentemente correlata all'entità degli incentivi di ciascuna categoria.

Per approfondire l'analisi del cambiamento nel comportamento auto-citazionale dopo l'introduzione dell'ASN, abbiamo applicato due versioni alternative del modello econometrico. Nella prima versione, *PostASN* non è interagito con *StatusASN*, ottenendo un unico coefficiente che misura il cambiamento complessivo: il rapporto dei tassi risulta pari a 1,095 (s.e. 0,013), ossia, controllando le covariate disponibili, il tasso di auto-citazione complessivo è aumentato del 9,5% dopo il 2012. Nella seconda versione, *PostASN* è ulteriormente interagito con il rango ac-

cademico e l'ADU, ottenendo l'insieme dei coefficienti specifici riportati nella Fig. 1. È significativo che, nonostante una certa eterogeneità, l'aumento del tasso di auto-citazione si sia verificato in ogni ADU e per ogni livello d'inquadramento.

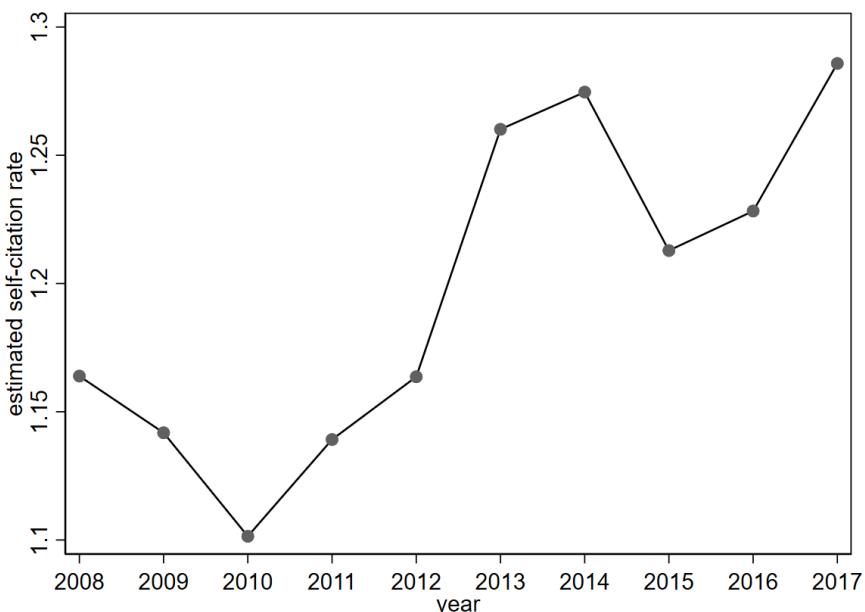
Fig. 1 - Stime dell'aumento del tasso di auto-citazione per rango accademico e area disciplinare (ADU) per i professori ordinari, e per i ricercatori e professori associati non abilitati (tutte le altre covariate sono fissate a zero)



* 1 – Matematica e informatica, 2 – Fisica, 3 – Chimica, 4 – Scienze della Terra, 5 – Biologia, 6 – Medicina, 7 – Scienze agrarie e veterinarie, 8 – Ingegneria civile, 9 – Ingegneria industriale e dell'informazione, 11 – Psicologia

Un'ultima considerazione è che i risultati del modello si basano sulla variabile dummy *PostASN*, che impone un salto nell'andamento temporale nel 2013, subito dopo l'ASN. Per verificare empiricamente se ciò si sia effettivamente verificato, abbiamo esplorato l'andamento temporale senza imporre alcuna struttura, formulando un modello in cui l'andamento temporale è considerato attraverso una serie di variabili *dummy* specifiche per anno. In effetti le stime di questo modello mostrano un salto considerevole nel 2013. Il tutto è esemplificato nella Fig. 2, che illustra l'andamento temporale del tasso di auto-citazione stimato dei professori ordinari nel SSD FIS/01 (Fisica sperimentale)¹⁴.

Fig. 2 - Stima del tasso di auto-citazione annuo (modello con dummies specifiche per anno), per i professori ordinari nel SSD FIS/01 (Fisica sperimentale) con tutte le covariate fissate a zero



14. La scelta di un SSD è necessaria per calcolare i tassi stimati, ma si noti che il modello assume la stessa struttura per tutti gli SSD.

6.3. Diffusione del cambiamento del comportamento auto-citazionale

L'analisi basata sul modello statistico ci ha permesso di stabilire che l'aumento delle auto-citazioni riguarda tutti i livelli d'inquadramento accademico e tutte le discipline. Tuttavia, i parametri del modello riflettono un comportamento medio, per cui sarebbe interessante valutare il grado di eterogeneità nel comportamento dei singoli accademici, e capire gli effetti dei casi anomali. A tal fine, abbiamo condotto un'analisi descrittiva del cambiamento del comportamento individuale, attraverso la misura della differenza tra il tasso medio di auto-citazione prima e dopo l'ASN. In questo modo, senza aggiustare le variabili di controllo, misuriamo direttamente il cambiamento nel tasso di auto-citazione. Ebbene, la differenza media è di 0,251 con deviazione standard di 1,111. Poiché il tasso medio di auto-citazioni era di 1,201 prima dell'ASN, l'aumento medio è del 20,9% ($0,251/1,201$). Questo dato è notevolmente superiore alla stima del modello, poiché non è aggiustato per le covariate. I *boxplot* nella Fig. 3 mostrano la distribuzione della differenza per ciascuna categoria di *StatusASN*.

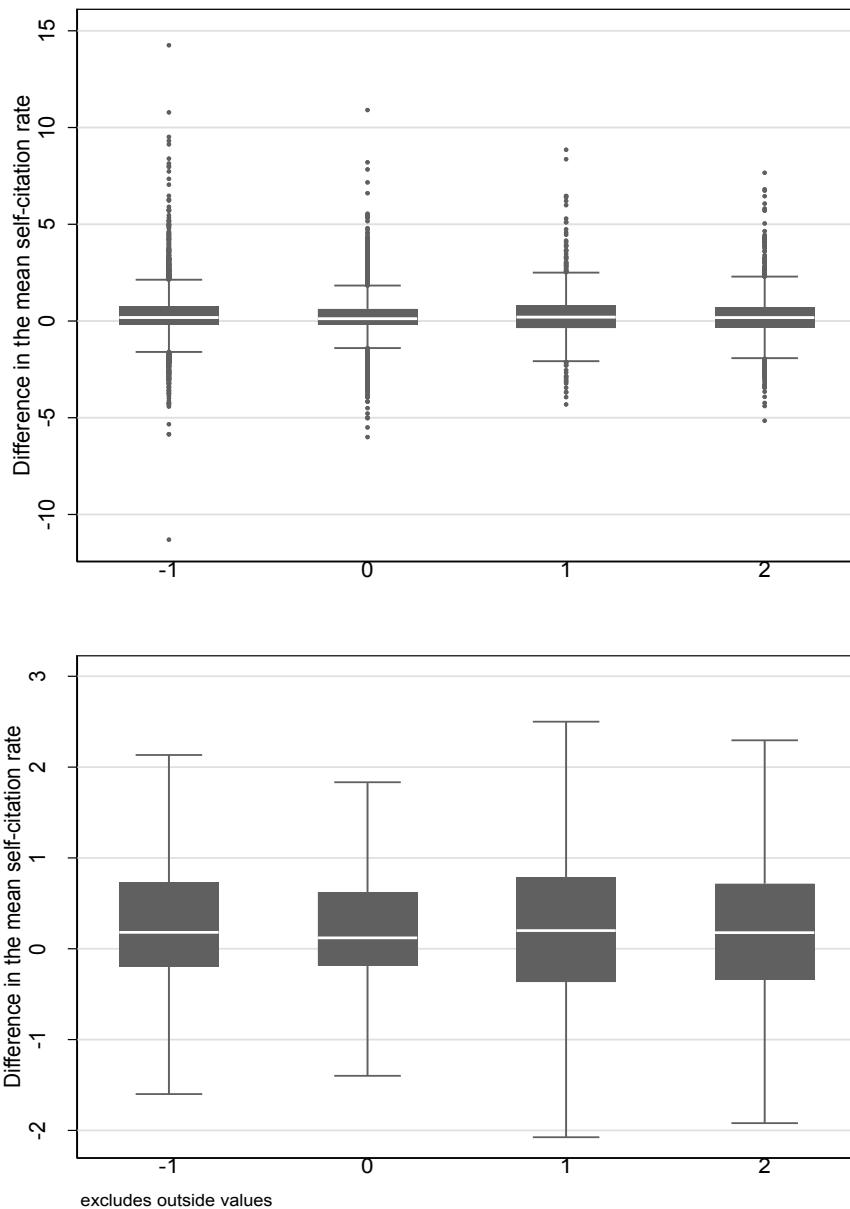
Le distribuzioni sono simili: con un'asimmetria a destra e diversi casi estremi, sia positivi che negativi. È interessante notare che i casi estremi positivi sono solo leggermente superiori a quelli negativi. Se escludiamo il punto estremo a -11,302, il minimo è -6,000, a una distanza di 6,251 dalla media 0,251. Questo porta a chiedersi quanti casi si trovino oltre una tale distanza dalla media, ovvero quante osservazioni si trovino oltre $0,251+6,251=6,502$. La risposta è 22, cioè solo 22 professori su 14259 (0,15%) mostrano un elevato aumento anomalo del loro tasso di auto-citazione.

Su 14.259 professori, il 41,1% non ha aumentato il tasso di auto-citazione e tra coloro che lo hanno fatto (il 58,9%) si osserva una grande variabilità. L'analisi dei casi estremi rivela solo alcuni casi di aumento esagerato.

7. Conclusioni

I sistemi incentivanti volti a perseguire il miglioramento continuo non sono esenti da effetti collaterali, quali l'adozione di comportamenti opportunistici. Il fenomeno delle auto-citazioni illegittime associato alla valutazione bibliometrica della performance di ricerca ne è un esempio. L'analisi del fenomeno è complessa, poiché il tasso di auto-citazione dipende da una serie di variabili che devono essere controllate quando si studiano gli effetti dei sistemi incentivanti.

Fig. 3 - Boxplot della differenza nel tasso medio di auto-citazione (2013-2017 vs 2008-2012). Il pannello superiore include i casi anomali (outlier), quello inferiore no. Da sinistra a destra, nell'ordine: professori ordinari; professori associati e ricercatori non abilitati nell'ASN 2012; ricercatori abilitati; professori associati abilitati



L'auto-citazione deriva dalla natura cumulativa del progresso scientifico, che include l'apporto personale, che l'autore ovviamente conosce e ricorda meglio di quello di altri. Può anche derivare dalla necessità di evitare l'autoplagio, o dal bisogno di gratificazione personale, o di maggiore visibilità e status scientifico o, infine, dall'intento di migliorare la propria performance di ricerca qualora misurata attraverso le citazioni.

In questo studio abbiamo dimostrato che il tasso di auto-citazione, cioè il numero medio di lavori auto-citati per pubblicazione in un arco temporale, aumenta con il numero di articoli citabili, mentre diminuisce, in ordine di importanza, con il numero di autori, la distanza cognitiva delle pubblicazioni citabili (grado di diversificazione) e la presenza di autori internazionali.

Le analisi monovariate del fenomeno spesso proposte in letteratura, sono quindi puramente descrittive e insufficienti a spiegare un rapporto di causa-effetto tra sistemi incentivanti e potenziali comportamenti opportunistici. Sarebbe inoltre scorretto soddisfare le più o meno esplicite richieste del decisore politico di poter disporre di soglie relative a un indicatore deterministico per discriminare il ricercatore onesto da quello disonesto, ridurre le distorsioni nella valutazione della performance e applicare eventuali sanzioni.

Tuttavia, da una prospettiva politica e gestionale della ricerca, rimane utile poter capire se e fino a che punto un miglioramento complessivo della produttività scientifica sia il frutto degli incentivi adottati o del comportamento opportunistico dei ricercatori, nel caso specifico, attraverso il ricorso scorretto all'auto-citazione. Nel caso di comportamento opportunistico, sarebbe anche utile sapere se occorrono differenze tra discipline, livelli d'inquadramento, genere, o aree geografiche, al fine di scontare le misure di performance coerentemente e modulare eventuali azioni correttive.

Considerando gli scopi dell'ASN o di schemi di incentivazione simili, è altrettanto importante appurare se il fenomeno delle auto-citazioni illegittime sia concentrato tra pochi o diffuso tra molti. Un comportamento concentrato sarebbe più grave, poiché qualcuno ne trarrebbe vantaggio a danno di altri, mentre un comportamento diffuso aumenterebbe semplicemente i valori delle soglie bibliometriche in generale, senza alcun reale beneficio per nessuno.

I risultati di questo lavoro mostrano che i professori italiani hanno risposto all'incentivo ASN con un aumento medio del tasso di auto-citazione del 9,5%. In pratica, ciò significa che dopo l'introduzione dell'ASN, a parità di altre covariate, un accademico che prima dell'ASN auto-citava una media di dieci lavori in ogni nuova pubblicazione, dopo ne auto-citerebbe circa undici. Abbiamo osservato che l'incremento è comune a tutte le disci-

pline e a tutti i livelli d'inquadramento, ma è sensibile all'entità dell'incen-tivo: più alto per ricercatori e professori associati non abilitati nella prima ASN (+15%), meno per i professori ordinari (+9,9%), ricercatori abilitati (+4,2%) e professori associati abilitati (+3,3%). Da ciò si evince che una volta conseguita l'abilitazione, gli accademici tendano contenere il ricorso all'auto-citazione.

È interessante osservare che il fenomeno dell'aumento delle auto-citazioni in seguito all'introduzione dell'ASN non è stato diffuso. Oltre il 41% dei professori non ha aumentato il proprio tasso di auto-citazione, il che potrebbe sollevare dubbi sull'equità del processo di abilitazione o sugli esercizi di valutazione della performance di ricerca basati sulle citazioni. In uno studio futuro si potrebbero confrontare le performance di coloro che hanno aumentato il tasso di auto-citazione e di coloro che non lo hanno fatto.

Rispetto all'unico studio multivariato sul tema (Seeber, *et al.*, 2019), il nostro campo di osservazione è molto più ampio, in termini di numero di osservazioni (15.037 contro 866), di SSD (199 SSD contro 4) e di finestra temporale successiva all'ASN (5 anni contro 2). In termini metodologici il modello inferenziale è simile a quello di Seeber *et al.* (2019), con l'aggiunta di un'ulteriore categoria di status, i professori abilitati nel 2012. Le principali variabili di controllo sono le stesse, mentre poche altre sono diverse, il che non dovrebbe cambiare sostanzialmente i risultati, che infatti sono allineati.

I dati disponibili non ci permettono di stabilire in che misura gli au-menti osservati siano il frutto di un puro ricorso all'auto-citazione illegittima, o a una maggiore attenzione a citare legittimamente i propri lavori in aggiunta o al posto di quelli di altri. Questo aspetto potrebbe essere oggetto di studi di caso specifici.

Certamente, non tutti gli incrementi possono essere condannati a pri-ori come frutto di comportamento non etico. Del resto, in un precedente studio sul fenomeno della “donazione” della firma di una pubblicazione (*authorship* illegittime) a seguito del primo esercizio di valutazione della ricerca in Italia (VQR 2004-2010), non sono emerse evidenze di compor-tamento opportunistico da parte degli accademici italiani (Abramo, D'An-gelo e Di Costa, 2019b). Dobbiamo inoltre ricordare che prima di essere pubblicati, tutti i lavori sono sottoposti al vaglio di revisori ed editori, che dovrebbero notare eventuali eccessi nel ricorso alle auto-citazione ed esse-re in grado di discriminare le auto-citazioni legittime da quelle illegittime.

È consapevolezza diffusa che tutti i sistemi incentivanti sono suscetti-bili di indurre un certo livello di comportamento opportunistico, e quan-do sono emerse possibili evidenze di ciò ha talvolta portato a sostenere o

suggerire l'abbandono delle metriche di valutazione e dei relativi sistemi incentivanti (Baccini *et al.*, 2019). Nel caso specifico della valutazione della ricerca, prima di abbandonare tali sistemi, sarebbe almeno opportuno valutare se, al netto degli effetti di comportamenti opportunistici, gli strumenti incentivanti messi in essere abbiano effettivamente raggiunto gli obiettivi di una migliore selezione del personale di ricerca e di un aumento della produttività. Molti studi si limitano a esaminare solo un lato della medaglia, ossia il comportamento opportunistico. D'altra parte, in un precedente lavoro sull'efficacia del reclutamento nelle università italiane dopo l'introduzione dell'ASN, è stato dimostrato un netto calo dell'ingresso o progressione nel sistema accademico di professori con scarsa produttività di ricerca, che è il principale obiettivo dell'ASN e un impatto positivo sulla produttività di ricerca dei professori italiani a seguito dell'introduzione della VQR (Abramo e D'Angelo, 2021).

Il sospetto di comportamento opportunistico ha portato anche a suggerire di escludere le auto-citazioni dalle metriche di valutazione (Baccini *et al.*, 2019). Il dibattito sulla questione è in corso e il consenso è ancora lontano. Alcuni studiosi ritengono che le auto-citazioni non siano importanti quanto le citazioni di altri accademici nel determinare l'autorevolezza di un accademico all'interno della propria comunità epistemica (Fowler e Aksnes, 2007). Alcuni hanno iniziato a filtrare le auto-citazioni, mentre altri suggeriscono di correggere gli indicatori di performance escludendo le auto-citazioni (Brown, 2009). Altri viceversa sostengono che non considerare le auto-citazioni sia indifendibile e confonde inutilmente qualsiasi tentativo di comprendere l'impatto bibliometrico del lavoro di ricerca” (Kacem *et al.*, 2020).

A nostro avviso, escludere le auto-citazioni dal conteggio delle citazioni, quando si intraprende un'analisi comparativa dei risultati della ricerca, è una decisione politica delicata e prima di prenderla si dovrebbero considerare anche i suoi effetti collaterali. L'esclusione delle auto-citazioni, infatti, penalizzerebbe i ricercatori onesti, in particolare quelli che pubblicano molto e sono altamente specializzati nel proprio settore. Dall'analisi condotta in questo lavoro risulta che, per ogni auto-citazione aggiuntiva post-ASN frutto di possibile comportamento opportunistico, che andremmo ad escludere, ne escluderemmo altre dieci, considerate legittime fino all'introduzione dell'ASN. La nostra posizione personale è che non si dovrebbe cercare di correggere un comportamento opportunistico attraverso misure che penalizzerebbero coloro che non lo hanno mai adottato. Sarebbe invece più sensato che revisori e editori fossero più attenti nell'identificare e inibire eventuali auto-citazioni illegittime. Come affermano Szomszor *et al.* (2020): “Se c'è davvero un numero significativo di individui che adotta

comportamenti non etici, perché il sistema editoriale non lo rileva e non lo sanziona al momento della revisione, prima della pubblicazione? Oppure anche la gestione delle riviste è un problema?”.

Per quanto riguarda le considerazioni più generali sui difetti o sull'eccessivo affidamento alle metriche citazionali per le decisioni concernenti il reclutamento, l'avanzamento di carriera, e il finanziamento delle istituzioni di ricerca, la nostra opinione personale è che, almeno in Italia, le potenzialità di tali metriche non siano sfruttate a pieno e che, laddove si ricorra alla bibliometria, siano piuttosto le modalità applicativa e la scelta degli indicatori di performance a dover essere corretti. Le evidenze empiriche mostrano che perfino il semplice utilizzo delle citazioni (piuttosto che più sofisticati indicatori basati sia sulle citazioni che su altre *feature*) è più efficace della *peer review* per la valutazione dell'impatto degli articoli scientifici (Abramo, D'Angelo e Reale, 2019). Il punto è piuttosto come migliorare e diffondere l'uso appropriato della bibliometria, anziché insistere sul ritorno alla *peer review* quale unica tecnica di valutazione dell'impatto.

In una fase storica in cui il sistema di *peer review* sta letteralmente esplodendo e si cercano soluzioni anche nell'ambito dell'intelligenza artificiale (Thelwall *et al.*, 2023), è davvero insostenibile continuare ad affidarsi a questo approccio per esercizi su larga scala. Alcuni studiosi propongono sistemi di ricompensa appropriati per incentivare i pari a svolgere il lavoro, ma può davvero essere una soluzione, una soluzione sostenibile? I revisori sono già sovraccarichi di lavoro. Aczel *et al.* (2021) hanno calcolato che “il tempo totale in cui i revisori a livello globale hanno lavorato alle revisioni tra pari è stato di oltre 100 milioni di ore nel 2020, pari a oltre 15 mila anni”.

Eppure, stiamo assistendo a un dibattito schizofrenico, con quasi tutte le voci più influenti che sostengono il ritorno ad approcci esclusivamente qualitativi, in cui la revisione paritaria svolge un ruolo centrale. Fa eccezione l'appassionata lettera di Torres-Salinas *et al.* (2023) che parlano di “negazionismo bibliometrico”, riferendosi ai molti studiosi che “negano categoricamente l'utilità degli indicatori bibliometrici, scoraggiandone fortemente l'uso nei processi di valutazione scientifica... Nell'era dei Big Data e delle decisioni informate dai dati, sembra che nel nostro settore si stia promuovendo l'opposto, proprio quando abbiamo più metodi e indicatori bibliometrici a disposizione”. Sulla stessa lunghezza d'onda è il presidente dell'International Society for Scientometrics and Informetrics (ISSI) Giovanni Abramo: attraverso evidenze empiriche e argomentazioni logiche interne, egli contesta la supremazia dei metodi qualitativi su quelli quantitativi in termini di robustezza, accuratezza, validità, funzionalità e convenienza, ma anche in termini di tempi e costi, almeno nella valutazione nelle aree STEMM (Abramo, 2024).

La decisione di utilizzare metodi di *peer review* o scientometrici dovrebbe essere basata su una valutazione completa degli obiettivi specifici, delle risorse disponibili e dei contesti culturali rilevanti. Entrambi i metodi hanno i loro limiti e suggeriscono una combinazione pragmatica dei due, sfruttando i punti di forza di ciascuno. La soluzione non è quindi quella di scartare le metriche, ma di usarle in modo saggio e appropriato. Se applicate in modo responsabile, le metriche offrono oggettività, coerenza e scalabilità. Nei campi in cui le banche dati bibliografiche non forniscono una copertura adeguata, la revisione tra pari rimane essenziale. Tuttavia, nelle discipline STEMM, dove la maggior parte dell'avanzamento della conoscenza è diffusa attraverso pubblicazioni indicizzate, la scientometria può fornire indicazioni preziose sull'impatto della ricerca e consentire valutazioni più precise delle prestazioni della ricerca.

Riguardo poi al miglioramento e all'applicazione delle metriche bibliometriche, chiosiamo questo studio con un'osservazione di Heinze e Jappe (2020), sul ricorso a modelli bibliometrici in Olanda e in Italia: “L’Olanda segue un modello di professionalità bibliometrica, mentre l’Italia segue un modello burocratico centralizzato che coopta le élite accademiche”, e nonostante “il Paese disponga di organizzazioni di esperti con competenze bibliometriche, la progettazione e l’attuazione degli esercizi di valutazione nazionali è rimasta al di fuori delle loro competenze”. Per l’Italia, la sfida consiste nell’attingere alle conoscenze disponibili per sviluppare metriche sempre più accurate, e applicarle con cognizione di causa, senza perdere di vista i loro principali limiti.

Conclusioni

di Emanuela Reale

Il progetto PRIN del quale abbiamo presentato alcuni risultati aveva come scopo quello di approfondire gli effetti sul lavoro degli accademici della valutazione della ricerca promossa dal governo in Italia attraverso gli strumenti della VQR e dell'ASN, avendo come punti di attenzione principali gli effetti di cambiamento su a) qualità della produzione accademica, b) produttività dei ricercatori, c) effetti di stimolo alle collaborazioni scientifiche e alla mobilità internazionale, d) reclutamento dei ricercatori e relazioni con il mentore nei giovani ricercatori.

In queste sintetiche conclusioni cerchiamo di fare il punto sui principali risultati raggiunti dal nostro studio, quanto essi si accordino con evidenze già rilevate in letteratura e cosa aggiungano di nuovo, quali siano le prospettive di ricerca che si aprono sul tema degli effetti della valutazione, nonché le indicazioni che possono essere di interesse per i decisori politici che hanno il compito di disegnare la valutazione in Italia.

Qualità percepita

Una parte rilevante dei risultati generati dal nostro studio si riferisce agli effetti della valutazione sulla qualità della ricerca, nel senso di un suo miglioramento e della capacità di influenzare domande di ricerca, problemi affrontati, struttura metodologica scelta, modi per disseminare i risultati.

Nel capitolo 1 la *survey* nazionale restituisce un quadro secondo il quale è molto ampia la numerosità degli accademici che ritengono che la necessità di produrre pubblicazioni in linea con le linee guida della VQR abbia influenzato la qualità del proprio lavoro scientifico, interessando una percentuale di soggetti intervistati largamente superiore alla metà del campione coinvolto nella rilevazione. Tuttavia, non è trascurabile anche la quota di soggetti che invece non ha ritenuto che la valutazione abbia avuto effetti sulla qualità della ricerca prodotta, indicando quindi che l'e-

sercizio della VQR ha prodotto cambiamenti solo su alcuni soggetti e non su altri.

L'area disciplinare (in particolare le scienze economiche e statistiche), la posizione geografica (appartenere a un'università del Sud), il ruolo accademico (essere un giovane ricercatore non ancora stabilizzato), e la dimensione dell'università di appartenenza (lavorare in una piccola università), sono risultati i fattori più importanti nella percezione degli effetti sulla qualità del lavoro scientifico. Invece, le variabili demografiche come il genere e la presenza di figli non sembrano influenzare significativamente la percezione della VQR. Dunque l'incentivo prodotto dalla valutazione della ricerca promossa dal governo in Italia ha selettivamente generato effetti sulle comunità che hanno modalità di disseminazione dei risultati propri e diversi da quelli delle aree scientifiche sperimentali, sulle comunità giovani, precarie e collocate in atenei più svantaggiati, nei confronti dei quali è stato esercitato un intervento esterno per riconfigurare i parametri di qualità scientifica riportandoli all'interno di uno schema valoriale estrinseco rispetto a quello delle comunità di riferimento.

Le storie di vita (capitolo 2) approfondiscono questo risultato e consentono di esplorare le esperienze personali dei giovani ricercatori (ricercatori non ancora stabilizzati all'interno del percorso accademico) e comprendere gli effetti sulle loro scelte di carriera. Qui il differenziale prodotto dal settore disciplinare nella percezione degli effetti della valutazione ha un'evidenza molto forte che si lega al conflitto tra i criteri e indicatori dell'ASN e a come è valutata la qualità del lavoro scientifico all'interno dei settori scientifici di appartenenza. Tale differenza si lega anche all'influenza esercitata dal mentore per l'avvio alla carriera scientifica, che risulta estremamente rilevante, e alla pressione avvertita dai giovani per assicurarsi la propria indipendenza scientifica, di riuscire ad attrarre fondi di ricerca competitivi, i quali sono distribuiti sulla base di processi di valutazione nei quali lo *standing* del proponente prevale sulla qualità della proposta. Gli effetti della valutazione sui giovani sono così incisivi da influire anche sulla struttura dell'agenda di ricerca, la selezione dei problemi da affrontare e dei metodi da applicare nel lavoro scientifico (capitolo 3). Il *focus* sulle scienze sociali nell'analisi degli effetti della valutazione sull'agenda di ricerca ha consentito di entrare nei particolari di queste discipline e nell'effetto trasformativo prodotto dalla valutazione nazionale sia attraverso VQR, sia attraverso ASN.

Letteratura e comportanti osservati nelle università e nelle riviste scientifiche

L'analisi della letteratura nei capitoli 5 e 6 ha consentito di delineare un quadro delle criticità e delle sfide che riguardano gli attuali sistemi di valutazione e le implicazioni prodotte dalla prevalenza degli indicatori

quantitativi nella misurazione delle varie dimensioni della qualità della ricerca. In particolare, sono stati evidenziati a) i limiti associati all'uso degli indicatori citazionali come misura della qualità della ricerca, b) la necessità di applicare metodi diversificati, sia quantitativi che qualitativi, in base al settore disciplinare considerato, nonché c) l'opportunità di considerare il contesto nel quale la ricerca viene svolta.

Attraverso osservazioni legate a risultati di altri paesi europei la letteratura ha confermato anche la distanza epistemica tra la qualità della ricerca e i sistemi di valutazione basati sulla performance. Tali sistemi influenzano le prassi di ricerca promuovendo determinati tipi di ricerca o di metodologie e la scelta delle domande di ricerca per aumentare la probabilità di pubblicare su riviste ad alto impatto. L'introduzione di sistemi di valutazione che pongono l'accento sull'internazionalizzazione e sull'impatto accademico ha portato a conseguenze indesiderate, tra cui la creazione di una cultura *publish or perish*, favorendo un'omogeneizzazione dei risultati della ricerca dovuta alla tendenza dei ricercatori a produrre lavori conformi ai criteri dell'esercizio di valutazione nazionale.

L'approfondimento sul settore Organizzazione aziendale dell'Area 13 conferma che, anche nel caso italiano, l'attuale sistema di valutazione ha generato una perdita di diversità epistemica. I ricercatori sono indotti a pubblicare più frequentemente e solo su riviste indicizzate, internazionali o inserite nei *ranking* nazionali come la Fascia A dell'ANVUR, scoraggiando linee di ricerca non convenzionali. I risultati suggeriscono che, nel periodo considerato, le riforme introdotte in Italia hanno modificato i comportamenti non solo degli accademici ma anche delle istituzioni, almeno nell'ambito del settore concorsuale considerato. Infatti, alcune università, per competere nel settore dell'alta formazione a livello globale, hanno reclutato accademici di fama internazionale titolari di decine di pubblicazioni su riviste di Classe A che sono state citate migliaia di volte. Lo strumento ha quindi indirizzato le istituzioni a interpretare le regole introdotte in modo opportunistico; in questo modo le regole fatte per aumentare la qualità del sistema nazionale diventano una leva importante per vincere la sfida della produttività e migliorare l'impatto scientifico di alcune università, quelle appunto che possono disporre di risorse sufficienti.

Tuttavia, prescindendo dai casi specifici di singole università, nell'ambito disciplinare considerato in questo studio, a una produttività maggiore non corrisponde un analogo incremento delle citazioni. Emerge dunque un disallineamento delle dimensioni della qualità che viene confermato anche dalle analisi bibliometriche: a un contenuto ritenuto scientificamente rilevante (perché pubblicato in riviste di Classe A) non sempre corrisponde un

interesse nella comunità accademica e, conseguentemente, in alcuni domini del dibattito scientifico.

Inoltre, uno specifico studio che ha coinvolto un campione di *editor* di riviste scientifiche in settori bibliometrici e non bibliometrici (capitolo 6) ha cercato di entrare nei processi di accettazione dei manoscritti sottomessi per identificare i fattori che influenzano le decisioni degli *editor* nelle fasi iniziali del processo di pubblicazione e che, in ultima analisi, legittimano la qualità di un manoscritto. I risultati mostrano che la valutazione della ricerca introduce nel sistema logiche estranee alla scienza in senso stretto, quali quelle di mercato, anche nell'editoria scientifica. La standardizzazione del processo di *peer review* ha orientato il focus del giudizio dei revisori verso caratteristiche estrinseche della qualità, al di là del contenuto della ricerca, trasformando il ruolo dell'*editor* delle riviste scientifiche ‘da custode del futuro della disciplina a dominus della carriera degli accademici’, collegato alla possibilità di pubblicare su riviste a elevato impatto. Così, la logica di mercato declinata attraverso il richiamo agli standard di performance, si è introdotta nel contesto accademico modificando le prospettive e le priorità dei ricercatori.

Produttività, qualità e reclutamento del personale accademico

Le analisi bibliometriche sviluppate nei capitoli 7, 8 e 9 controllano elementi relativi alla produttività, effetti distorsivi nella disseminazione dei risultati attraverso le pratiche citazionali, e cambiamenti nel sistema di reclutamento del personale accademico.

Per quanto riguarda la produttività della ricerca si segnala un aumento relativo delle università meridionali rispetto a quelle del centro-nord, utilizzando indicatori e metodi più adatti al confronto tra la situazione pre e post valutazione, pur con i limiti evidenziati in relazione alla copertura delle pubblicazioni in alcuni settori disciplinari. L'effetto di bilanciamento della performance tra le due circoscrizioni geografiche del nostro paese sarebbe dunque un effetto positivo prodotto dalla VQR. Tuttavia, l'aumento della mobilità degli studenti del sud verso le università del nord e le analisi che sono state condotte sulle cause di questa mobilità lungo la penisola (Viesti, 2019) evidenziano fattori diversi dalla produttività scientifica delle università. Emergono sia fattori di contesto sia il persistere e l'aumentare di un divario reputazionale, anche a causa del sistema di valutazione, che insieme producono effetti cumulativi negli anni deprimente il finanziamento di base (FFO). Il risultato descritto dalle analisi bibliometriche non contrasta con quanto rilevato nelle *survey*, in particolare quella condotta a livello nazionale, dove l'aumento della produttività è segnalato da una larghissima percentuale di rispondenti, con risultati particolarmente accentuati nei rispondenti appartenenti agli atenei meridionali.

Inoltre, l'aumento della produttività non risulta associato né a un miglioramento della qualità della produzione scientifica, né a un miglioramento del reclutamento delle università. Se infatti la percentuale di neoassunti improduttivi nelle attività di ricerca è diminuita, questo effetto positivo è compensato dal fatto che il rendimento medio complessivo non è aumentato, come dimostrano i valori medi e mediani dei *rating* di produttività delle coorti osservate. In sostanza, se da una parte diminuisce la percentuale di neoassunti con performance di ricerca mediocri, dall'altra diminuisce anche quella di coloro che hanno performance superiori alla media. Questo risultato potrebbe essere legato all'altissima percentuale di successo che i candidati interni delle università ricevono nei concorsi, fenomeno ben conosciuto dal mondo accademico italiano e i cui effetti negativi sono stati analizzati in letteratura. Questo fenomeno è certamente collegato a una programmazione interna delle università non sempre strategica sotto il profilo del recupero di competitività e di qualità del personale accademico. Tuttavia, l'effetto potrebbe essere stato ulteriormente accentuato dalla mancata crescita del finanziamento di base collegato all'FFO che ha accompagnato gli anni di introduzione della valutazione in Italia, e che ha segnato il limite più visto- so per una politica il cui obiettivo era aumentare la performance del sistema universitario attraverso la leva del finanziamento.

L'analisi relativa ai comportamenti citazionali degli individui a seguito dell'intervento dell'ASN disegna invece un quadro in parte allineato e in parte diverso da quello che emerge da altri studi svolti sul medesimo tema in Italia. Il controllo è stato diretto a individuare la presenza di comportamenti opportunistiche degli studiosi per alterare i risultati degli indici citazionali attraverso un incremento delle autocitazioni come risposta all'incentivo introdotto dall'ASN verso chi raggiunge alti livelli citazionali. Utilizzando solo le citazioni basate sulla presenza dell'autore citante come autore anche del lavoro citato nel periodo precedente l'introduzione dell'ASN e in quello successivo, i risultati indicano che l'aumento delle autocitazioni è modesto e interessa i professori di tutti i livelli di inquadramento e in tutte le discipline scientifiche. Il fenomeno, tuttavia, non è diffuso in quanto una larga quota di accademici non modifica il proprio comportamento dopo l'avvento dell'ASN, e pochi sono i comportamenti eccessivi chiaramente identificabili come opportunistiche.

Collaborazioni e mobilità internazionale

Il lavoro ha esplorato in particolare l'eventuale ruolo della valutazione nazionale tra i fattori che motivano o promuovono l'intrapresa di collaborazioni scientifiche e la partecipazione a iniziative di mobilità verso l'estero.

La valutazione nazionale della ricerca assume un ruolo relativamente marginale quale *driver* nell'avvio di attività di ricerca congiunta: in so-

stanza, essa non agisce come fattore di stimolo per la creazione di network nazionali o internazionali. Similmente, la *survey* nazionale indica con chiarezza che essa viene raramente annoverata dagli accademici tra le motivazioni per la decisione di intraprendere mobilità verso l'estero.

In sintesi, è emerso come la valutazione nazionale abbia un impatto limitato sul processo decisionale relativo all'intrapresa di collaborazioni scientifiche e alla partecipazione a iniziative di mobilità, non producendo effetti significativi di incentivo. Coloro che denunciano un effetto della valutazione nazionale per questo tipo di scelta sono comunque caratterizzati dall'appartenenza a piccole università localizzate nel sud d'Italia.

Prospettive future

I risultati ottenuti hanno dunque confermato e precisato molte delle osservazioni già presenti in letteratura, anche internazionale, sugli effetti generati dalla ‘valutazione sommativa’ promossa dal governo per l’allocazione basata sulle performance FFO. Vale comunque la pena evidenziare alcune novità, tra le molte presenti nei capitoli che compongono questo volume.

Primo: l’attribuzione alla valutazione di molteplici effetti negativi senza una precisa individuazione di nessi causali tra effetti (autocitazioni, iperproduttività, iper-autorialità, non riproducibilità, falsi, uso incontrollato dell’intelligenza artificiale, riviste predatorie, cfr. ROARS, 5 giugno 2024) e cause, e senza individuare chi sono gli individui (o le organizzazioni) effettivamente interessati, non appare corretta dal punto di vista dei risultati del nostro lavoro. Il focus sugli individui ha evidenziato, infatti, che la valutazione non ha effetti diffusi su tutta la comunità accademica: al contrario, gli effetti sono concentrati e incidono con maggiore importanza su alcune comunità piuttosto che su altre. Questo risultato rappresenta, a nostro modo di vedere, un aspetto assai critico e discutibile per uno strumento come la valutazione massiva, in linea di principio destinato a tutto il mondo dell’istruzione superiore e volto all’aumento della qualità di tutto il sistema universitario.

Secondo: in linea con quanto detto al punto precedente, gli effetti della valutazione si concentrano maggiormente su alcune comunità epistemiche, piuttosto che su altre, e gli effetti tendono a differenziarsi maggiormente in relazione ai diversi settori scientifici sui quali la valutazione stessa insiste, ponendosi quindi come fattore estrinseco che guida il cambiamento del modo di creazione e diffusione della conoscenza scientifica, almeno per la parte rilevabile attraverso le pubblicazioni.

Terzo: sul piano metodologico, l’approccio contributivo aiuta a selezionare gli effetti sui quali la valutazione ha effettivamente avuto un ruolo importante rispetto ad altri fattori potenzialmente influenti. Il risultato, sotto

questo profilo, è l'importanza forte e più diffusa rispetto alla valutazione che ha il finanziamento su base competitiva per l'acquisizione di progetti di ricerca. In questo senso, resta valido l'invito a riflettere se sia davvero necessario e attuale, nel contesto dei cambiamenti trasformativi che la ricerca sta vivendo anche in relazione all'affermarsi del paradigma della scienza aperta, continuare a utilizzare i risultati della valutazione per usi distributivi piuttosto che valorizzare gli aspetti formativi che essa può avere. In relazione a questo punto, conviene accennare, seppur brevemente, a due tra i tanti elementi in favore di una prospettiva formativa. Da una parte, occorre rispondere alla crescente necessità di valorizzare la valutazione non solo dei risultati ma anche dei processi di ricerca da cui i risultati stessi derivano, non solo per finalità di trasparenza rispetto al lavoro che ha portato al risultato, ma anche per apprezzare la qualità stessa del risultato. Dall'altra, l'importanza attribuita alla produzione d'impatto sociale della ricerca riceverà risposte fallibili e parziali se si continua ad associarlo a una prospettiva di performance. La generazione d'impatto sulla società è fattore che deve essere conciliato con la ricerca della migliore qualità scientifica, obiettivo questo molto complesso, il cui perseguitamento necessita di osservazione e valutazione in chiave di lezioni che possono essere apprese per aumentare la capacità delle università di essere luoghi di ricerca responsabile e inclusiva dei bisogni della società.

Quarto: l'uso di metodi misti, che integrino osservazioni provenienti da approcci diversi e da diverse fonti di dati, potrebbe aiutare a controllare la robustezza dei risultati e rispondere a domande di ricerca diverse.

Molti suggerimenti per future linee di ricerca sul tema degli effetti prodotti dalla valutazione sul lavoro individuale sono contenuti nei diversi capitoli del volume. Una proposta di carattere più generale, che emerge dall'esperienza di questo lavoro, è migliorare la capacità di individuare il nesso causale tra eventi osservati e imputazione dei medesimi alla valutazione rispetto ad altre concause possibili nel corso del tempo. A questo riguardo, il limite più severo recentemente proposto all'attenzione degli studiosi di scienze sociali è quello della complessità e sostanziale difficoltà di collegare in modo univoco gli effetti di un singolo strumento di *policy*, come ad esempio la valutazione VQR e ASN, ai risultati prodotti (Gläser, 2024). La proposta avanzata dall'autore è quella di cambiare la prospettiva di analisi e operare con un percorso di tipo abduttivo: partire dall'osservazione del fenomeno prodotto cercando di ricostruire retrospettivamente le diverse cause del cambiamento ed eventuali spiegazioni alternative. In questo modo, il cambiamento sarebbe circoscritto e la difficoltà risiederebbe nella selezione di strategie di comparazione tra effetti osservabili nelle diverse comunità epistemiche.

Sul piano delle indicazioni di *policy*, il punto principale non è la proposta di eventuali correttivi al sistema di valutazione sulla base dei risultati esposti, ma l'invito a considerare in prospettiva critica se e quanto sia opportuno e vantaggioso per la crescita della qualità e della competitività delle università italiane, continuare a mantenere in vita un sistema di finanziamento della ricerca fortemente basato sulla performance guidata dai risultati di una valutazione esterna massiva. Gli effetti positivi collegati alla prima attuazione dello strumento si sono rapidamente esauriti. Oggi lo strumento non sembra garantire il miglioramento della qualità atteso, specie se consideriamo la riduzione delle risorse pubbliche destinate all'università che mettono in crisi anche la possibilità di svolgere le normali funzioni didattiche.

La logica sottesa alla misurazione della performance attraverso indicatori, *rating* e *ranking* appare dunque sempre più problematica per valutare la ricerca. Alla ricerca vengono richiesti crescenti livelli di innovazione dei risultati, forte direzionalità verso la produzione d'impatto, coinvolgimento di soggetti esterni al mondo accademico (i cd. stakeholder), responsabilità nell'uso degli indicatori e inclusività nei confronti di scienziati appartenenti a culture diverse e con diverse tradizioni scientifiche. Se si considera la necessità di garantire la completa trasparenza e accessibilità alla ricerca lungo tutto il percorso scientifico, al risultato finale prodotto, seguendo i principi già espressi per la modifica dei sistemi di valutazione nazionale nell'ambito del CoARA Agreement (Coalition for Advancing Research Assessment), risulta difficile negare come l'attuale sistema non possa più essere corretto per stratificazioni successive. Esso dovrebbe essere ripensato e ridisegnato abbracciando le logiche della 'scienza aperta', che con il *performance assessment* ha pochi (e problematici) punti in comune.

Riferimenti bibliografici

- Aagaard K. (2015), “How incentives trickle down: Local use of a national bibliometric indicator system”, *Science and Public Policy*, 42, 5: 725-737. <https://doi.org/10.1093/scipol/scu087>
- Aagaard K., Bloch C. & Schneider J. W. (2015), “Impacts of performance-based research funding systems: The case of the Norwegian Publication Indicator”, *Research Evaluation*, 24, 2: 106-117. <https://doi.org/10.1093/reseval/rvv003>
- Aagaard K. & Schneider J. W. (2017), “Some considerations about causes and effects in studies of performance-based research funding systems”, *Journal of Informetrics*, 11, 3: 923-926. <https://doi.org/10.1016/j.joi.2017.05.018>
- Abramo G. (2018), “Revisiting the scientometric conceptualization of impact and its measurement”, *Journal of Informetrics*, 12, 3: 590-597. <https://doi.org/10.1016/j.joi.2018.05.001>
- Abramo G. (2024), “The forced battle between peer review and scientometric research assessment: Why the CoARA initiative is unsound”, *Research Evaluation*. rvae021. <https://doi.org/10.1093/reseval/rvae021>
- Abramo G., Aksnes D. W. & D'Angelo C. A. (2020), “Comparison of research performance of Italian and Norwegian professors and universities”, *Journal of Informetrics*, 14, 2. <https://doi.org/10.1016/j.joi.2020.101023>
- Abramo G., Aksnes D. W. & D'Angelo C. A. (2021), “Gender differences in research performance within and between countries: Italy vs Norway”, *Journal of Informetrics*, 15, 2. <https://doi.org/10.1016/j.joi.2021.101144>
- Abramo G., Cicero T. & D'Angelo C. A. (2011a), “Assessing the varying level of impact measurement accuracy as a function of the citation window length”, *Journal of Informetrics*, 5, 4: 659-667. <https://doi.org/10.1016/j.joi.2011.06.004>
- Abramo G., Cicero T. & D'Angelo C. A. (2011b), “The dangers of performance-based research funding in non-competitive higher education systems”, *Scientometrics*, 87, 3: 641-654. <https://doi.org/10.1007/s11192-011-0355-4>

- Abramo G., Cicero T. & D'Angelo C. A. (2012a), "A sensitivity analysis of researchers' productivity rankings to the time of citation observation", *Journal of Informetrics*, 6, 2: 192-201. <https://doi.org/10.1016/j.joi.2011.12.003>
- Abramo G., Cicero T. & D'Angelo C. A. (2012b), "Revisiting size effects in higher education research productivity", *Higher Education*, 63, 6: 701-717. <https://doi.org/10.1007/s10734-011-9471-6>
- Abramo G., Cicero T. & D'Angelo C. A. (2012c), "The dispersion of research performance within and between universities as a potential indicator of the competitive intensity in higher education systems", *Journal of Informetrics*, 6, 2: 155-168. <https://doi.org/10.1016/j.joi.2011.11.007>
- Abramo G., Cicero T. & D'Angelo C. A. (2013a), "Individual research performance: A proposal for comparing apples to oranges", *Journal of Informetrics*, 7, 2: 528-539. <https://doi.org/10.1016/j.joi.2013.01.013>
- Abramo G., Cicero T. & D'Angelo C. A. (2013b), "National peer review research assessment exercises for the hard sciences can be a complete waste of money: The Italian case", *Scientometrics*, 95, 1: 311-324. <https://doi.org/10.1007/s11192-012-0875-6>
- Abramo G., Cicero T. & D'Angelo C. A. (2014), "Are the authors of highly cited articles also the most productive ones?", *Journal of Informetrics*, 8, 1: 89-97. <https://doi.org/10.1016/j.joi.2013.10.011>
- Abramo G. & D'Angelo C. A. (2011), "Evaluating research: From informed peer review to bibliometrics", *Scientometrics*, 87, 3: 499-514. <https://doi.org/10.1007/s11192-011-0352-7>
- Abramo G. & D'Angelo C. A. (2014a), "Assessing national strengths and weaknesses in research fields", *Journal of Informetrics*, 8, 3: 766-775. <https://doi.org/10.1016/j.joi.2014.07.002>
- Abramo G. & D'Angelo C. A. (2014b), "How do you define and measure research productivity?", *Scientometrics*, 101, 2: 1129-1144. <https://doi.org/10.1007/s11192-014-1269-8>
- Abramo G. & D'Angelo C. A. (2015a), "An assessment of the first 'scientific habilitation' for university appointments in Italy", *Economia Politica*, 32, 3: 329-357. <https://doi.org/10.1007/s40888-015-0016-9>
- Abramo G. & D'Angelo C. A. (2015b), "Ranking research institutions by the number of highly-cited articles per scientist", *Journal of Informetrics*, 9, 4: 915-923. <https://doi.org/10.1016/j.joi.2015.09.001>
- Abramo G. & D'Angelo C. A. (2015c), "The VQR, Italy's second national research assessment: Methodological failures and ranking distortions", *Journal of the Association for Information Science and Technology*, 66, 11: 2202-2214. <https://doi.org/10.1002/asi.23323>
- Abramo G. & D'Angelo C. A. (2016a), "Refrain from adopting the combination of citation and journal metrics to grade publications, as used in the Italian national research assessment exercise (VQR 2011-2014)", *Scientometrics*, 109, 3: 2053-2065. <https://doi.org/10.1007/s11192-016-2153-5>
- Abramo G. & D'Angelo C. A. (2017), "On tit for tat: Franceschini and Maisano versus ANVUR regarding the Italian research assessment exercise VQR 2011-2014", *Journal of Informetrics*, 11, 3: 783-787. <https://doi.org/10.1016/j.joi.2017.06.003>

- Abramo G. & D'Angelo C. A. (2021), "Were the Italian policy reforms to contrast favoritism and foster effectiveness in faculty recruitment successful?", *Science and Public Policy*, 47, 5: 604-615. <https://doi.org/10.1093/scipol/scaa048>
- Abramo G., D'Angelo C. A. & Cicero T. (2012), "What is the appropriate length of the publication period over which to assess research performance?", *Scientometrics*, 93, 3: 1005-1017. <https://doi.org/10.1007/s11192-012-0714-9>
- Abramo G., D'Angelo C. A. & Di Costa F. (2009a), "Mapping excellence in national research systems: The case of Italy", *Evaluation Review*, 33, 2. <https://doi.org/10.1177/0193841X08322871>
- Abramo G., D'Angelo C. A. & Di Costa F. (2009b), "Research collaboration and productivity: is there correlation?", *Higher Education*, 57, 2: 155-171. <https://doi.org/10.1007/s10734-008-9139-z>
- Abramo G., D'Angelo C. A. & Di Costa F. (2011), "Research productivity: Are higher academic ranks more productive than lower ones?", *Scientometrics*, 88, 3: 915-928. <https://doi.org/10.1007/s11192-011-0426-6>
- Abramo G., D'Angelo C. A. & Di Costa F. (2014a), "Inefficiency in selecting products for submission to national research assessment exercises", *Scientometrics*, 98, 3: 2069-2086. <https://doi.org/10.1007/s11192-013-1177-3>
- Abramo G., D'Angelo C. A. & Di Costa F. (2014b), "Investigating returns to scope of research fields in universities", *Higher Education*, 68, 1: 69-85. <https://doi.org/10.1007/s10734-013-9685-x>
- Abramo G., D'Angelo C. A. & Di Costa F. (2017), "Specialization versus diversification in research activities: the extent, intensity and relatedness of field diversification by individual scientists", *Scientometrics*, 112, 3: 1403-1418. <https://doi.org/10.1007/s11192-017-2426-7>
- Abramo G., D'Angelo C. A. & Di Costa F. (2019), "A nation's foreign and domestic professors: which have better research performance? (the Italian case)", *Higher Education*, 77, 5: 917-930. <https://doi.org/10.1007/s10734-018-0310-x>
- Abramo G., D'Angelo C. A. & Di Costa F. (2019a), "Diversification versus specialization in scientific research: Which strategy pays off?", *Technovation*, 82-83: 51-57. <https://doi.org/10.1016/j.technovation.2018.06.010>
- Abramo G., D'Angelo C. A. & di Costa F. (2019b), "When research assessment exercises leave room for opportunistic behavior by the subjects under evaluation", *Journal of Informetrics*, 13, 3: 830-840. <https://doi.org/10.1016/j.joi.2019.07.006>
- Abramo G., D'Angelo C. A. & Di Costa F. (2021), "On the relation between the degree of internationalization of cited and citing publications: A field level analysis, including and excluding self-citations", *Journal of Informetrics*, 15, 1: 101101. <https://doi.org/10.1016/j.joi.2020.101101>
- Abramo G., D'Angelo C. A. & Grilli L. (2021), "The effects of citation-based research evaluation schemes on self-citation behavior", *Journal of Informetrics*, 15, 4: 101204. <https://doi.org/10.1016/j.joi.2021.101204>
- Abramo G., D'Angelo C. A. & Murgia G. (2017), "The relationship among research productivity, research collaboration, and their determinants", *Journal of Informetrics*, 11, 4: 1016-1030. <https://doi.org/10.1016/j.joi.2017.09.007>

- Abramo G., D'Angelo C. A. & Reale E. (2019), "Peer review versus bibliometrics: Which method better predicts the scholarly impact of publications?", *Scientometrics*, 121, 1: 537-554. <https://doi.org/10.1007/s11192-019-03184-y>
- Abramo G., D'Angelo C. A. & Rosati F. (2014a), "Relatives in the same university faculty: nepotism or merit?", *Scientometrics*, 101, 1: 737-749. <https://doi.org/10.1007/s11192-014-1273-z>
- Abramo G., D'Angelo C. A. & Rosati F. (2014b), "Career advancement and scientific performance in universities", *Scientometrics*, 98, 2: 891-907. [https://doi.org/10.1007/S11192-013-1075-8/METRICS](https://doi.org/10.1007/S11192-013-1075-8)
- Abramo G., D'Angelo C. A. & Rosati F. (2015), "The determinants of academic career advancement: Evidence from Italy", *Science and Public Policy*, 42, 6: 761-774. <https://doi.org/10.1093/scipol/scu086>
- Abramo G., D'Angelo C. A. & Rosati F. (2016a), "A methodology to measure the effectiveness of academic recruitment and turnover", *Journal of Informetrics*, 10, 1: 31-42. <https://doi.org/10.1016/j.joi.2015.10.004>
- Abramo G., D'Angelo C. A. & Rosati F. (2016b), "The north-south divide in the Italian higher education system", *Scientometrics*, 109, 3: 2093-2117. <https://doi.org/10.1007/s11192-016-2141-9>
- Abramo G., D'Angelo C. A., & Soldatenkova A. (2017). "An investigation on the skewness patterns and fractal nature of research productivity distributions at field and discipline level", *Journal of Informetrics*, 11, 1, 324-335. <https://doi.org/10.1016/j.joi.2017.02.001>
- Abritis A. & McCook A. (2017), "Cash incentives for papers go global", *Science*, 357, 6351: 541. <https://doi.org/10.1126/science.357.6351.541>
- Ackers L. (2008), "Internationalisation, Mobility and Metrics: A New Form of Indirect Discrimination?", *Minerva*, 46, 4: 411-435. <https://doi.org/10.1007/s11024-008-9110-2>
- Aczel B., Szaszi B. & Holcombe A. O. (2021), "A billion-dollar donation: estimating the cost of researchers' time spent on peer review", *Research Integrity and Peer Review*, 6, 1. <https://doi.org/10.1186/s41073-021-00118-2>
- Adams J. (2013), "Collaborations: The fourth age of research", *Nature*, 497, 7451: 557-560. <https://doi.org/10.1038/497557a>
- Adams J. & Gurney K. (2010), *Funding selectivity, concentration and excellence - how good is the UK's research?*, HEPI Publications - Higher Education Policy Institute. Report disponibile all'indirizzo www.hepi.ac.uk/2010/03/25/funding-selectivity-concentration-and-excellence-how-good-is-the-uks-research/
- Adenot D., Latour B. & Woolgar S. (1983), "Labory Life. The Social Construction of Scientific Facts", *Revue Française de Sociologie*, 24, 1: 156. <https://doi.org/10.2307/3321794>
- Agre P. (2005), *Networking on the Network: A Guide to Professional Skills for PhD Students*. Testo disponibile al: https://blog.regehr.org/extra_files/Networking_on_the_Network.pdf (Consultato il 23 Aprile 2023)
- Ajiferuke I., Burrell Q. & Tague J. (1988), "Collaborative coefficient: A single measure of the degree of collaboration in research", *Scientometrics*, 14, 5-6: 421-433. <https://doi.org/10.1007/BF02017100>

- Åkerlind G. S. (2008), "An academic perspective on research and being a researcher: an integration of the literature", *Studies in Higher Education*, 33, 1: 17-31. <https://doi.org/10.1080/03075070701794775>
- Aksnes D. W. (2003), "A macro study of self-citation", *Scientometrics*, 56, 2: 235-246. <https://doi.org/10.1023/A:1021919228368>
- Aksnes D. W. (2005), *Citations and their use as indicators in science policy. Studies of validity and applicability issues with a particular focus on highly cited papers.* Tesi di dottorato. https://ris.utwente.nl/ws/portalfiles/portal/6087461/thesis_Aksnes.pdf
- Aksnes D. W., Langfeldt, L. & Wouters, P. (2019), "Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories", *SAGE Open*, 9, 1: 2158244019829575. <https://doi.org/10.1177/2158244019829575>
- Albarrán P., Ortúño I. & Ruiz-Castillo J. (2011), "High- and low-impact citation measures: Empirical applications", *Journal of Informetrics*, <https://doi.org/10.1016/j.joi.2010.10.001>
- Alerasoul S. A., Afeltra G., Hakala H., Minelli E. & Strozzi F. (2022), "Organisational learning, learning organisation, and learning orientation: An integrative review and framework", *Human Resource Management Review*, 32, 3: 100854. <https://doi.org/10.1016/J.HRMR.2021.100854>
- Allen N. (1988), "Aspects of promotion procedures in Australian universities", *Higher Education*, 17, 3: 267-280. <https://doi.org/10.1007/BF00163790>
- Allesina S. (2011), "Measuring nepotism through shared last names: The case of Italian academia", *PLoS ONE*, 6, 8. <https://doi.org/10.1371/journal.pone.0021160>
- Altbach P. G. (1996), "The Academic Profession in International Perspective". In P. G. Altbach (Ed.), *The International Academic Profession: Portraits of Fourteen Countries*, 3-48. Carnegie Foundation for the Advancement of Teaching.
- Altman Y., Baruch Y., Zoghbi Manrique-de-Lara P. & Viera Armas M. M. (2020), "Baby boomers at the cusp of their academic career: Storming ahead, Hanging on, or Calling it a day", *Studies in Higher Education*, 45, 7: 1335-1350. <https://doi.org/10.1080/03075079.2019.1610864>
- Ancaiani A., Anfossi A. F., Barbara A., Benedetto S., Blasi B., Carletti V., Cicero T., Ciolfi A., Costa F., Colizza G., Costantini M., di Cristina F., Ferrara A., Lacatena R. M., Malgarini M., Mazzotta I., Nappi C. A., Romagnosi S. & Sileoni S. (2015), "Evaluating scientific research in Italy: The 2004-10 research evaluation exercise", *Research Evaluation*, 24, 3: 242-255. <https://doi.org/10.1093/reseval/rvv008>
- Andrej M., Vladimir B. & Wouter de N. (2018), *Exploratory Social Network Analysis with Pajek*. In *Complex Adaptive Systems Modeling* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781108565691>
- ANVUR (2014), *Rapporto sullo stato del sistema universitario e della ricerca 2013. Rapporto Sullo Stato Del Sistema Universitario e Della Ricerca 2013*. Disponibile all'indirizzo www.anvur.it/archivio-documenti-ufficiali/rapporto-sullo-stato-del-sistema-universitario-e-della-ricerca-2013/
- Arocena R., Göransson B. & Sutz J. (2019), "Towards making research evaluation more compatible with developmental goals", *Science and Public Policy*, 46, 2: 210-218. <https://doi.org/10.1093/scipol/scy051>

- Atkinson M. (2001), "Peer Review Culture", *Science and Engineering Ethics*, 7, 2: 193-204. [https://doi.org/10.1007/S11948-001-0040-8/METRICS](https://doi.org/10.1007/S11948-001-0040-8)
- Aung H. H., Zheng H., Erdt M., Aw A. S., Sin S. C. J. & Theng Y. L. (2019), "Investigating familiarity and usage of traditional metrics and altmetrics", *Journal of the Association for Information Science and Technology*, 70, 8: 872-887. <https://doi.org/10.1002/asi.24162>
- Auranen O. & Nieminen M. (2010), "University research funding and publication performance-An international comparison", *Research Policy*, 39, 6: 822-834. <https://doi.org/10.1016/j.respol.2010.03.003>
- Aydogan I. (2012), "The existence of favoritism in organizations", *African Journal of Business Management*, 6, 12: 4577-4586.
- Babbie E. R. (2020), *The practice of social research*. Cengage Learning.
- Baccini A. (2014), "La VQR di Area 13: una riflessione di sintesi", *Statistica & Società*, 3, 3: 32-37.
- Baccini A. & De Nicolao G. (2016), "Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise", *Scientometrics*, 108, 3: 1651-1671. <https://doi.org/10.1007/s11192-016-1929-y>
- Baccini A., De Nicolao G. & Petrovich E. (2019), "Citation gaming induced by bibliometric evaluation: A country-level comparative analysis", *PLoS ONE*, 14, 9. <https://doi.org/10.1371/journal.pone.0221212>
- Baffy G., Burns M. M., Hoffmann B., Ramani S., Sabharwal S., Borus J. F., Pories S., Quan S. F. & Ingelfinger J. R. (2020), "Scientific Authors in a Changing World of Scholarly Communication: What Does the Future Hold?", *American Journal of Medicine*, 133, 1: 26-31. <https://doi.org/10.1016/j.amjmed.2019.07.028>
- Baldarelli M. G., del Baldo M. & Vignini S. (2016), "Pink accounting in Italy: Cultural perspectives over discrimination and/or lack of interest", *Meditari Accountancy Research*, 24, 2: 269-292. <https://doi.org/10.1108/MEDAR-10-2015-0065/FULL/XML>
- Baldwin M. (2017a), "In referees we trust?", *Physics Today*, 70, 2: 44. <https://doi.org/10.1063/PT.3.3463>
- Baldwin M. (2017b), "What it was like to be peer reviewed in the 1860s", *Physics Today*. <https://doi.org/10.1063/PT.5.9098>
- Balstad M. T. & Berg T. (2020), "A long-term bibliometric analysis of journals influencing management accounting and control research", *Journal of Management Control*, 30, 4: 357-380. <https://doi.org/10.1007/S00187-019-00287-8/TABLES/6>
- Bangani S. & Onyancha O. B. (2021), "Evaluation of the national research foundation-rated researchers' output at a South African university", *Global Knowledge, Memory and Communication*, 70, 1-2: 187-202. <https://doi.org/10.1108/GKMC-02-2020-0017>
- Barjak F. & Robinson S. (2008), "International collaboration, mobility and team diversity in the life sciences: impact on research performance", *Social Geography*, 3, 1: 23-36. <https://doi.org/10.5194/sg-3-23-2008>
- Beaver D. Deb. (2001), "Reflections on Scientific Collaboration (and its study): Past, Present, and Future", *Scientometrics*, 52, 3: 365-377. <https://doi.org/10.1023/A:1014254214337>

- Beaver D. Deb & Rosen R. (1978), "Studies in scientific collaboration", *Scientometrics*, 1, 1: 65-84. <https://doi.org/10.1007/BF02016840>
- Bergé L. R. (2017), "Network proximity in the geography of research collaboration", *Papers in Regional Science*, 96, 4: 785-815. <https://doi.org/10.1111/pirs.12218>
- Bertaux D., ed. (1981), *Biography and Society: The Life History Approach in the Social Sciences*, Sage Publications, Beverly Hills, Calif.
- Biagioli M. (2016), "Watch out for cheats in citation game", *Nature*, 535, 7611: 201. <https://doi.org/10.1038/535201a>
- Bichi R. (2000), "La tecnica di intervista nelle storie di vita: il rimando neutro, forme e modalità di riconoscimento tra aree di significato", *Studi Di Sociologia*, 38, 2: 175-188.
- Bichi R. (2002), *L'intervista biografica. Una proposta metodologica*, Vita e Pensiero.
- Binswanger M. (2014), *Excellence by nonsense: The competition for publications in modern science*, in Bartling S., Friesike S. eds., *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, Springer Open, 49-72.
- Bleiklie I., Enders J., Lepori B. & Musselin C. (2013), "NPM, network governance and the university as a changing professional organization", In *The Ashgate Research Companion to New Public Management*. Available at: <https://sciencespo.hal.science/hal-00972968/document>
- Bloch C. & Schneider J. W. (2016), "Performance-based funding models and researcher behavior: An analysis of the influence of the Norwegian Publication Indicator at the individual level", *Research Evaluation*, 25, 4: 371-382. <https://doi.org/10.1093/reseval/rvv047>
- Boch Waldorff S., Reay T. & Goodrick E. (2013), A Tale of Two Countries: How Different Constellations of Logics Impact Action, in Lounsbury M. and Boxenbaum, E., ed. *Institutional Logics in Action, Part A (Research in the Sociology of Organizations, Vol. 39 Part A)*, Emerald Group Publishing Limited, Leeds, pp. 99-129. [https://doi.org/10.1108/S0733-558X\(2013\)0039AB008](https://doi.org/10.1108/S0733-558X(2013)0039AB008)
- Bonaccorsi A. (2015), *La valutazione possibile. Teoria e pratica nel mondo della ricerca*, Il Mulino, Bologna.
- Bonaccorsi A. & Daraio C. (2005), "Exploring size and agglomeration effects on public research productivity", *Scientometrics*, 63, 1: 87-120. <https://doi.org/10.1007/s11192-005-0205-3>
- Bongaerts J. C. (2022), "The Humboldtian Model of Higher Education and its Significance for the European University on Responsible Consumption and Production", *BHM Berg- Und Hüttmännische Monatshefte*, 167, 10: 500-507. <https://doi.org/10.1007/S00501-022-01280-W>
- Borgman C. L. (1989), "Bibliometrics and Scholarly Communication: Editor's Introduction", *Communication Research*. <https://doi.org/10.1177/009365089016005002>
- Bornmann L. & Daniel H. D. (2007), "What do we know about the h index?", *Journal of the American Society for Information Science and Technology*, 58, 9: 1381-1385. <https://doi.org/10.1002/asi.20609>

- Bornmann L. & Daniel H.-D. (2008), "What do citation counts measure? A review of studies on citing behavior", *Journal of Documentation*, 64, 1: 45-80. <https://doi.org/10.1108/00220410810844150>
- Bornmann L., Haunschild R. & Hug S. E. (2018), "Visualizing the context of citations referencing papers published by Eugene Garfield: a new type of keyword co-occurrence analysis", *Scientometrics*, 114, 2: 427-437. <https://doi.org/10.1007/s11192-017-2591-8>
- Borrelli D. & Stazio M. (2018), "La "grande trasformazione" dell'università italiana", *Rivista Trimestrale Di Scienza dell'amministrazione*, 1: 1-19.
- Boschma R. (2005), "Proximity and Innovation: A Critical Assessment", *Regional Studies*, 39, 1: 61-74. <https://doi.org/10.1080/0034340052000320887>
- Bourdieu P. (1988), *Homo academicus*. Stanford University Press, Stanford, Calif.
- Bozeman B. & Sarewitz D. (2011), "Public Value Mapping and Science Policy Evaluation", *Minerva*, 49, 1: 1-23. <https://doi.org/10.1007/s11024-011-9161-7>
- Bozzon R., Murgia A., Poggio B. & Rapetti E. (2017), "Work-life interferences in the early stages of academic careers: The case of precarious researchers in Italy", *European Educational Research Journal*, 16, 2-3: 332-351. <https://doi.org/10.1177/1474904116669364>
- Brantnell A. & Baraldi E. (2020), "Following unique logics despite institutional complexity: An inductive study of academic inventors and institutional logics", *European Management Journal*, 38, 5: 684-697. <https://doi.org/10.1016/j.emj.2020.02.005>
- Bratti M., Cicero T., Lippo E., Nappi C. A. & Turri M. (2021), "The effect of research evaluation exercises on research output: fifteen years of evidence from Italy", *Politica Economica*, 37, 1: 13-42.
- Brew A. (2003), "Teaching and research: New relationships and their implications for inquiry-based teaching and learning in higher education", *Higher Education Research and Development*, 22, 1: 3-18. <https://doi.org/10.1080/0729436032000056571>
- Bridges D. (2011), "Research Quality Assessment: Intended and Unintended Consequences", *Power and Education*, 3, 1: 31-38. <https://doi.org/10.2304/POWER.2011.3.1.31>
- Bromley P. & Powell W. W. (2012), "From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World", *Academy of Management Annals*, 6, 1: 483-530. <https://doi.org/10.1080/19416520.2012.684462>
- Brown M., Gómez González A. & Munté Pascual A. (2013), "Procesos dialógicos de planificación de los servicios sociales: el proceso de cambio en los barrios de La Milagrosa y La Estrella (Albacete)", *Scripta Nova: Revista Electrónica de Geografía y Ciencias Sociales*, 17, 277. www.um.edu.mt/library/oar/handle/123456789/118516
- Brown R. J. C. (2009), "A simple method for excluding self-citation from the h-index: The b-index", *Online Information Review*, 33, 6: 1129-1136. <https://doi.org/10.1108/14684520911011043>
- Brysbaert M. & Smyth S. (2011), "Self-enhancement in scientific research: The self-citation bias", *Psychologica Belgica*, 51, 2: 129-137. <https://doi.org/10.5334/pb-51-2-129>

- Buckle R. A. & Creedy J. (2019), "The evolution of research quality in New Zealand universities as measured by the performance-based research fund process", *New Zealand Economic Papers*, 53, 2: 144-165. <https://doi.org/10.1080/00779954.2018.1429486>
- Buckle R. A., Creedy J. & Gemmell N. (2020), "Is external research assessment associated with convergence or divergence of research quality across universities and disciplines? Evidence from the PBRF process in New Zealand", *Applied Economics*, 52, 36: 3919-3932. <https://doi.org/10.1080/00036846.2020.1725235>
- Butler L. (2003a), "Explaining Australia's increased share of ISI publications - the effects of a funding formula based on publication counts", *Research Policy*, 32, 1: 143-155. [https://doi.org/10.1016/S0048-7333\(02\)00007-0](https://doi.org/10.1016/S0048-7333(02)00007-0)
- Butler L. (2003b), "Modifying publication practices in response to funding formulas", *Research Evaluation*, 12, 1 SPEC ISS.: 39-46. <https://doi.org/10.3152/147154403781776780>
- Cafagna L. (1989), "La questione delle origini del dualismo economico italiano", *Dualismo E Sviluppo Nella Storia D'Italia*, 187-220.
- Cagan R. (2013), "The San Francisco Declaration on Research Assessment", *DMM Disease Models and Mechanisms*, 6, 4: 869-870. <https://doi.org/10.1242/dmm.012955>
- Cai Y. & Mountford N. (2022), "Institutional logics analysis in higher education research", *Studies in Higher Education*, 47, 8: 1627-1651. <https://doi.org/10.1080/03075079.2021.1946032>
- Campbell D. F. J. (2013), *New University Governance: How the Academic Profession Perceives the Evaluation of Research and Teaching*, in Teichler U. & Höhle E. A., eds., *The Work Situation of the Academic Profession in Europe: Findings of a Survey in Twelve Countries*, Springer Netherlands, 205-228. https://doi.org/10.1007/978-94-007-5977-0_10
- Canhilal S. K., Lepori B. & Seeber M. (2016), *Decision-making power and institutional logic in higher education institutions: A comparative analysis of European universities*, in Pinheiro R., Geschwind L., Ramirez F.O. and Vrangbæk, K., eds., *Towards A Comparative Institutionalism: Forms, Dynamics And Logics Across The Organizational Fields Of Health Care And Higher Education (Research in the Sociology of Organizations, Vol. 45)*, Emerald Group Publishing Limited, Leeds, pp. 169-194. <https://doi.org/10.1108/S0733-558X20150000045019>
- Cañibano C., D'Este P., Otamendi F. J. & Woolley R. (2020), "Scientific careers and the mobility of European researchers: an analysis of international mobility by career stage", *Higher Education*, 80, 6: 1175-1193. <https://doi.org/10.1007/s10734-020-00536-z>
- Cañibano C., Vilardell I., Corona C. & Benito-Amat C. (2018), "The evaluation of research excellence and the dynamics of knowledge production in the humanities: The case of history in Spain", *Science and Public Policy*, 45, 6: 775-789. <https://doi.org/10.1093/scipol/scy025>
- Capano G. (2010), "A Sisyphean Task: Evaluation and Institutional Accountability in Italian Higher Education", *Higher Education Policy*, 23, 1: 39-62. <https://doi.org/10.1057/hep.2009.19>

- Casnici N., Grimaldo F., Gilbert N. & Squazzoni F. (2017), "Attitudes of referees in a multidisciplinary journal: An empirical analysis", *Journal of the Association for Information Science and Technology*, 68, 7: 1763-1771. <https://doi.org/10.1002/asi.23665>
- Castells M. (2000), *The rise of the network society* (2nd ed.). Blackwell Publishers. <https://doi.org/10.1002/9781444319514>
- Cattaneo M., Meoli M. & Signori A. (2016), "Performance-based funding and university research productivity: the moderating effect of university legitimacy", *Journal of Technology Transfer*, 41, 1: 85-104. <https://doi.org/10.1007/s10961-014-9379-2>
- Cerejo C. (2014), "International journal editors and East Asian authors: Two surveys", *Learned Publishing*, 27, 1: 63-75. <https://doi.org/10.1087/20140110>
- Chandler J., Barry J. & Clark H. (2002), "Stressing academe: The wear and tear of the New Public Management", *Human Relations*, 55, 9: 1051-1069. <https://doi.org/10.1177/0018726702055009019>
- Charlton B. G. (2004), "Conflicts of interest in medical science: Peer usage, peer review and 'CoI consultancy'", *Medical Hypotheses*, 63, 2: 181-186. Churchill Livingstone. <https://doi.org/10.1016/j.mehy.2004.06.001>
- Chavarro D., Ràfols I. & Tang P. (2018), "To what extent is inclusion in the Web of Science an indicator of journal 'quality?", *Research Evaluation*, 27, 3: 284-284. <https://doi.org/10.1093/reseval/rvy015>
- Chavarro D., Tang P. & Ràfols I. (2017), "Why researchers publish in non-mainstream journals: Training, knowledge bridging, and gap filling", *Research Policy*, 46, 9: 1666-1680. <https://doi.org/10.1016/j.respol.2017.08.002>
- Checchi D., Malgarini M. & Sarlo S. (2019), "Do performance-based research funding systems affect research production and impact?", *Higher Education Quarterly*, 73, 1: 45-69. <https://doi.org/10.1111/hequ.12185>
- Checchi D., Mazzotta I., Momigliano S. & Olivanti F. (2020), "Convergence or polarisation? The impact of research assessment exercises in the Italian case", *Scientometrics*, 124, 2: 1439-1455. <https://doi.org/10.1007/s11192-020-03517-2>
- Chubin D. E. & Hackett E. J. (1990), "Peerless Science: Peer Review and U.S. Science Policy", *Social Forces*, 69, 2: 634. <https://doi.org/10.2307/2579681>
- Colarusso S. & Giancola O. (2020), *Università e nuove forme di valutazione: Strategie individuali, produzione scientifica, effetti istituzionali*, Collana Materiali e documenti, Vol. 59. Sapienza Università Editrice.
- Colicchia C. & Strozzi F. (2012), "Supply chain risk management: A new methodology for a systematic literature review", *Supply Chain Management*. <https://doi.org/10.1108/13598541211246558>
- Collins J. C. & Porras J. I. (1991), "Organizational Vision and Visionary Organizations", *California Management Review*, 34, 1: 30-52. <https://doi.org/10.2307/41166682>
- Conrath-Hargreaves A. & Wüstemann S. (2019), "Managing Multiple Institutional Logics and the Use of Accounting: Insights from a German Higher Education Institution", *Abacus*, 55, 3: 483-510. <https://doi.org/10.1111/abac.12164>
- COPE (2020), *Editing of reviewer comments - Survey results*. <https://doi.org/10.24318/gYZ99iXA>

- Cora-Bramble D. (2006), “Minority faculty recruitment, retention and advancement: Applications of a resilience-based theoretical framework”, *Journal of Health Care for the Poor and Underserved*, 17, 2: 251-255. <https://doi.org/10.1353/hpu.2006.0057>
- Corbetta P. (2014), *Metodologia e Tecniche Della Ricerca Sociale*, Il Mulino, Bologna.
- Corsi M., D’Ippoliti C. & Zacchia G. (2018), “A Case Study of Pluralism in Economics: The Heterodox Glass Ceiling in Italy”, *Review of Political Economy*, 30, 2: 172-189. <https://doi.org/10.1080/09538259.2018.1423974>
- Corsi M., D’Ippoliti C. & Zacchia G. (2019), “On the evolution of the glass ceiling in Italian academia: The case of economics”, *Science in Context*. <https://doi.org/10.1017/S0269889720000046>
- Costas R., van Leeuwen T. N. & Bordons M. (2010), “Self-citations at the meso and individual levels: Effects of different calculation methods”, *Scientometrics*, 82, 3: 517-537. <https://doi.org/10.1007/s11192-010-0187-7>
- Costas R., Zahedi Z. & Wouters P. (2014), “Disentangling the meaning of ‘altmetrics’: content analysis of Web of Science scientific publications”. Disponibile al sito: <http://altmetrics.org/altmetrics14/> (consultato il 12 marzo 2023).
- Costas R., Zahedi Z. & Wouters P. (2015), “Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective”, *Journal of the Association for Information Science and Technology*, 66, 10: 2003-2019. <https://doi.org/10.1002/asi.23309>
- Coulthard D. & Keller S. (2016), “Publication Anxiety, Quality, and Journal Rankings: Researcher Views”, *Australasian Journal of Information Systems*, 20. <https://doi.org/10.3127/ajis.v20i0.1262>
- Creswell J. & Creswell D. (2018), *Research design: qualitative, quantitative, and mixed methods approaches*. SAGE Publications, Inc. (US). SAGE.
- Creswell J. W., Plano Clark V. L., Gutmann M. L. & Hanson W. E. (2003), *Advanced Mixed Methods Research Designs.*, in A. Tashakkori A., Teddlie C. eds, *Handbook of Mixed Methods in Social and Behavioral Research*, (pp. 209-240), Sage, Thousand Oaks.
- Cronin B. (1984), “The Citation Process: The Role and Significance of Citations in Scientific Communication”, *College & Research Libraries*, 46, 1. https://doi.org/10.5860/crl_46_01_86
- Cruz-Castro L. & Sanz-Menéndez L. (2010), “Mobility versus job stability: Assessing tenure and productivity outcomes”, *Research Policy*, 39, 1: 27-38. <https://doi.org/10.1016/j.respol.2009.11.008>
- Csiszar A. (2016), “Peer review: Troubled from the start”, *Nature*, 532, 7599: 306-308. <https://doi.org/10.1038/532306a>
- Cunliffe A. L. (2022), “Must I Grow a Pair of Balls to Theorize about Theory in Organization and Management Studies”, *Organization Theory*, 3, 3. <https://doi.org/10.1177/26317877221109277>
- Dahler-Larsen P. (2011), *The Evaluation Society*. Stanford University Press, Stanford, CA.

- D'Angelo C. A., Giuffrida C. & Abramo G. (2011), "A heuristic approach to author name disambiguation in bibliometrics databases for large-scale research assessments", *Journal of the American Society for Information Science and Technology*, 62, 2: 257-269. <https://doi.org/10.1002/asi.21460>
- Daniele V. (2015), "Two Italies? Genes, intelligence and the Italian North-South economic divide", *Intelligence*, 49: 44-56. <https://doi.org/10.1016/j.intell.2014.12.004>
- D'Antuono P. & Ciavarella M. (2019), "The false myth of the rise in self-citations, and the impressively positive effect of bibliometric evaluations on the increase of the impact of Italian research", Arxiv preprint. <https://arxiv.org/abs/1910.02948>.
- de Boer H., Jongbloed B., Benneworth P., Cremonini L., Kolster R., Kottmann A., Lemmens-Krug K. & Vossensteyn H. (2015), *Performance-based funding and performance agreements in fourteen higher education systems*, Report for the Ministry of Education, Culture and Science, Center for Higher Education Policy Studies, Universiteit Twente. Available at: <https://ris.utwente.nl/ws/portalfiles/portal/5139542/jongbloed%20ea%20performance-based-funding-and-performance-agreements-in-fourteen-higher-education-systems.pdf>
- De Filippo D., Casado E. S. & Gómez I. (2009), "Quantitative and qualitative approaches to the study of mobility and scientific performance: a case study of a Spanish university", *Research Evaluation*, 18, 3: 191-200. <https://doi.org/10.3152/095820209X451032>
- de Jong S. P. L., van Arensbergen P., Daemen F., van der Meulen B. & van den Besselaar P. (2011), "Evaluation of research in context: An approach and two cases", *Research Evaluation*, 20, 1: 61-72. <https://doi.org/10.3152/095820211X12941371876346>
- De Kleijn, M., Jayabalasingham, B., Falk-Krzesinski, HJ., Collins, T., Kuiper-Hooyng, L., Cingolani, I., Zhang, J., Roberge, G., et al. (2020), "The Researcher Journey Through a Gender Lens: An Examination of Research Participation, Career Progression and Perceptions Across the Globe", *Elsevier*, Retrieved from <https://elsevier.com/insights/gender-and-diversity-in-research/researcher-journey-2020>
- De Philippis M. (2015), "Multitask Agents and Incentives: The Case of Teaching and Research for University Professors", *The Economic Journal*, 131, 636: 1643-1681. <https://doi.org/10.1093/ej/ueaa119>
- de Rijcke S., Wouters P. F., Rushforth A. D., Franssen T. P. & Hammarfelt B. (2016), "Evaluation practices and effects of indicator use - a literature review", *Research Evaluation*, 25, 2: 161-169. <https://doi.org/10.1093/reseval/rvv038>
- de Weert E. (2013), "Support for continued data collection and analysis concerning mobility patterns and career paths of researchers", Deliverable 8 - Final report MORE2 prepared for European Commission, Research Directorate-General. Directorate B - European Research Area, available at: https://euraxess.ec.europa.eu/sites/default/files/policy_library/final_report_0.pdf
- Della Sala S. & Brooks J. (2008), "Multi-authors' self-citation: A further impact factor bias?", *Cortex*, 44, 9: 1139-1145. <https://doi.org/10.1016/j.cortex.2008.07.001>

- Demetrescu C., Finocchi I., Ribichini A. & Schaerf M. (2020a), "On bibliometrics in academic promotions: a case study in computer science and engineering in Italy", *Scientometrics*, 124, 3: 2207-2228. <https://doi.org/10.1007/s11192-020-03548-9>
- Demetrescu C., Ribichini A. & Schaerf M. (2020b), "Are Italian research assessment exercises size-biased?", *Scientometrics*, 125, 1: 533-549. <https://doi.org/10.1007/s11192-020-03643-x>
- Di Franco G. (2011), *Tecniche e modelli di analisi multivariata*, Franco Angeli, Milano.
- Donnellon A., Cummings L. L. & Frost P. J. (1986), "Publishing in the Organizational Sciences", *Administrative Science Quarterly*, 31, 2: 313. <https://doi.org/10.2307/2392796>
- Dundar H. & Lewis D. R. (1998), "Determinants of research productivity in higher education", *Research in Higher Education*, 39, 6: 607-631. <https://doi.org/10.1023/A:1018705823763>
- Dunn M. B. & Jones C. (2010), "Institutional logics and institutional pluralism: The contestation of care and science logics in medical education, 1967-2005", *Administrative Science Quarterly*, 55, 1: 114-149. <https://doi.org/10.2189/asqu.2010.55.1.114>
- Dunn T. & Holtz-Eakin D. (1996), *Financial capital, human capital and the transition to self-employment: Evidence from intergeneration links*, Center for Policy Research. 452. <https://surface.syr.edu/cpr/452>
- Durante R., Labartino G. & Perotti R. (2011), "Academic Dynasties: Decentralization and Familism in the Italian Academia", *NBER WORKING PAPER SERIES, Working paper 17572*, <https://doi.org/10.3386/w17572>.
- Eco U. (1964), *Apocalittici e integrati: comunicazioni di massa e teorie della cultura di massa*, Bompiani, Milano.
- Edwards M. A. & Roy S. (2017), "Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition", *Environmental Engineering Science*, 34, 1: 51-61. <https://doi.org/10.1089/ees.2016.0223>
- el Gibari S., Gómez T. & Ruiz F. (2018), "Evaluating university performance using reference point based composite indicators", *Journal of Informetrics*, 12, 4: 1235-1250. <https://doi.org/10.1016/j.joi.2018.10.003>
- Ellegaard O. & Wallin J. A. (2015), "The bibliometric analysis of scholarly production: How great is the impact?", *Scientometrics*, 105, 3: 1809-1831. <https://doi.org/10.1007/s11192-015-1645-z>
- Enders J. (2001), "A chair system in transition: Appointments, promotions, and gate-keeping in German higher education", *Higher Education*, 41: 3-25.
- Enders J., de Boer H., File J., Jongbloed B. & Westerheijden D. (2011), *Reform of higher education in Europe*, in Enders J., de Boer H.F., Westerheijden D.F., eds., *Reform of Higher Education in Europe*. SensePublishers. https://doi.org/10.1007/978-94-6091-555-0_1
- Enders J., Kehm B. M. & Schimank U. (2015), "Turning universities into actors on quasi-markets: How new public management reforms affect academic research", *Higher Education Dynamics*, 43: 89-103. https://doi.org/10.1007/978-3-319-09677-3_5

- Epstein S. (1995), "What Can Be Done to Improve the Journal Review Process", *American Psychologist*, 50: 883-885. <https://doi.org/10.1037/0003-066X.50.10.883>
- Ertmer P. A. & Glazewski K. D. (2014), "Developing a research agenda: Contributing new knowledge via intent and focus", *Journal of Computing in Higher Education*, 26: 54-68. <https://doi.org/10.1007/S12528-013-9076-4>
- METRICS
- Espeland W. N. & Sauder M. (2007), "Rankings and reactivity: How public measures recreate social worlds", *American Journal of Sociology*, 113: 1-40. <https://doi.org/10.1086/517897>
- European Commission (2012), *Excellence, Equality and Entrepreneurialism Building Sustainable Research Careers in the European Research Area*. Publications Office of the European Union. Disponibile al link https://euraxess.ec.europa.eu/sites/default/files/policy_library/expertgrouponresearchprofession.pdf.
- European Commission (2017), *MORE3 study - Support data collection and analysis concerning mobility patterns and career paths of researchers*. Publications Office of the European Union. <https://doi.org/doi/10.2777/710643>
- European Ministers in Charge of Higher Education (1999), "The European higher education area. The Bologna declaration of 19 June, 1999". Joint Declaration of the European Ministers of Education. Available at: https://ehea.info/Upload/document/ministerial_declarations/1999_Bologna_Declaration_English_553028.pdf
- Evans C. (1995), "Choosing People: Recruitment and selection as leverage on subjects and disciplines", *Studies in Higher Education*, 20: 253-265. <https://doi.org/10.1080/03075079512331381535>
- Faber M. & Westerheijden D. (2011), *European degree structure and national reform: Constitutive dynamics of the Bologna Process*, in Enders J., de Boer H.F., Westerheijden D.F., Reform of Higher Education in Europe Sense Publishers, pp. 11-28. https://doi.org/10.1007/978-94-6091-555-0_2
- Facchini C. & Pennisi C. (2020), "Gli usi della valutazione tra regole ed algoritmi: il cambiamento degli Atenei nella percezione dei Direttori di Dipartimento", *RIV Rassegna Italiana Di Valutazione*, 74: 11-34. <https://doi.org/10.3280/RIV2019-074002>
- Fanelli D. (2010), "Do pressures to publish increase scientists' bias? An empirical support from US states data", *PLoS ONE*, 5, 4. <https://doi.org/10.1371/journal.pone.0010271>
- Fang F.C., Steen R.G. & Casadevall A. (2012), "Misconduct accounts for the majority of retracted scientific publications", *Proceedings of the National Academy of Sciences of the United States of America*, 109, 42: 17028-17033. <https://doi.org/10.1073/pnas.1212247109>
- Fasanella A. & Martire F. (2017), "Considerazioni metodologiche sulla Vqr 2011-2014 e possibili sviluppi della valutazione", *Sociologia e Ricerca Sociale*, 2017/114: 89-116. <https://doi.org/10.3280/SR2017-114005>
- Fasanella A. & Martire F., a cura di (2020), *Valutazione della ricerca e ricerca sulla valutazione. Riflessioni, analisi e proposte per la VQR*, Franco Angeli, Milano.

- Fassari L.G. & Spanò E. (2019), "Capturing phantasm in academia: Evaluation artefacts and gender", *Italian Journal of Sociology of Education*, 11, 1: 310-330. <https://doi.org/10.14658/pupj-ijse-2019-1-15>
- Favale T., Soro F., Trevisan M., Drago I. & Mellia M. (2020), "Campus traffic and e-Learning during COVID-19 pandemic", *Computer Networks*, 176: 107290. <https://doi.org/10.1016/J.COMNET.2020.107290>
- Fernández M.T., Morillo F., Bordons M. & Gómez I. (2002), "Bibliometric study of a scientific-technological area of the Spanish Research Plan (2000-2003)", *Revista Española de Documentación Científica*, 25, 4. <https://doi.org/10.3989/redc.2002.v25.i4.272>
- Ferrara E., Menczer F., Flammini A. & Radicchi F. (2015), "Quality versus quantity in scientific impact", *Journal of Informetrics*, 9, 4: 800-808. <https://doi.org/10.1016/j.joi.2015.07.008>
- Ferrer-Sapena, A., Aleixandre-Benavent, R., Peset, F. & Sánchez-Pérez, E. A. (2018), "Citations to arXiv preprints by indexed journals and their impact on research evaluation", *Journal of Information Science Theory and Practice*, 6, 4: 6-16. <https://doi.org/10.1633/JISTaP.2018.6.4.1>
- Fister, I., Fister, I. & Perc, M. (2016), "Toward the discovery of citation cartels in citation networks", *Frontiers in Physics*, 4, DEC. <https://doi.org/10.3389/fphy.2016.00049>
- Flatt, J. W., Blasimme, A. & Vayena, E. (2017), "Improving the measurement of scientific success by reporting a self-citation index", *Publications*, 5, 3. <https://doi.org/10.3390/publications5030020>
- Flecha, A., Pulido, C. & Christou, M. (2011), "Transforming Violent Selves Through Reflection in Critical Communicative Research", *Qualitative Inquiry*, 17, 3: 246-255. <https://doi.org/10.1177/1077800410397803>
- Fligstein, N. & McAdam, D. (2012), *A Theory of Fields*, Oxford University Press, USA <https://doi.org/10.1093/acprof:oso/9780199859948.001.0001>
- Fondazione CRUI. (2015). *L'utilizzo dei dati VQR per la valutazione dipartimentale negli atenei italiani Metodologie ed esperienze*. https://www2.crui.it/crui/vqr/vqr_pubblicazione_crui.pdf
- Fondazione RES (2015), "Nuovi divari. Un'indagine sulle Università del Nord e del Sud", *Nuovi Scenari. Una Indagine Sulle Università Del Nord E Del Sud*. www.resricerche.it/nuovi-divari-unindagine-sulle-universita-del-nord-e-del-sud/
- Fontana, R. & Valentini, E., a cura di (2020), *Conseguenze della valutazione. Idee e pratiche dei docenti universitari nelle scienze sociali*, Franco Angeli, Milano.
- Fowler, J. H. & Aksnes, D. W. (2007), "Does self-citation pay?", *Scientometrics*, 72, 3: 427-437. <https://doi.org/10.1007/s11192-007-1777-2>
- Fox, M. F. & Faver, C. A. (1984), "Independence and Cooperation in Research: The Motivations and Costs of Collaboration", *The Journal of Higher Education*, 55, 3: 347. <https://doi.org/10.2307/1981888>
- Franceschini, F. & Maisano, D. (2011), "Proposals for evaluating the regularity of a scientist's research output", *Scientometrics*. <https://doi.org/10.1007/s11192-011-0371-4>

- Franceschini, F. & Maisano, D. (2017), "Critical remarks on the Italian research assessment exercise VQR 2011-2014", *Journal of Informetrics*, 11, 2: 337-357. <https://doi.org/10.1016/j.joi.2017.02.005>
- Franzoni, C., Scellato, G. & Stephan, P. (2014), "The mover's advantage: The superior performance of migrant scientists", *Economics Letters*, 122, 1: 89-93. <https://doi.org/10.1016/j.econlet.2013.10.040>
- Frey, B. S., Eichenberger, R. & Frey, R. L. (2009), "Editorial Ruminations: Publishing Kyklos", *Kyklos*, 62, 2: 151-160. <https://doi.org/10.1111/J.1467-6435.2009.00428.X>
- Friedland, R. & Alford, R. R. (1991), *Bringing Society Back in: Symbols, Practices, and Institutional Contradictions*, in DiMaggio P. J., Powell W. W., eds., *The New Institutionalism in Organisational Analysis*, pp. 232-267, University of Chicago Press.
- Fumasoli T., Goastellec G. & Kehm B. M. (2015), "Academic careers and work in Europe: Trends, challenges, perspectives", *Academic Work and Careers in Europe: Trends, Challenges, Perspectives*, 201-214. https://doi.org/10.1007/978-3-319-10720-2_10/COVER
- Fyfe A., Coate K., Curry S., Lawson S., Moxham N. & Røstvik C. M. (2017), "Untangling Academic Publishing: A history of the relationship between commercial interests, academic prestige and the circulation of research", <https://doi.org/10.5281/ZENODO.546100>
- Garfield E. (1979), "Is citation analysis a legitimate evaluation tool?", *Scientometrics*, 1, 4: 359-375. <https://doi.org/10.1007/BF02019306>
- Geuna A. & Martin B. R. (2003), "University research evaluation and funding: An international comparison", *Minerva*, 41, 4: 277-304. <https://doi.org/10.1023/B:MINE.0000005155.70870.bd>
- Geuna A. & Piolatto M. (2016), "Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while)", *Research Policy*, 45, 1: 260-271. <https://doi.org/10.1016/j.respol.2015.09.004>
- Gibson A. G. & Hazelkorn E. (2017), "Arts and humanities research, redefining public benefit, and research prioritization in Ireland", *Research Evaluation*, 26, 3: 199-210. <https://doi.org/10.1093/reseval/rvx012>
- Giglioli P. P. (1979), *Baroni e burocrati : il ceto accademico italiano*. Il Mulino, Bologna.
- Gill R. (2009), "Breaking the silence: The hidden injuries of neo-liberal academia", *Feministische Studien*, 34, 1: 228-244.
- Giuffrida C., Abramo G. & D'Angelo C. A. (2019), "Are all citations worth the same? Valuing citations by the value of the citing items", *Journal of Informetrics*, 13, 2: 500-514. <https://doi.org/10.1016/j.joi.2019.02.008>
- Glänzel W. & Chi P. S. (2020), "The big challenge of Scientometrics 2.0: exploring the broader impact of scientific research in public health", *Scientometrics* 125: 1011-1031 <https://doi.org/10.1007/s11192-020-03473-x>
- Glänzel W., Thijs B. & Schlemmer B. (2004), "A bibliometric approach to the role of author self-citations in scientific communication", *Scientometrics*, 59, 1: 63-77. <https://doi.org/10.1023/B:SCIE.0000013299.38210.74>
- Gläser B. G. & Strauss A. L. (1999), *The discovery of grounded theory : strategies for qualitative research*. Routledge.

- Gläser J. (2017), “A fight on epistemological quicksand: Comment on the dispute between van den Besselaar et al. and Butler”, *Journal of Informetrics*, 11, 3: 927-932. <https://doi.org/10.1016/j.joi.2017.05.019>
- Gläser J. (2024). “From Effects to Governance to Causes of Epistemic Change”, *Minerva* (2024), <https://doi.org/10.1007/s11024-024-09526-2>
- Gläser J. & Laudel G. (2016), “Governing Science: How Science Policy Shapes Research Content”, *European Journal of Sociology / Archives Européennes de Sociologie*, 57, 1: 117-168. <https://doi.org/10.1017/S0003975616000047>
- Gläser J. & Whitley R. (2007), *The Social Orders of Research Evaluation Systems*, in Whitley R., Gläser J., eds. *The Changing Governance of the Sciences. Sociology of the Sciences Yearbook*, vol 26. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6746-4_12
- Golden J., Carstensen F. v. (1992), “Academic research productivity, department size and organization: Further results, rejoinder”, *Economics of Education Review*, 11, 2: 169-171. [https://doi.org/10.1016/0272-7757\(92\)90007-P](https://doi.org/10.1016/0272-7757(92)90007-P)
- Gomez A. (2017), “Communicative methodology and social impact”, in N. K. Denzin & M. D. Giardina (Eds.), *Qualitative Inquiry in Neoliberal Times*, pp. 166-178. Routledge.
- Gómez A., Puigvert L., Flecha R. (2011), “Critical Communicative Methodology: Informing Real Social Transformation Through Research”, *Qualitative Inquiry*, 17, 3: 235-245. <https://doi.org/10.1177/1077800410397802>
- González-Teruel A., González-Alcaide G., Barrios M., Abad-García M.-F. (2015), “Mapping recent information behavior research: an analysis of co-authorship and co-citation networks”, *Scientometrics*, 103, 2: 687-705. <https://doi.org/10.1007/s11192-015-1548-z>
- Goodrick E., Reay T. (2011), “Constellations of institutional logics: Changes in the professional work of pharmacists”, *Work and Occupations*, 38, 3: 372-416. <https://doi.org/10.1177/0730888411406824>
- Gorraiz J., Gumpenberger C. (2015), “A flexible bibliometric approach for the assessment of professorial appointments”, *Scientometrics*, 105, 3: 1699-1719. <https://doi.org/10.1007/s11192-015-1703-6>
- Greenwood R., Raynard M., Kodeih F., Micelotta E. R., Lounsbury M. (2011), “Institutional complexity and organizational responses”, *Academy of Management Annals*, 5, 1: 317-371. <https://doi.org/10.1080/19416520.2011.590299>
- Grisorio M. J., Prota F. (2020), “Italy’s national research assessment: some unpleasant effects”, *Studies in Higher Education*, 45, 4: 736-754. <https://doi.org/10.1080/03075079.2019.1693989>
- Grossi G., Dobija D., Strzelczyk W. (2020), “The Impact of Competing Institutional Pressures and Logics on the Use of Performance Measurement in Hybrid Universities”, *Public Performance and Management Review*, 43, 4: 818-844. [https://doi.org/10.1080/15309576.2019.1684328/SUPPL_FILE/MPMR_A_1684328_SM2811.DOCX](https://doi.org/10.1080/15309576.2019.1684328)
- Groves T., López E. M., Carvalho T. (2018), “The impact of international mobility as experienced by Spanish academics”, *European Journal of Higher Education*, 8, 1: 83-98. <https://doi.org/10.1080/21568235.2017.1388187>

- Guglielmi G. (2019), “Clubby and ‘disturbing’ citation behavior by researchers in Italy has surged”. Article. <https://doi.org/10.1126/science.aaz4708>
- Gunnulfsen A. E. (2021), “Applying the Integration Dimensions of Quantitative and Qualitative Methods in Education Policy Research: Lessons Learned From Investigating Micro Policymaking in Norwegian Schools”, *International Journal of Qualitative Methods*, 20: 160940692110283. <https://doi.org/10.1177/16094069211028349>
- Gupta S. (2016), “How has publishing changed in the last twenty years?”, *Notes and Records of the Royal Society of London*, 70, 4: 391. <https://doi.org/10.1098/RSNR.2016.0035>
- Gush J., Jaffe A., Larsen V., Laws A. (2018), “The effect of public funding on research output: the New Zealand Marsden Fund”, *New Zealand Economic Papers*, 52, 2: 227-248. <https://doi.org/10.1080/00779954.2017.1325921>
- Guthrie S., Lichten C., Harte E., Parks S., Wooding S. (2017), *International mobility of researchers: A survey of researchers in the UK*. Research report. RAND Corporation. <https://doi.org/10.7249/RR1991>
- Hall J., Hundley V., van Teijlingen E. (2015), “The journal editor: Friend or foe?”, *Women and Birth*, 28, 2: e26-e29. <https://doi.org/10.1016/j.wombi.2015.01.008>
- Hamilton D. G., Fraser H., Hoekstra R., Fidler F. (2020), “Journal policies and editors’ opinions on peer review”, *eLife*, 9: 1-14. <https://doi.org/10.7554/eLife.62529>
- Hammarfelt B. (2017), “Recognition and reward in the academy: Valuing publication oeuvres in biomedicine, economics and history”, *Aslib Journal of Information Management*, 69, 2: <https://doi.org/10.1108/AJIM-01-2017-0006>
- Hammarfelt B. & de Rijcke S. (2015), “Accountability in context: Effects of research evaluation systems on publication practices, disciplinary norms, and individual working routines in the faculty of Arts at Uppsala University”, *Research Evaluation*, 24, 1: 63-77. <https://doi.org/10.1093/reseval/rvu029>
- Hardeman S., Van-Roy V., Vertesy D. & Saisana M. (2013), *An analysis of national research systems (I): A composite indicator for scientific and technological research excellence*, JRC Scientific and Policy Report, Publication Office of the European Union. <https://doi.org/10.2788/95887>
- Hargens L. L. (2012), “Academic Labor Markets and Assistant Professors’ Employment Outcomes”, *Research in Higher Education*, 53, 3: 311-324. <https://doi.org/10.1007/s11162-011-9228-1>
- Hassan S. U., Imran M., Gillani U., Aljohani N. R., Bowman T. D. & Didegah F. (2017), “Measuring social media activity of scientific literature: an exhaustive comparison of scopus and novel altmetrics big data”, *Scientometrics*, 113, 2: 1037-1057. <https://doi.org/10.1007/s11192-017-2512-x>
- Hatch, M. J. (2018). *Organization theory: Modern, symbolic, and postmodern perspectives* (4th ed.). Oxford University Press.
- Haunschild R. & Bornmann L. (2018), “Field- and time-normalization of data with many zeros: an empirical analysis using citation and Twitter data”, *Scientometrics*, 116, 2: 997-1012. <https://doi.org/10.1007/s11192-018-2771-1>

- Haunschild R., Leydesdorff L., Bornmann L., Hellsten I. & Marx W. (2019), “Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags”, *Journal of Informetrics*, 13, 2: 695-707. <https://doi.org/10.1016/j.joi.2019.03.008>
- Haupt J. P. (2022), “Short-Term Internationally Mobile Academics and Their Research Collaborations Upon Return: Insights From the Fulbright U.S. Scholar Program”, *Journal of Studies in International Education*, 26, 4: 511-530. <https://doi.org/10.1177/1028315321990760>
- Haustein S. (2016), “Grand challenges in altmetrics: heterogeneity, data quality and dependencies”, *Scientometrics*, 108, 1: 413-423. <https://doi.org/10.1007/s11192-016-1910-9>
- Haustein S., Bowman T. D., Holmberg K., Tsou A., Sugimoto C. R. & Larivière V. (2016), “Tweets as impact indicators: Examining the implications of automated ‘bot’ accounts on Twitter”, *Journal of the Association for Information Science and Technology*, 67, 1: 232-238. <https://doi.org/10.1002/asi.23456>
- Haustein S. & Larivière V. (2015), “The use of bibliometrics for assessing research: Possibilities, limitations and adverse effects”, in: Welpe I., Wollersheim J., Ringelhan S., Osterloh M., eds., *Incentives and Performance*, Springer, Cham. https://doi.org/10.1007/978-3-319-09785-5_8
- Hazelkorn, E., 2010. *Pros and cons of research assessment*, in *ISSC/UNESCO World Social Science Report 2010*, UNESCO Press, Paris, pp. 255-258
- Hazelkorn E. (2015), *Rankings and the Reshaping of Higher Education: The Battle for World-Class Excellence: Second Edition*, Palgrave Macmillan London. <https://doi.org/10.1057/9781137446671>
- Heinze T. & Jappe A. (2020), “Quantitative science studies should be framed with middle-range theories and concepts from the social sciences”, *Quantitative Science Studies*, 1, 3: 983-992. https://doi.org/10.1162/qss_a_00059
- Heinze T., Shapira P., Rogers J. D. & Senker J. M. (2009), “Organizational and institutional influences on creativity in scientific research”, *Research Policy*, 38, 4: 610-623. <https://doi.org/10.1016/j.respol.2009.01.014>
- Hemlin S. (1993), “Scientific quality in the eyes of the scientist. A questionnaire study”, *Scientometrics*, 27, 1: 3-18. <https://doi.org/10.1007/BF02017752>
- Hemlin S., Rasmussen S. B. (2006), “The shift in academic quality control”, *Science Technology and Human Values*, 31, 2: 173-198. <https://doi.org/10.1177/0162243905283639>
- Hendrik P. & van Dalen K.H. (2012), “Intended and Unintended Consequences of a Publish-or-Perish Culture: A Worldwide Survey”, *Journal of the American Society for Information Science and Technology*, 63, 7: 1282-1293. <https://doi.org/10.1002/asi.22636>
- Herschberg C., Benschop Y. & van den Brink M. (2018), “Precarious postdocs: A comparative study on recruitment and selection of early-career researchers”, *Scandinavian Journal of Management*, 34, 4: 303-310. <https://doi.org/10.1016/j.scaman.2018.10.001>
- Hessels L. K. & van Lente H. (2008), “Re-thinking new knowledge production: A literature review and a research agenda”, *Research Policy*, 37, 4: 740-760. <https://doi.org/10.1016/j.respol.2008.01.008>

- Hicks D. (2010), *Overview of models of performance-based research funding systems*, in *OECD, Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*. *OECD Publishing, Paris, 2010*, pp. 23-52. <https://doi.org/10.1787/9789264094611-4-en>
- Hicks D. (2012), “Performance-based university research funding systems”, *Research Policy*, 41, 2: 251-261. <https://doi.org/10.1016/j.resp.2011.09.007>
- Hicks D. (2017), “What year? Difficulties in identifying the effect of policy on university output”, *Journal of Informetrics*, 11, 3: 933-936. <https://doi.org/10.1016/j.joi.2017.05.020>
- Hicks D. & Holbrook J. B. (2020), “A Cartography of Philosophy’s Engagement with Society”, *Minerva*, 58, 1: 25-45. <https://doi.org/10.1007/s11024-019-09384-3>
- Hicks D., Wouters P., Waltman L., de Rijcke S. & Rafols I. (2015), “Bibliometrics: The Leiden Manifesto for research metrics”, *Nature*, 520, 7548: 429-431. <https://doi.org/10.1038/520429a>
- Higgins A. C., Chubin D. E. & Hackett E. J. (1990), “Peerless Science: Peer Review and U.S. Science Policy”, *Social Forces*, 69, 2: 634. <https://doi.org/10.2307/2579681>
- Hilbe J. M. (2011), *Negative Binomial Regression*, Cambridge University Press. <https://doi.org/10.1017/CBO9780511973420>
- Hing C. B., Higgs D., Hooper L., Donell S. T. & Song F. (2011), “A survey of orthopaedic journal editors determining the criteria of manuscript selection for publication”, *Journal of Orthopaedic Surgery and Research*, 6, 1. <https://doi.org/10.1186/1749-799X-6-19>
- Hiruy K., Murphy R., Lewis T., White W. T. & Purcell S. W. (2019), “Measuring scientific impact of fisheries and aquaculture research-for-development projects in South East Asia and the Pacific”, *Research Evaluation*, 28, 4: 313-325. <https://doi.org/10.1093/reseval/rvz019>
- Honig B. & Bedi A. (2012), “The fox in the hen house: A critical examination of plagiarism among members of the academy of management”, *Academy of Management Learning and Education*, 11, 1: 101-123. <https://doi.org/10.5465/amle.2010.0084>
- Hood, C. (1995). The “new public management” in the 1980s: Variations on a theme. *Accounting, organizations and society*, 20, 2-3: 93-109. [https://doi.org/10.1016/0361-3682\(93\)E0001-W](https://doi.org/10.1016/0361-3682(93)E0001-W)
- Horbach S. P. J. M. & Halffman W. (2018), “The changing forms and expectations of peer review”, *Research Integrity and Peer Review*. <https://doi.org/10.1186/s41073-018-0051-5>
- Horta H. & Santos J. M. (2016), “An instrument to measure individuals’ research agenda setting: the multi-dimensional research agendas inventory”, *Scientometrics*, 108, 3: 1243-1265. <https://doi.org/10.1007/S11192-016-2012-4>/METRICS
- Hosmer D. W. & Lemeshow S. (2000), *Applied Logistic Regression*, Wiley. <https://doi.org/10.1002/0471722146>
- Huang F., Finkelstein M. & Rostan M., eds. (2014), *The Internationalization of the Academy*, Springer Netherlands. <https://doi.org/10.1007/978-94-007-7278-6>

- Huet S. (2019), “Recherche scientifique: Le faux miracle Italien”, *Recherche Scientifique: Le Faux Miracle Italien*, 1, 4: 598-614. <https://doi.org/10.1007/BF01390690/METRICS>
- Hunt G. E., Cleary M. & Walter G. (2010), “Psychiatry and the Hirsch h -index: The Relationship Between Journal Impact Factors and Accrued Citations”, *Harvard Review of Psychiatry*, 18, 4: 207-219. <https://doi.org/10.3109/10673229.2010.493742>
- Husu L. (2000), “Gender discrimination in the promised land of gender equality”, *Higher Education in Europe*, 25, 2: 221-228.
- Hyland K. (2003), “Self-citation and self-reference: Credibility and promotion in academic publication”, *Journal of the American Society for Information Science and Technology*, 54, 3: 251-259. <https://doi.org/10.1002/asi.10204>
- Ioannidis J. P. A. (2015), “A generalized view of self-citation: Direct, co-author, collaborative, and coercive induced self-citation”, *Journal of Psychosomatic Research*, 78, 1: 7-11. <https://doi.org/10.1016/j.jpsychores.2014.11.008>
- Ioannidis J. P. A., Baas J., Klavans R. & Boyack K. W. (2019), “A standardized citation metrics author database annotated for scientific field”, *PLoS Biology*, 17, 8. <https://doi.org/10.1371/journal.pbio.3000384>
- Istat. (2022) *Livelli di istruzione e ritorni occupazionali. Anno 2021*. www.istat.it/it/archivio/276497
- Jackall R. (1988), “Moral mazes: The world of corporate managers”, *International Journal of Politics, Culture and Society*, 1, 4: 598-614. <https://doi.org/10.1007/BF01390690/METRICS>
- Jia A. Y., Rowlinson S., Loosemore M., Gilbert D. & Ciccarelli M. (2019), “Institutional logics of processing safety in production: The case of heat stress management in a megaproject in Australia”, *Safety Science*, 120: 388-401. <https://doi.org/10.1016/j.ssci.2019.07.004>
- Jia A. Y., Rowlinson S., Loosemore M., Xu M., Li B. & Gibb A. (2017), “Institutions and institutional logics in construction safety management: the case of climatic heat stress”, *Construction Management and Economics*, 35, 6: 338-367. <https://doi.org/10.1080/01446193.2017.1296171>
- Johann D., Neufeld J., Thomas K., Rathmann J. & Rauhut H. (2024), “The impact of researchers’ perceived pressure on their publication strategies”, *Research Evaluation*. <https://doi.org/10.1093/reseval/rvae011>
- Jonkers K. & Cruz-Castro L. (2013), “Research upon return: The effect of international mobility on scientific ties, production and impact”, *Research Policy*, 42, 8: 1366-1377. <https://doi.org/10.1016/j.respol.2013.05.005>
- Jonkers K. & Tijssen R. (2008), “Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity”, *Scientometrics*, 77, 2: 309-333. <https://doi.org/10.1007/s11192-007-1971-x>
- Kacem A., Flatt J. W. & Mayr P. (2020), “Tracking self-citations in academic publishing”, *Scientometrics*, 123, 2: 1157-1165. <https://doi.org/10.1007/s11192-020-03413-9>
- Kallio K. M., Kallio T. J., Grossi G. & Engblom J. (2021), “Institutional logic and scholars’ reactions to performance measurement in universities”, *Accounting, Auditing and Accountability Journal*, 34, 9: 135-161. <https://doi.org/10.1108/AAAJ-03-2018-3400>

- Kaltenbrunner W. (2018), "Situated Knowledge Production, International Impact: Changing Publishing Practices in a German Engineering Department", *Minerva*, 56, 3: 283-303. <https://doi.org/10.1007/s11024-017-9337-x>
- Kanchan T. & Krishan K. (2019), "The Leiden Manifesto and Research Assessment", *Science and Engineering Ethics*, 25, 2: 643-644. <https://doi.org/10.1007/S11948-017-0012-2/METRICS>
- Karlsson S. (2017), "Evaluation as a travelling idea: Assessing the consequences of Research Assessment Exercises", *Research Evaluation*, 26, 2: 55-65. <https://doi.org/10.1093/reseval/rvx001>
- Karpik L. (2010), "Valuing the unique : the economics of singularities", *Princeton University Press*.
- Karvalics L. Z. (2013), "From Scientific Literacy to Lifelong Research: A Social Innovation Approach", *Communications in Computer and Information Science*, 397 CCIS: 126-133. https://doi.org/10.1007/978-3-319-03919-0_15
- Kassirer J. P. & Campion E. W. (1994), "Peer Review: Crude and Understudied, but Indispensable", *JAMA: The Journal of the American Medical Association*, 272(2): 96-97. <https://doi.org/10.1001/jama.1994.03520020022005>
- Katz J. S. & Martin B. R. (1997), "What is research collaboration?", *Research Policy*, 26, 1: 1-18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)
- Kim D. H. & Bak H.-J. (2016), "How Do Scientists Respond to Performance-Based Incentives? Evidence From South Korea", *International Public Management Journal*, 19, 1: 31-52. <https://doi.org/10.1080/10967494.2015.1032460>
- Kim J. & Diesner J. (2014), "A network-based approach to coauthorship credit allocation", *Scientometrics*, 101, 1: 587-602. <https://doi.org/10.1007/s11192-014-1253-3>
- Kim T. & Locke W. (2010), *Transnational academic mobility and the academic profession*. The Open University, Centre for Higher Education Research and Information, London.
- King M. M., Bergstrom C. T., Correll S. J., Jacquet J. & West J. D. (2017), "Men Set Their Own Cites High: Gender and Self-citation across Fields and over Time", *Socius*, 3: 1-22. <https://doi.org/10.1177/2378023117738903>
- Kivistö J. & Kohtamäki V. (2016), *Does Performance-Based Funding Work? Reviewing the Impacts of Performance-Based Funding on Higher Education Institutions*, in Pritchard R., Pausits A., Williams J., eds., *Positioning Higher Education Institutions: From Here to There* (pp. 215-226). Sense Publishers, Rotterdam. https://doi.org/10.1007/978-94-6300-660-6_12
- Kivistö J. & Mathies C. (2023), *Incentives, rationales, and expected impact: linking performance-based research funding to internal funding distributions of universities*, in Lepori B., Jongbloed B. & Hicks D., eds., *Handbook of Public Funding of Research* (pp. 186-202). Edward Elgar Publishing.
- Knorr-Cetina K. (1981), *The Manufacture of knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Pergamon Press, Oxford.
- Komorita S. S. & Graham W. K. (1965), "Number of scale points and the reliability of scales", *Educational and Psychological Measurement*, 25, 4: 987-995. <https://doi.org/10.1177/001316446502500404>

- Korytkowski P. & Kulczycki E. (2019), "Examining how country-level science policy shapes publication patterns: The case of Poland", *Scientometrics*, 119, 3: 1519-1543. <https://doi.org/10.1007/s11192-019-03092-1>
- Krippendorff K. (2019), *Content Analysis: An introduction to its Methodology* (fourth). SAGE Publications, Inc. <https://doi.org/10.4135/9781071878781>
- Kumar S., Rohani V. A. & Ratnavelu K. (2014), "International research collaborations of ASEAN Nations in economics, 1979-2010", *Scientometrics*, 101, 1: 847-867. <https://doi.org/10.1007/s11192-014-1363-y>
- Kyvik S. & Reymert I. (2017), "Research collaboration in groups and networks: differences across academic fields", *Scientometrics*, 113, 2: 951-967. <https://doi.org/10.1007/s11192-017-2497-5>
- Lamont M. (2009), *How professors think: inside the curious world of academic judgment*, Harvard University Press, Cambridge, MA.
- Langfeldt L., Nedeva M., Sörlin S. & Thomas D. A. (2020), "Co-existing Notions of Research Quality: A Framework to Study Context-specific Understandings of Good Research", *Minerva*, 58, 1: 115-137. <https://doi.org/10.1007/s11024-019-09385-2>
- Larivière V. & Costas R. (2016), "How Many Is Too Many? On the Relationship between Research Productivity and Impact", *PLOS ONE*, 11, 9: e0162709. <https://doi.org/10.1371/journal.pone.0162709>
- Larivière V., Haustein S. & Mongeon P. (2015), "The Oligopoly of Academic Publishers in the Digital Era", *PLOS ONE*, 10, 6: e0127502. <https://doi.org/10.1371/JOURNAL.PONE.0127502>
- Larivière V., Ni C., Gingras Y., Cronin B. & Sugimoto C. R. (2013), "Bibliometrics: Global gender disparities in science", *Nature*, 504, 7479: 211-213. <https://doi.org/10.1038/504211a>
- Latour B. & Woolgar S. (1979), *Laboratory Life: The Social Construction of Scientific Facts*. Sage Publications, Beverly Hills, Calif.
- Laudel G. (2006), "The art of getting funded: How scientists adapt to their funding conditions", *Science and Public Policy*, 33, 7: 489-504. <https://doi.org/10.3152/147154306781778777>
- Laudel G. & Gläser J. (2007), "Interviewing Scientists", *Science, Technology & Innovation Studies*, 3, 2: 91-111.
- Lawani S. M. (1982), "On the Heterogeneity and Classification of Author Self-Citations", *Journal of the American Society for Information Science*, 33, 5: 280-284.
- Lee S. & Bozeman B. (2005), "The Impact of Research Collaboration on Scientific Productivity", *Social Studies of Science*, 35, 5: 673-702. www.jstor.org/stable/25046667
- Lemke S., Mehrazar M., Mazarakis A. & Peters I. (2018), "Are there different types of online research impact?", *Proceedings of the Association for Information Science and Technology*, 55, 1: 282-289. <https://doi.org/10.1002/pra2.2018.14505501031>
- Lentz B. F. & Laband D. N. (1989), "Why so many children of doctors become doctors: Nepotism vs. human capital transfers", *Journal of Human Resources*, 24, 3: 396-413.

- Levitt S. D. & Dubner S. J. (2005), *Freakonomics: A rogue economist explores the hidden side of everything*, William Morrow & Co.. <https://psycnet.apa.org/record/2005-05399-000>
- Lewis J. M. (2013), *Academic governance: disciplines and policy*, Routledge, New York. <https://doi.org/10.4324/9780203754085>
- Lindahl J., Colliander C. & Danell R. (2020), “Early career performance and its correlation with gender and publication output during doctoral education”, *Scientometrics*, 122, 1: 309-330. <https://doi.org/10.1007/s11192-019-03262-1>
- Lissoni F., Mairesse J., Montobbio F. & Pezzoni M. (2011), “Scientific productivity and academic promotion: A study on French and Italian physicists”, *Industrial and Corporate Change*, 20, 1: 253-294. <https://doi.org/10.1093/icc/dtq073>
- Liu X., Wei Y. & Zhao Z. (2020), “How researchers view altmetrics: an investigation of ISSI participants”, *Aslib Journal of Information Management*, 72, 3: 361-378. <https://doi.org/10.1108/AJIM-07-2019-0165>
- Lo C.O., Tan Y.S.M., Chrostowski M., et al. (2019), “Reenvisioning gifted education in British Columbia: a Qualitative research protocol of policy analysis in the context of a redesigned curriculum”, *International Journal of Qualitative Methods*, 18, January-December 2019. <https://doi.org/10.1177/1609406918822233>
- Locke E. A. & Latham G. P. (2002), “Building a practically useful theory of goal setting and task motivation: A 35-year odyssey”, *American Psychologist*, 57, 9: 705-717. <https://doi.org/10.1037/0003-066X.57.9.705>
- Lombi L. (2015), *Le web survey*. Franco Angeli, Milano.
- Louvel S. (2010), “Changing Authority Relations within French Academic Research Units since the 1960s”, in Whitley R., Gläser J. & Engwall L., eds., *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, Oxford University Press, pp. 184-210. <https://doi.org/10.1093/acpr/oso/9780199590193.003.0006>
- Lucas L. (2006), *The research game in academic life*, SRHE and Open University Press, Maidenhead, UK.
- Lucio-Arias D. & Leydesdorff L. (2008), “Main-path analysis and path-dependent transitions in HistCiteTM-based historiograms”, *Journal of the American Society for Information Science and Technology*, 59, 12: 1948-1962. <https://doi.org/10.1002/asi.20903>
- Luukkonen T. & Thomas D. A. (2016), “The ‘Negotiated Space’ of University Researchers’ Pursuit of a Research Agenda”, *Minerva*, 54, 1: 99-127. <https://doi.org/10.1007/S11024-016-9291-Z/METRICS>
- Maassen P. & Stensaker B. (2011), “The knowledge triangle, European higher education policy logics and policy implications”, *Higher Education*, 61, 6: 757-769. <https://doi.org/10.1007/s10734-010-9360-4>
- Machado-Taylor M. de L., Soares V. M. & Teichler U., eds. (2017), “Challenges and Options: The Academic Profession in Europe”, Springer International Publishing. <https://doi.org/10.1007/978-3-319-45844-1>

- Magnani A. (2019), Boom di auto-citazioni, ecco perché la ricerca italiana è “dopata”, Il Sole 24 Ore. Articolo on-line www.ilsole24ore.com/art/boom-auto-citazioni-studio-denuncia-doping-ricerca-italiana-ACnIPdj (Consultato il 16 novembre 2023).
- Mäkinen E. I., Evans E. D. & McFarland, D. A. (2024), “Interdisciplinary Research, Tenure Review, and Guardians of the Disciplinary Order”, *The Journal of Higher Education*, 1-28. <https://doi.org/10.1080/00221546.2024.2301912>.
- Margalida A. & Colomer M. À. (2016), “Improving the peer review process and editorial quality: Key errors escaping the review and editorial process in top scientific journals”, *PeerJ*, 4:e1670. <https://doi.org/10.7717/peerj.1670>
- Margherita A., Elia G. & Petti, C. (2022), “What Is Quality in Research? Building a Framework of Design, Process and Impact Attributes and Evaluation Perspectives”, *Sustainability*, 14, 5: 3034. <https://doi.org/10.3390/SU14053034>
- Mari L. (2003), “Epistemology of measurement”, *Measurement*, 34, 1: 17-30. [https://doi.org/10.1016/S0263-2241\(03\)00016-2](https://doi.org/10.1016/S0263-2241(03)00016-2)
- Mariani M. S., Medo M. & Zhang, Y. C. (2015), “Ranking nodes in growing networks: When PageRank fails”, *Scientific Reports*, 5, 16181. <https://doi.org/10.1038/srep16181>
- Marini G. (2017), “New promotion patterns in Italian universities: Less seniority and more productivity? Data from ASN”, *Higher Education*, 73, 2: 189-205. <https://doi.org/10.1007/s10734-016-0008-x>
- Marini G. & Meschitti, V. (2018), “The trench warfare of gender discrimination: Evidence from academic promotions to full professor in Italy”, *Scientometrics*, 115, 2: 989-1006. <https://doi.org/10.1007/S11192-018-2696-8/METRICS>
- Martin B. (2009), “Academic patronage”, *International Journal for Educational Integrity*, 5, 1: 3-19.
- Martin B. R. (2016), “What’s happening to our universities?”, *Prometheus*, 34, 1: 7-24. <https://doi.org/10.1080/08109028.2016.1222123>
- Martin B. R. (2017), “When social scientists disagree: Comments on the Butler-van den Besselaar debate”, *Journal of Informetrics*, 11, 3: 937-940. <https://doi.org/10.1016/j.joi.2017.05.021>
- Martin M. (2000), *Managing university-industry relations: A study of institutional practices from 12 different countries*, International Institute for Educational Planning/ UNESCO, Paris.
- Marzolla M. (2015), “Quantitative analysis of the Italian National Scientific Qualification”, *Journal of Informetrics*, 9, 2: 285-316. <https://doi.org/10.1016/j.joi.2015.02.006>
- Mathies C., Kivistö J. & Birnbaum, M. (2020), “Following the money? Performance-based funding and the changing publication patterns of Finnish academics”, *Higher Education*, 79, 1: 21-37. <https://doi.org/10.1007/s10734-019-00394-4>
- Mauss M. (1966), *The Gift: Forms and Functions in Archaic Societies*, Presses Universitaires de France.
- Mayo E. (1933), *The human problems of an industrial civilization*, Macmillan Co.
- Mcculloch S. (2017), “Hobson’s choice: the effects of research evaluation on academics’ writing practices in England”, *Aslib Journal of Information Management*, 69, 5: 503-515. <https://doi.org/10.1108/AJIM-12-2016-0216>.

- McNay I. (2015), "Debate: Does research quality assessment increase output and give value for money?", *Public Money and Management*, 35, 1: 67-68. <https://doi.org/10.1080/09540962.2015.986888>
- MEC-Ministerio de Educación y Ciencia. (2007), "Real decreto 1312/2007, por el que se establece la acreditación nacional para el acceso a los cuerpos docentes universitarios". www.aneca.es/var/media/185245/academia_rd1312.pdf.
- Meho, L. I. & Yang, K. (2007), "Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar", *Journal of the American Society for Information Science and Technology*, 58, 13: 2105-2125. <https://doi.org/10.1002/ASI.20677>.
- Merriam S. B. & Tisdell, E. (2015), *Qualitative Research: A Guide to Design and Implementation*, Jossey-Bass, San Francisco, CA.
- Merton, R. K. (1938). Science and the social order. *Philosophy of science*, 5, 3: <321-337. <https://doi.org/10.1086/286513>
- Merton R. K. (1973a), *Priorities in scientific discovery*, in Merton R. K., *The Sociology of Science: Theoretical and Empirical Investigations*, The University of Chicago Press, Chicago, pp. 286-324.
- Merton R. K. (1973b), "The Normative Structure of Science", in Merton R. K., *The Sociology of Science: Theoretical and Empirical Investigations*, The University of Chicago Press, Chicago, pp. 267-278.
- Meyer J. W. & Rowan, B. (1977), "Institutionalized Organizations: Formal Structure as Myth and Ceremony", *American Journal of Sociology*, 83, 2: 340-363. <https://doi.org/10.1086/226550>
- Meyer R. E., Egger-Peitler I., Höllerer M. A. & Hammerschmid, G. (2014), "Of bureaucrats and passionate public managers: Institutional logics, executive identities, and public service motivation", *Public Administration*, 92, 4: 861-885. <https://doi.org/10.1111/j.1467-9299.2012.02105.x>.
- Micheli P. & Mari, L. (2014), "The theory and practice of performance measurement", *Management Accounting Research*, 25, 2: 147-156. <https://doi.org/10.1016/J.MAR.2013.07.005>
- Miguel Da Costa J., Marques S. & Santos, D. (2019), "The science of scientific agendas: insights into the agenda setting process of researchers", Istituto Universitario de Lisboa, Thesis specially presented for the fulfillment of the degree of Doctor in Psychology. <https://repositorio.iscte-iul.pt/handle/10071/19906>.
- Minelli E., Rebora, G. & Turri, M. (2008), "Structure and significance of the Italian research assessment exercise (VTR)", *European Universities in Transition: Issues, models and cases*. Edward Elgar Publishing.
- Minelli E., Rebora G. & Turri, M. (2012), "Waiting for the market: Where is the Italian university system heading?", *Higher Education Policy*, 25, 1: 131-145. <https://doi.org/10.1057/hep.2011.27>
- Minelli E., Rebora G., Turri M. & Huisman J. (2006), "The Impact of Research and Teaching Evaluation in Universities: Comparing an Italian and a Dutch case", *Quality in Higher Education*, 12, 2: 109-124. <https://doi.org/10.1080/13538320600916668>
- Mishra S., Fegley B. D., Diesner J. & Torvik V. I. (2018), "Self-citation is the hallmark of productive authors, of any gender", *PLoS ONE*, 13, 9: e0195773. <https://doi.org/10.1371/journal.pone.0195773>

- Mitroff I. I. (1974), "Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists", *American Sociological Review*, 39, 4: 579. <https://doi.org/10.2307/2094423>.
- Moed H. F. & Halevi G. (2015), "Multidimensional assessment of scholarly research impact", *Journal of the Association for Information Science and Technology*, 66, 10: 1988-2002. <https://doi.org/10.1002/asi.23314>
- Mohammadi E. & Thelwall M. (2014), "Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows", *Journal of the Association for Information Science and Technology*, 65, 8: 1627-1638. <https://doi.org/10.1002/asi.23071>
- Moher D., Naudet F., Cristea I. A., Miedema F., Ioannidis J. P. A. & Goodman S. N. (2018), "Assessing scientists for hiring, promotion, and tenure", *PLoS Biology*, 16, 3: e2004089. <https://doi.org/10.1371/journal.pbio.2004089>
- Molas-Gallart J. (2012), "Research Governance and the Role of Evaluation", *American Journal of Evaluation*, 33, 4: 583-598. <https://doi.org/10.1177/1098214012450938>
- Molas-Gallart J. (2015), "Research evaluation and the assessment of public value", *Arts and Humanities in Higher Education*, 14, 1: 111-126. <https://doi.org/10.1177/1474022214534381>
- Morawska E. (2018), *Qualitative migration research: Viable goals, open-ended questions, and multidimensional answers*, in Zapata-Barrero R., Yalaz E., eds. *Qualitative Research in European Migration Studies*. IMISCOE Research Series. Springer, Cham. https://doi.org/10.1007/978-3-319-76861-8_7
- Moscati R. (2001), "Italian university professors in transition", *Higher Education*, 41, 1/2: 103-129. <https://doi.org/10.1023/A:1026766829752>
- Mula-Falcón J. & Caballero K. (2023), "Early career academic's odyssey: A narrative study of her professional identity construction", *Research Evaluation*, 32, 2: 458-466. <https://doi.org/10.1093/reseval/rvad005>
- Mullen C. A., Pryor C. R., Browne-Ferrigno T. & Harris S. L. (2013), "An International Call for Democratizing the Academic Journal Culture From a Community of Editors", *Interchange*, 44, 179-201 <https://doi.org/10.1007/s10780-014-9206-6>
- Müller R. & de Rijcke S. (2017), "Exploring the epistemic impacts of academic performance indicators in the life sciences", *Research Evaluation*, 26, 3: 157-168. <https://doi.org/10.1093/reseval/rvx023>
- Murias P., de Miguel J. C. & Rodríguez D. (2008), "A composite indicator for university quality assessment: The case of Spanish higher education system", *Social Indicators Research*, 89, 1: 129-146. <https://doi.org/10.1007/s11205-007-9226-z>
- Musselin C. (2004), "Towards a European Academic Labour Market? Some Lessons Drawn from Empirical Studies on Academic Mobility", *Higher Education*, 48, 1: 55-78. <https://doi.org/10.1023/B:HIGH.0000033770.24848.41>
- Musselin C. (2013), "How peer review empowers the academic profession and university managers: Changes in relationships between the state, universities and the professoriate", *Research Policy*, 42, 5: 1165-1173. <https://doi.org/10.1016/j.respol.2013.02.002>
- Nagin D. S., Rebitzer J. B., Sanders S. & Taylor L. J. (2002), "Monitoring, motivation, and management: The determinants of opportunistic behavior in

- a field experiment”, *American Economic Review*, 92, 4: 850-873. <https://doi.org/10.1257/00028280260344498>.
- Nardon L., Hari A. & Aarma K. (2021), “Reflective Interviewing—Increasing Social Impact through Research”, *International Journal of Qualitative Methods*, 20: 160940692110652. <https://doi.org/10.1177/16094069211065233>.
- Neave G. (1988), “On the Cultivation of Quality, Efficiency and Enterprise: An Overview of Recent Trends in Higher Education in Western Europe, 1986-1988”, *European Journal of Education*, 23, 1/2: 7. <https://doi.org/10.2307/1502961>
- Neave G. (1998), “The Evaluative State Reconsidered”, *European Journal of Education*, 33, 3: 265-284.
- Neave G. (2012), *The Evaluative State, Institutional Autonomy and Re-engineering Higher Education in Western Europe*, Palgrave Macmillan London.
- Neill U. S. (2008), “Publish or perish, but at what cost?”, *Journal of Clinical Investigation*, 118, 7: 2368. <https://doi.org/10.1172/JCI36371>
- Nicholas D., Rodríguez-Bravo B., Watkinson A., Boukacem-Zeghmouri C., Herman E., Xu J., Abrizah A. & Świgon M. (2017), “Early career researchers and their publishing and authorship practices”, *Learned Publishing*, 30, 3: 205-217. <https://doi.org/10.1002/leap.1102>.
- Nielsen M. (2011), *Reinventing discovery. The new era of networked science*, Princeton University Press. <https://doi.org/10.1515/9781400839452>
- Nielsen M. W. (2016), “Gender inequality and research performance: moving beyond individual-meritocratic explanations of academic advancement”, *Studies in Higher Education*, 41, 11: 2044-2060. <https://doi.org/10.1080/03075079.2015.1007945>
- Northam S., Greer D. B., Rath L. & Toone A. (2014), “Nursing journal editor survey results to help nurses publish”, *Nurse Educator*, 39, 6: 290-297. <https://doi.org/10.1097/NNE.0000000000000086>
- Noruzi A. (2005), “Google Scholar: the new generation of citation indexes”, *International Journal of Libraries and Information Services*, 55, 4: 170-180. <https://doi.org/10.1515/LIBR.2005.170>
- Nykl M., Ježek K., Fiala D., Dostal M. (2014), “PageRank variants in the evaluation of citation networks”, *Journal of Informetrics*, 8, 3: 683-692. <https://doi.org/10.1016/j.joi.2014.06.005>
- OECD (2008), *Handbook on constructing composite indicators: methodology and user guide - OECD*. Testo disponibile al sito: www.oecd.org/els/soc/handbookonconstructingcompositeindicatorsmethodologyanduserguide.htm (Consultato il 8 marzo 2023)
- OECD (2010), *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, OECD Publishing, Paris. <https://doi.org/10.1787/9789264094611-en>.
- Osterloh M. & Frey B. S. (2015), “Ranking Games”, *Evaluation Review*, 39, 1: 102-129. <https://doi.org/10.1177/0193841X14524957>.
- Osuna C., Cruz-Castro L., & Sanz-Menéndez L. (2011), “Overturning some assumptions about the effects of evaluation systems on publication performance”, *Scientometrics*, 86, 3: 575-592. <https://doi.org/10.1007/s11192-010-0312-7>

- Palumbo L. (2002), “Verità sofistica e verità filosofica nel Teeteto di Platone”, In Casertano G., a cura di, *Il Teeteto di Platone: struttura e problematiche*, Loffredo Editore.
- Palumbo M. & Pennisi C. (2015), “Criteri corretti e condivisi per una valutazione buona e utile della ricerca”, *Sociologia e Politiche Sociali*, 2: 73-89. <https://doi.org/10.3280/SP2015-002004>
- Paradeise, C., Bleiklie, I., Enders, J., Goastellec, G., Michelsen, S., Reale, E., et al. (2009). *Reform policies and change processes in Europe*. In Huisman J., ed., *International perspectives on the governance of higher education. Alternative frameworks for coordination*. New York: Routledge.
- Park H. & Park H. W. (2018), “Research evaluation of Asian countries using altmetrics: comparing South Korea, Japan, Taiwan, Singapore, and China”, *Scientometrics*, 117, 771-788. <https://doi.org/10.1007/s11192-018-2884-6>
- Patrício M. T., Santos P., Loureiro P. M. & Horta H. (2018), “Faculty-exchange programs promoting change: motivations, experiences, and influence of participants in the Carnegie Mellon University-Portugal Faculty Exchange Program”, *Tertiary Education and Management*, 24, 1: 1-18. <https://doi.org/10.1080/13583883.2017.1305440>
- Patti F. P. (2019), “Career Paths into Legal Academia in Italy”, *The Rabel Journal of Comparative and International Private Law*, 84, 2020, Bocconi Legal Studies Research Paper No. #3895852, Available at SSRN: <https://ssrn.com/abstract=3895852>
- Pautasso M. (2015), “The Italian university habilitation and the challenge of increasing the representation of women in academia”, *Challenges*, 6, 1: 26-41. <https://doi.org/10.3390/challe6010026>
- Pearse N. (2011), “Deciding on the Scale Granularity of Response Categories of Likert type Scales: The Case of a 21 Point Scale”, *Electronic Journal of Business Research Methods*, 9, 2: 159-171. <https://academic-publishing.org/index.php/ejbrm/article/view/1278>
- Peroni S., Ciancarini P., Gangemi A., Nuzzolese A. G., Poggi F. & Presutti V. (2020), “The practice of self-citations: a longitudinal study”, *Scientometrics*, 123, 1: 253-282. <https://doi.org/10.1007/s11192-020-03397-6>
- Perotti R. (2008), *L'università truccata*, Einaudi, Torino.
- Persson O., Glänzel W. & Danell R. (2004), “Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies”, *Scientometrics*, 60, 3: 421-432. <https://doi.org/10.1023/B:SCIE.0000034384.35498.7d>
- Petersen A. M., Pan R. K., Pammolli F. & Fortunato S. (2019), “Methods to account for citation inflation in research evaluation”, *Research Policy*, 48, 7: 1855-1865. <https://doi.org/10.1016/j.respol.2019.04.009>
- Pichappan P. & Sarasvady S. (2002), “The other side of the coin: The intricacies of author self-citations”, *Scientometrics*, 54, 2: 285-290. <https://doi.org/10.1023/A:1016070029935>
- Pinheiro D., Melkers J. & Youtie J. (2014), “Learning to play the game: Student publishing as an indicator of future scholarly success”, *Technological Forecasting and Social Change*, 81, 1: 56-66. <https://doi.org/10.1016/J.TECHFORE.2012.09.008>

- Platone, *Opere complete*, vol.II, trad. it. Manara Valgimigli, Laterza, Bari 1998.
- Polanyi, M. (1958). *Personal Knowledge: Towards a Post-critical Philosophy*. Routledge and Kegan Paul, London.
- Pontille D. & Torny D. (2015), "From Manuscript Evaluation to Article Valuation: The Changing Technologies of Journal Peer Review", *Human Studies*, 38, 1: 57-79. <https://doi.org/10.1007/s10746-014-9335-z>
- Popkin R. H. (1989), "The scholarly communication process in the humanities: The role of the editor", *Serials Librarian*, 17, 3-4: 25-32. https://doi.org/10.1300/J123v17n03_03
- Price E. G., Gozu A., Kern D. E., Powe N. R., Wand G. S., Golden S. & Cooper L. A. (2005), "The role of cultural diversity climate in recruitment, promotion, and retention of faculty in academic medicine", *Journal of General Internal Medicine*, 20, 7: 565-571. <https://doi.org/10.1111/j.1525-1497.2005.0127.x>
- Prieto-Flores Ó., Feu J. & Casademont X. (2016), "Assessing Intercultural Competence as a Result of Internationalization at Home Efforts", *Journal of Studies in International Education*, 20, 5: 437-453. <https://doi.org/10.1177/1028315316662977>
- Primack R. B., Regan T. J., Devictor V., Zipf L., Godet L., Loyola R., Maas B., Pakeman R. J., Cumming G. S., Bates A. E., Pejchar L. & Koh L. P. (2019), "Are scientific editors reliable gatekeepers of the publication process?", *Biological Conservation*, 238: 108232. <https://doi.org/10.1016/j.biocon.2019.108232>
- Pruvot E. B. & Estermann T. Lisi V. (2019), "Public Funding Observatory". Testo disponibile al sito: <https://eua.eu/101-projects/586-public-funding-observatory.html> (Consultato il 16 novembre 2023).
- Puigvert L. (2012), "The Dialogic Turn: Dialogue or Violence?", *International and Multidisciplinary Journal of Social Sciences*, 1, 1: 78-96. <https://doi.org/10.4471/rimcis.2012.04>
- Rafols I., Leydesdorff L., O'Hare A., Nightingale P. & Stirling A. (2012), "How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management", *Research Policy*, 41, 7: 1262-1282. <https://doi.org/10.1016/j.respol.2012.03.015>
- Rauhvargers A. (2011), "Global university rankings and their impact". Testo disponibile al sito: www.ehea.info/media.ehea.info/file/EUA_seminar_June_2011/76/5/Rankings_Seminar_Rauhvargers_Presentation_of_Report_607765.pdf (Consultato il 2 Febbraio 2023)
- Reale E. (2017), *Analysis of National Public Research Funding - PREF. Final Report*, Publications Office of the European Union, Luxembourg.
- Reale E., Barbara A. & Costantini A. (2007), "Peer review for the evaluation of academic research: lessons from the Italian experience", *Research Evaluation*, 16, 3: 216-228. <https://doi.org/10.3152/095820207X227501>
- Reale E. & Marini G. (2017), *The Transformative Power of Evaluation on University Governance*. In Bleiklie I., Enders J. & Lepori B., eds., *Managing Universities: Policy and Organizational Change in a Western European Comparative Perspective*. Palgrave Macmillan Cham, pp. 107-137. https://doi.org/10.1007/978-3-319-53865-5_5

- Reale E., Morettini L. & Zinilli A. (2019), "Moving, remaining, and returning: international mobility of doctorate holders in the social sciences and humanities", *Higher Education*, 78, 1: 17-32. <https://doi.org/10.1007/s10734-018-0328-0>
- Reale E. & Pennisi C. (2013), *La valutazione dell'Università e della ricerca in Italia: stato dell'arte e prospettive*, in Vergani A., a cura di, *Prove di valutazione. Libro Bianco sulla valutazione in Italia*, Franco Aneli, Milano, pp. 23-62.
- Reale E. & Seeber M. (2013), "Instruments as empirical evidence for the analysis of Higher Education policies", *Higher Education*, 65, 1: 135-151. <https://doi.org/10.1007/s10734-012-9585-5>
- Reale E. & Zinilli A. (2017), "Evaluation for the allocation of university research project funding: Can rules improve the peer review?", *Research Evaluation*, 26, 3: 190-198. <https://doi.org/10.1093/reseval/rvx019>
- Reardon J. (2008), "Barriers to entry: heterodox publishing in mainstream journals", *On the Horizon*, 16, 4: 185-197. <https://doi.org/10.1108/10748120810912510>
- Reategui E., Pires A., Carniato M. & Franco S. R. K. (2020), "Evaluation of Brazilian research output in education: confronting international and national contexts", *Scientometrics*, 125, 1: 427-444. <https://doi.org/10.1007/s11192-020-03617-z>
- Reay T. & Hinings B. C. R. (2005), "The recombination of an organizational field: Health care in Alberta", *Organization Studies*, 26, 3: 351-384. <https://doi.org/10.1177/0170840605050872>
- Reay T. & Hinings B. C. R. (2009), "Managing the rivalry of competing institutional logics", *Organization Studies*, 30, 6: 629-652. <https://doi.org/10.1177/0170840609104803>
- Rebora G. & Turri M. (2013), "The UK and Italian research assessment exercises face to face", *Research Policy*, 42, 9: 1657-1666. <https://doi.org/10.1016/j.respol.2013.06.009>
- Reed M. S., Ferré M. Martin-Ortega J., Blanche R., Lawford-Rolfe R., Dallimer M. & Holden J., (2021), "Evaluating impact from research: A methodological framework", *Research Policy*, 50, 4: 104147. <https://doi.org/10.1016/j.respol.2020.104147>
- Rees T. (2004), *Measuring excellence in scientific research: The UK research assessment exercise*, in *Gender and Excellence in the Making*, European Commission, Brussels, pp. 117-123.
- Regan Á. & Henchion M. (2019), "Making sense of altmetrics: The perceived threats and opportunities for academic identity", *Science and Public Policy*, 46, 4: 479-489. <https://doi.org/10.1093/scipol/scz001>
- Robinson-Garcia N., van Leeuwen T. N. & Ràfols I. (2018), "Using altmetrics for contextualised mapping of societal impact: From hits to networks", *Science and Public Policy*, 45, 6: 815-826. <https://doi.org/10.1093/SCIPOL/SCY024>
- Rosa M. J. & Amaral A. (2007), "A Self-assessment of Higher Education Institutions from the Perspective of the EFQM Excellence Model", *Higher Education Dynamics*, 20: 181-207. https://doi.org/10.1007/978-1-4020-6012-0_7/COVER

- Rosenblum S., st. Clair R. L., Isett K. R. & Johnson R. (2020), "Publishing while editor: Transparency and behaviour in public administration journals", *Australian Journal of Public Administration*, 79, 4: 407-425. <https://doi.org/10.1111/1467-8500.12411>
- Ross-Hellauer T., Deppe A. & Schmidt B. (2017), "Survey on open peer review: Attitudes and experience amongst editors, authors and reviewers", *PLoS ONE*, 12, 12: e0189311. <https://doi.org/10.1371/journal.pone.0189311>
- Rossi P. (2016), "Luci e ombre della VQR", *Analysis*, 18: 7-11.
- Rostan M., a cura di (2011), *La professione accademica in Italia. Aspetti, problemi e confronti nel contesto europeo*, LED - Edizioni Universitarie di Lettere Economia Diritto.
- Ruiz-Eugenio L., Puigvert L. Ríos O. & Cisneros R. M. (2020), "Communicative Daily Life Stories: Raising Awareness About the Link Between Desire and Violence", *Qualitative Inquiry*, 26, 8-9: 1003-1009. <https://doi.org/10.1177/1077800420938880>
- Rushforth A. & de Rijcke S. (2015), "Accounting for Impact? The Journal Impact Factor and the Making of Biomedical Research in the Netherlands", *Minerva*, 53, 2: 117-139. <https://doi.org/10.1007/s11024-015-9274-5>
- Sà C. M., Kretz A. & Sigurdson K. (2013), "Accountability, performance assessment, and evaluation: Policy pressures and responses from research councils", *Research Evaluation*, 22, 2: 105-117 <https://doi.org/10.1093/reseval/rvs041>
- Salimi N. (2017), "Quality assessment of scientific outputs using the BWM", *Scientometrics*, 112: 195-213 <https://doi.org/10.1007/s11192-017-2284-3>
- Sanders J. & Munford R. (2005), "Activity and Reflection", *Qualitative Social Work*, 4, 2: 197-209. <https://doi.org/10.1177/1473325005052393>
- Santos J. M. & Horta H. (2018), "The research agenda setting of higher education researchers", *Higher Education*, 76, 4: 649-668. <https://doi.org/10.1007/s10734-018-0230-9>
- Santos J. M. & Horta H. (2020), "The Association between Researchers' Conceptions of Research and Their Strategic Research Agendas", *Journal of Data and Information Science*, 5, 4: 56-74. <https://doi.org/10.2478/JDIS-2020-0032>
- Sanz-Menéndez L. & Cruz-Castro L. (2019), "University academics' preferences for hiring and promotion systems", *European Journal of Higher Education*, 9, 2: 153-171. <https://doi.org/10.1080/21568235.2018.1515029>
- Sauder M. & Espeland W. N. (2009), "The discipline of rankings: Tight coupling and organizational change", *American Sociological Review*, 74, 1: 63-82. <https://doi.org/10.1177/000312240907400104>
- Sayed O. H. (2019), "Critical Treatise on University Ranking Systems", *Open Journal of Social Sciences*, 7, 12: 39-51. <https://doi.org/10.4236/OJSS.2019.712004>
- Scarpa F., Bianco V. & Tagliafico L. A. (2018), "The impact of the national assessment exercises on self-citation rate and publication venue: an empirical investigation on the engineering academic sector in Italy", *Scientometrics*, 117, 2: 997-1022. <https://doi.org/10.1007/s11192-018-2913-5>

- Schneider J. W., Aagaard K. & Bloch C. W. (2016), "What happens when national research funding is linked to differentiated publication counts? A comparison of the Australian and Norwegian publication-based funding models", *Research Evaluation*, 25, 3: 244-256. <https://doi.org/10.1093/reseval/rvv036>
- Seeber M., Cattaneo M., Meoli M. & Malighetti P. (2019), "Self-citations as strategic response to the use of metrics for career decisions", *Research Policy*, 48, 2: 478-491. <https://doi.org/10.1016/j.respol.2017.12.004>
- Seglen P. O. (1997), "Citations and journal impact factors: questionable indicators of research quality", *Allergy*, 52, 11: 1050-1056. <https://doi.org/10.1111/J.1398-9995.1997.TB00175.X>
- Seglen P. O. & Aksnes D. W. (2000), "Scientific productivity and group size: A bibliometric analysis of Norwegian microbiological research", *Scientometrics*, 49, 1: 125-143. <https://doi.org/10.1023/A:1005665309719>
- Serpa S., Sá M. J., Santos A. I. & Ferreira C. M. (2020), "Challenges for the academic editor in the scientific publication", *Academic Journal of Interdisciplinary Studies*, 9, 3: 12-16. <https://doi.org/10.36941/ajis-2020-0037>
- Shalley C. E. (1995), "Effects of Coaction, Expected Evaluation, and Goal Setting on Creativity and Productivity", *Academy of Management Journal*, 38, 2: 483-503. <https://doi.org/10.2307/256689>
- Shapiro D. L. (2017), "2016 Presidential Address: Making the Academy Full-Voice Meaningful", *Academy of Management*, 42, 2: 165-173. <https://doi.org/10.5465/AMR.2016.0539>
- Shattock M. (2014), "Can we still speak of there being an academic profession?", *History of Education*, 43, 6: 727-739. <https://doi.org/10.1080/0046760X.2014.964008>
- Shelomi M. (2014), "Editorial Misconduct—Definition, Cases, and Causes", *Publications*, 2, 2: 51-60. <https://doi.org/10.3390/publications2020051>
- Shore C. & Wright S. (2004), "Whose accountability? Governmentality and the auditing of universities". *Parallax*, 10, 2: 100-116. <https://doi.org/10.1080/1353464042000208558/ASSET//CMS/ASSET/18216D2A-D160-4B2B-95CC-FEA38A3D1C39/1353464042000208558.FP.PNG>
- Shriberg M. (2002), "Institutional assessment tools for sustainability in higher education", *International Journal of Sustainability in Higher Education*, 3, 3: 254-270. <https://doi.org/10.1108/14676370210434714>
- Shrivastava R. & Mahajan P. (2017), "An altmetric analysis of ResearchGate profiles of physics researchers: A study of University of Delhi (India)", *Performance Measurement and Metrics*, 18, 1: 52-66. <https://doi.org/10.1108/PMM-07-2016-0033>
- Sidiropoulos A., Katsaros D. & Manolopoulos Y., (2007), "Generalized Hirsch h-index for disclosing latent facts in citation networks", *Scientometrics*, 72, 2: 253-280. <https://doi.org/10.1007/s11192-007-1722-z>
- Simon R. J., Clark S. M. & Tifft L. L. (1966), "Of nepotism, marriage, and the pursuit of an academic career", *Sociology of Education*, 39, 4: 344-358.
- Singh Chawla D. (2019), *Researchers in Italy are increasingly citing researchers in Italy*, Articolo on-line sul sito di Physics Today. <https://doi.org/10.1063/PT.6.2.20190919a>

- Slaughter S. & Leslie L. L. (1997), *Academic Capitalism: Politics, Policies, and the Entrepreneurial University*, The Johns Hopkins University Press, Baltimore.
- Snyder H. & Bonzi S. (1998), “Patterns of self-citation across disciplines (1980-1989)”, *Journal of Information Science*, 24, 6: 431-435. <https://doi.org/10.1177/016555159802400606>
- Sobey A. J., Townsend N. C., Metcalf C. D. Bruce K. D. & Fazi F. M. (2013), “Incorporation of Early Career Researchers within multidisciplinary research at academic institutions”, *Research Evaluation*, 22, 3: 169-178. <https://doi.org/10.1093/reseval/rvt004>
- Sperling J. (2000), *Rebel with a cause: The Entrepreneur Who Created the University of Phoenix and the For-Profit Revolution in Higher Education*, John Wiley & Sons.
- Spier R. (2002). “The history of the peer review process”. In *Trends in biotechnology*, 20, 8: 357-358. [https://doi.org/10.1016/S0167-7799\(02\)01985-6](https://doi.org/10.1016/S0167-7799(02)01985-6)
- Spinello A. O., Reale E. & Zinilli A. (2021). “Design and implementation of a web survey on the effects of evaluation on academic research”, *CNR-IRCrES Working Paper*, 3/2021. <http://dx.doi.org/10.23760/2421-7158.2021.003>
- Squires B. P. (1991). “The role of the editor in peer review”. *Serials Librarian*, 19, 3-4: 87-89. https://doi.org/10.1300/J123v19n03_10
- Stål H. (2011). “Examining the Relationship between Emerging and Prevailing Institutional Logics in an Early Stage of Institutional Entrepreneurship”. *Journal of Change Management*, 11, 4: 421-443. <https://doi.org/10.1080/14697017.2011.627365>
- Stan M. R. (2023). *Insights into evaluation of research quality: what role for the researcher?*. Tesi di dottorato. LIUC-Università Cattaneo.
- Stanley J. M., Capers C. F. & Berlin L. E. (2007). “Changing the Face of Nursing Faculty: Minority Faculty Recruitment and Retention”. *Journal of Professional Nursing*, 23, 5: 253-261. <https://doi.org/10.1016/j.profnurs.2007.07.005>
- Stephan P. (2015). *How economics shapes science*, Harvard University Press.
- Stephan P., Veugelers R. & Wang J. (2017). Reviewers are blinkered by bibliometrics. *Nature*, 544(7651), 411-412. <https://doi.org/10.1038/544411a>
- Street M. D., Bozeman D. P. & Whitfield J. M. (1998). “Author perceptions of positive and negative editor behaviors in the manuscript review process”. *Journal of Social Behavior and Personality*, 13, 1: 1-22.
- Strozzi F., Colicchia C., Creazza A. & Noè C. (2017). “Literature review on the ‘smart factory’ concept using bibliometric tools”, *International Journal of Production Research* 55, 22: 6572-6591 <https://doi.org/10.1080/00207543.2017.1326643>
- Szomszor M., Pendlebury D. A. & Adams J. (2020). “How much is too much? The difference between research influence and self-citation excess”. *Scientometrics*, 123, 1119-1147. <https://doi.org/10.1007/s11192-020-03417-5>
- Tahamtan I. & Bornmann L. (2018). “Core elements in the process of citing publications: Conceptual overview of the literature”, *Journal of Informetrics*, 12, 1: 203-216. <https://doi.org/10.1016/j.joi.2018.01.002>

- Tahamtan I. & Bornmann L. (2019). "What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018". *Scientometrics*, 121: 1635-1684. <https://doi.org/10.1007/s11192-019-03243-4>
- Tahamtan I., Safipour Afshar A. & Ahamdzadeh K. (2016). "Factors affecting number of citations: a comprehensive review of the literature". *Scientometrics*, 107: 1195-1225. <https://doi.org/10.1007/s11192-016-1889-2>
- Teichler U. (2017). "Internationally mobile academics: concept and findings in Europe". *European Journal of Higher Education*, 7, 1, 15-28. <https://doi.org/10.1080/21568235.2017.1254921>
- Teichler U. & Höhle E. A., eds. (2013). *The Work Situation of the Academic Profession in Europe: Findings of a Survey in Twelve Countries*, Springer Netherlands. <https://doi.org/10.1007/978-94-007-5977-0>
- Tellado I., Lepori B. & Morla-Folch T. (2020), "WIEGO: Communicative Daily Life Stories to Assess Social Impact in the Lives of Informal Workers", *Qualitative Inquiry*, 26, 8-9: 962-969. <https://doi.org/10.1177/1077800420938680>
- Tennant J. P., Dugan J. M., Graziotin D., et al. (2017), "A multi-disciplinary perspective on emergent and future innovations in peer review", *F1000Research*, 6:1151 <https://doi.org/10.12688/f1000research.12037.3>
- Tetlock P. E. (1985), "Accountability: A Social Check on the Fundamental Attribution Error", *Social Psychology Quarterly*, 48, 3: 227. <https://doi.org/10.2307/3033683>
- Thelwall M. (2017), "Are Mendeley reader counts high enough for research evaluations when articles are published?", *Aslib Journal of Information Management*, 69, 2: 174-183. <https://doi.org/10.1108/AJIM-01-2017-0028>
- Thelwall M., Kousha K., Wilson P., Makita M., Abdoli M., Stuart E., Levitt J., Knoth P. & Cancellieri M. (2023), "Predicting article quality scores with machine learning: The U.K. Research Excellence Framework", *Quantitative Science Studies*, 4, 2: 547-573. https://doi.org/10.1162/qss_a_00258
- Thomas D. & Nedeva M. (2012), "Characterizing researchers to study research funding agency impacts: The case of the European Research Council's Starting Grants", *Research Evaluation*, 21, 4: 257-269. <https://doi.org/10.1093/reseval/rvs020>
- Thombs B. D., Levis A. W., Razykov I., Syamchandra A., Leentjens A. F. G., Levenson J. L. & Lumley M. A. (2015), "Potentially coercive self-citation by peer reviewers: A cross-sectional study", *Journal of Psychosomatic Research*, 78, 1: 1-6. <https://doi.org/10.1016/j.jpsychores.2014.09.015>
- Thornton P. H. (2004), *Markets from Culture: Institutional Logics and Organizational Decisions in Higher Education Publishing*. Stanford University Press. <https://doi.org/10.1515/9781503619098>
- Thornton P. H. & Ocasio W. (1999), "Institutional logics and the historical contingency of power in organizations: Executive succession in the higher education publishing industry, 1958-1990", *American Journal of Sociology*, 105, 3: 801-843. <https://doi.org/10.1086/210361>
- Thornton P. H. & Ocasio W. (2012), *Institutional Logics*, in Greenwood R., Oliver C., Suddaby R. & Sahlin K., *The SAGE Handbook of Organizational*

- Institutionalism*, SAGE Publications Ltd, pp. 99-128. <https://doi.org/10.4135/9781849200387.n4>
- Timmermans S. & Oh H. (2010), “The Continued Social Transformation of the Medical Profession”, *Journal of Health and Social Behavior*, 51, 1_suppl: S94-S106. <https://doi.org/10.1177/0022146510383500>
- Tonta Y. & Akbulut M. (2020), “Does monetary support increase citation impact of scholarly papers?”, *Scientometrics*, 125, 2: 1617-1641. <https://doi.org/10.1007/s11192-020-03688-y>
- Torres-Salinas D., Arroyo-Machado W. & Robinson-Garcia N. (2023), “Bibliometric denialism”, *Scientometrics*, 128: 5357-5359. <https://doi.org/10.1007/s11192-023-04787-2>
- Tourangeau R. & Yan T. (2007), “Sensitive Questions in Surveys”, *Psychological Bulletin*, 133, 5: 859-883. <https://doi.org/10.1037/0033-2909.133.5.859>
- Townley B., Cooper D. J. & Oakes L. (2003), “Performance Measures and the Rationalization of Organizations”, *Organization Studies*, 24, 7: 1045-1071. <https://doi.org/10.1177/01708406030247003>
- Tran T., Hoang K. L., La V. P., Ho M. T. & Vuong Q. H. (2020), “Scrambling for higher metrics in the Journal Impact Factor bubble period: A real-world problem in science management and its implications”, *Problems and Perspectives in Management*, 18, 1: 48-56. [https://doi.org/10.21511/ppm.18\(1\).2020.05](https://doi.org/10.21511/ppm.18(1).2020.05)
- Trotman C.-A., Bennett E., Scheffler N. & Tulloch J. C. (2002), “Faculty recruitment, retention, and success in dental academia”, *American Journal of Orthodontics and Dentofacial Orthopedics*, 122, 1: 2-8. <https://doi.org/10.1067/mod.2002.124996>
- van Balen B. & van Arensbergen P., van der Weijden I., van den Besselaar P. (2012), “Determinants of Success in Academic Careers”, *Higher Education Policy*, 25, 3: 313-334. <https://doi.org/10.1057/hep.2012.14>
- van Dalen H. P. & Henkens K. (2012), “Intended and unintended consequences of a publish-or-perish culture: A worldwide survey”, *Journal of the American Society for Information Science and Technology*, 63, 7: 1282-1293. <https://doi.org/10.1002/asi.22636>
- van den Besselaar P., Hemlin S. & van der Weijden I. (2012), “Collaboration and competition in research”, *Higher Education Policy*, 25, 3: 263-266. <https://doi.org/10.1057/HEP.2012.16/METRICS>
- van den Besselaar P., Heyman U. & Sandström U. (2017), “Perverse effects of output-based research funding? Butler’s Australian case revisited”, *Journal of Informetrics*, 11, 3: 905-918. <https://doi.org/10.1016/j.joi.2017.05.016>
- van der Ploeg F. & Veugelers R. (2008), “Towards evidence-based reform of European universities”, *CESifo Economic Studies*, 54, 2: 99-120. <https://doi.org/10.1093/cesifo/ifn015>
- van der Weijden I., Belder R., van Arensbergen P. & van den Besselaar P. (2015), “How do young tenured professors benefit from a mentor? Effects on management, motivation and performance”, *Higher Education*, 69, 2: 275-287. <https://doi.org/10.1007/s10734-014-9774-5>
- van Gestel N. & Hillebrand B. (2011), “Explaining stability and change: The rise and fall of logics in pluralistic fields”, *Organization Studies*, 32, 2: 231-252. <https://doi.org/10.1177/0170840610397475>

- van Noorden R. (2019), *Italy's rise in research impact pinned on 'citation doping'*, Articolo on-line sul sito web di Nature. <https://doi.org/10.1038/d41586-019-02725-y>
- van Raan A. F. J. (1996), "Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises", *Scientometrics*, 36: 397-420. <https://doi.org/10.1007/BF02129602>
- van Raan A. F. J. (2008), "Self-citation as an impact-reinforcing mechanism in the science system", *Journal of the American Society for Information Science and Technology*, 59, 10: 1631-1643. <https://doi.org/10.1002/asi.20868>
- van Rijnsoever F. J., Hessels L. K. & Vandeberg R. L. J. (2008), "A resource-based view on the interactions of university researchers", *Research Policy*, 37, 8: 1255-1266. <https://doi.org/10.1016/j.respol.2008.04.020>
- van Vught F. A. & Westerheijden D. F. (1994), "Towards a general model of quality assessment in higher education", *Higher Education*, 28, 3: 355-371. <https://doi.org/10.1007/BF01383722>
- Ventura O. N. & Mombrú A. W. (2006), "Use of bibliometric information to assist research policy making. A comparison of publication and citation profiles of Full and Associate Professors at a School of Chemistry in Uruguay", *Scientometrics*, 69, 2: 287-313. <https://doi.org/10.1007/s11192-006-0154-5>
- Vessuri H., Guédon J. C. & Cetto A. M. (2014), "Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development", *Current Sociology*, 62, 5: 647-665. <https://doi.org/10.1177/0011392113512839>
- Veugelers R. & van Bouwel L. (2015), "The Effects of International Mobility on European Researchers: Comparing Intra-EU and U.S. Mobility", *Research in Higher Education*, 56, 4: 360-377. <https://doi.org/10.1007/s11162-014-9347-6>
- Viesti G. (2019), Un'analisi della mobilità geografica degli studenti universitari in Italia, *Rivista Economica del Mezzogiorno*, n.2
- Waaijer C. J. F., Teelken C., Wouters P. F. & van der Weijden I. C. M. (2018), "Competition in Science: Links between Publication Pressure, Grant Pressure and the Academic Job Market", *Higher Education Policy*, 31, 2: 225-243. <https://doi.org/10.1057/s41307-017-0051-y>
- Wager E., Fiack S., Graf C., Robinson A. & Rowlands I. (2009), "Science journal editors' views on publication ethics: results of an international survey", *Journal of Medical Ethics*, 35, 6: 348-353. <https://doi.org/10.1136/jme.2008.028324>
- Wagner C. S. & Leydesdorff L. (2005), "Network structure, self-organization, and the growth of international collaboration in science", *Research Policy*, 34, 10: 1608-1618. <https://doi.org/10.1016/j.respol.2005.08.002>
- Walker R. & Rocha da Silva P. (2015), "Emerging trends in peer review - a survey", *Frontiers in Neuroscience*, 9: 169. <https://doi.org/10.3389/fnins.2015.00169>
- Waltman L. (2016), "A review of the literature on citation impact indicators", *Journal of Informetrics*, 10, 2: 365-391. <https://doi.org/10.1016/j.joi.2016.02.007>
- Waltman L. & van Eck N. J. (2015), "Field-normalized citation impact indicators and the choice of an appropriate counting method", *Journal of Informetrics*, 9, 4: 872-894. <https://doi.org/10.1016/j.joi.2015.08.001>
- Waltman L., van Eck N. J., van Leeuwen T. N., Visser M. S. & van Raan A. F. J. (2011), "Towards a new crown indicator: Some theoretical considerations", *Journal of Informetrics*, 5, 1: 37-47. <https://doi.org/10.1016/j.joi.2010.08.001>

- Wang J. & Hicks D. (2013), "Detecting structural change in university research systems: A case study of british research policy", *Research Evaluation*, 22, 4: 258-268. <https://doi.org/10.1093/reseval/rvt016>
- Wang X. (2022), *Italian Research Assessment VQR: Framework, Achievements and Controversies*, in *ACM International Conference Proceeding Series*, pp. 261-267. <https://doi.org/10.1145/3543321.3543364>
- Weingart P. (2005), "Impact of bibliometrics upon the science system: Inadvertent consequences?", *Scientometrics*, 62, 1: 117-131. <https://doi.org/10.1007/s11192-005-0007-7>
- Welch A. R. (1997), "The peripatetic professor: the internationalisation of the academic profession", *Higher Education*, 34, 3: 323-345. <https://doi.org/10.1023/A:1003071806217>
- Welsh D. T. & Ordóñez L. D. (2014), "The dark side of consecutive high performance goals: Linking goal setting, depletion, and unethical behavior", *Organizational Behavior and Human Decision Processes*, 123, 2: 79-89. <https://doi.org/10.1016/J.OBHD.P.2013.07.006>
- Westerheijden D. F. (2007), *States and Europe and Quality of Higher Education*, in Westerheijden D.F., Stensaker B., Rosa M.J., eds., *Quality Assurance In Higher Education. Higher Education Dynamics*, vol 20, Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6012-0_3
- Wetzelhütter D. (2020), "Scale-sensitive response behavior!? Consequences of offering versus omitting a "don't know" option and/or a middle category", *Survey Practice*, 13, 1: 1-13. <https://doi.org/10.29115/SP-2020-0012>
- Whitley R. (2003), "Competition and pluralism in the public sciences: the impact of institutional frameworks on the organisation of academic science", *Research Policy*, 32, 6: 1015-1029. [https://doi.org/10.1016/S0048-7333\(02\)00112-9](https://doi.org/10.1016/S0048-7333(02)00112-9)
- Whitley R. (2007), *Changing Governance of the Public Sciences*, in Whitley, R., Gläser, J., eds. *The Changing Governance of the Sciences. Sociology of the Sciences Yearbook*, vol 26. Springer, Dordrecht, pp. 3-27. https://doi.org/10.1007/978-1-4020-6746-4_1
- Whitley R., Gläser J. & Engwall L., eds. (2010), *Reconfiguring Knowledge Production: Changing Authority Relationships in the Sciences and their Consequences for Intellectual Innovation*, Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199590193.001.0001>
- Wilson J., Allen L., Belfiore E., et al. (2015), *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. HEFCE. <https://doi.org/10.13140/RG.2.1.4929.1363>
- Woeleert P. (2015), "Governing Knowledge: The Formalization Dilemma in the Governance of the Public Sciences", *Minerva*, 53, 1: 1-19. <https://doi.org/10.1007/s11024-015-9266-5>
- Woeleert P. & McKenzie L. (2018), "Follow the money? How Australian universities replicate national performance-based funding mechanisms", *Research Evaluation*, 27, 3: 184-195. <https://doi.org/10.1093/reseval/rvy018>
- Wouters P. (1999), *The Citation Culture*, PhD thesis, University of Amsterdam. <http://garfield.library.upenn.edu/wouters/wouters.pdf>

- Wróblewska M. N., Balaban C., Derrick G. & Benneworth P. (2023), "The conflict of impact for early career researchers planning for a future in the academy", *Research Evaluation*, rvad024. <https://doi.org/10.1093/reseval/rvad024>
- Wuchty S., Jones B. F. & Uzzi B. (2007), "The Increasing Dominance of Teams in Production of Knowledge", *Science*, 316, 5827: 1036-1039. <https://doi.org/10.1126/science.1136099>
- Xu S., Mariani M. S., Lü L. & Medo M. (2020), "Unbiased evaluation of ranking metrics reveals consistent performance in science and technology citation data", *Journal of Informetrics*. 14, 1: 101005. <https://doi.org/10.1016/j.joi.2019.101005>
- Yin R. K. (2018), *Case Study Research and Applications, Design and Methods (Sixth Edition)*, SAGE Publishing.
- Za S. & Braccini A. M. (2017), "Tracing the roots of the organizational benefits of IT services", in: Za S., Drăgoicea M., Cavallari M., eds. *Exploring Services Science. IESS 2017. Lecture Notes in Business Information Processing*, vol 279. Springer, Cham , pp. 3-11. https://doi.org/10.1007/978-3-319-56925-3_1
- Za S., Minelli, E. A. & Resca A. (2021). *Publication trends in Organization Studies in Italy: a discussion about higher education policies and rankings*, in *ITAIS 2021 Proceedings*. 14. <https://aisel.aisnet.org/itaits2021/14>
- Zacharewicz T., Lepori B., Reale E. & Jonkers K. (2019), "Performance-based research funding in EU Member States – a comparative assessment", *Science and Public Policy*, 46, 1: 105-115. <https://doi.org/10.1093/scipol/scy041>
- Zagaria C. (2007), *Processo all'università. Cronache dagli atenei italiani tra inefficienze e malcostume*, Edizioni Dedalo.
- Zahedi Z., Costas R. & Wouters P. (2014), "How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications", *Scientometrics*, 101, 2: 1491-1513. <https://doi.org/10.1007/s11192-014-1264-0>
- Zhang Y. & Yu Q. (2020), "What is the best article publishing strategy for early career scientists?", *Scientometrics*, 122, 397-408. <https://doi.org/10.1007/s11192-019-03297-4>
- Zhao D. & Strotmann A. (2015), *Analysis and Visualization of Citation Networks*, Springer Cham. <https://doi.org/10.1007/978-3-031-02291-3>
- Zinilli A. (2016), "Competitive project funding and dynamic complex networks: evidence from Projects of National Interest (PRIN)", *Scientometrics*, 108, 2: 633-652. <https://doi.org/10.1007/s11192-016-1976-4>
- Zinilli A., Pierucci E. & Reale E. (2023), "Organizational factors affecting higher education collaboration networks: evidence from Europe", *Higher Education*, 88: 119-160. <https://doi.org/10.1007/s10734-023-01109-6>
- Zinilli A., Spinello A. O. & Reale E. (2020), "Finanziamento e valutazione della ricerca pubblica: l'approccio performance-based", *Welfare e Ergonomia*, 1/2019: 17-28. <https://doi.org/10.3280/WE2019-001003>
- Zinov'yeva N. & Bagues M. (2015), "The role of connections in academic promotions", *American Economic Journal: Applied Economics*, 7, 2: 264-292. <https://doi.org/10.1257/app.20120337>
- Zuccala A. (2012), "Quality and influence in literary work: Evaluating the "educated imagination""", *Research Evaluation*, 21, 3: 229-241. <https://doi.org/10.1093/reseval/rvs017>

Gli Autori

Emanuela Reale, sociologa, è stata direttrice dell'Istituto di Ricerca sulla Crescita Economica Sostenibile – CNR-IRCrES fino al 2023 e attualmente è Associato senior presso lo stesso istituto. È stata principal investigator e coordinatrice di numerosi progetti internazionali e nazionali, tra cui il più recente RISIS, l'infrastruttura europea per la ricerca scientifica, tecnologica e l'innovazione, il PRIN2017 sugli effetti della valutazione sul lavoro accademico e il progetto PNRR-IR “FOSSR”. È Presidente del Forum Europeo di Studi sulle Politiche di Ricerca e Innovazione-Eu-SPRI e co-Editor di Research Evaluation.

Giovanni Abramo, dopo circa quarant'anni al Consiglio Nazionale delle Ricerche (CNR), è passato nel 2023 all'Universitas Mercatorum come Professore di Management dell'Innovazione e Entrepreneurship. I suoi interessi scientifici concernono la valutazione della performance di ricerca e la scientometria valutativa. Giovanni ha co-fondato il Laboratorio CNR di Studi sulla Valutazione della Ricerca nel 2004 e Research Value, una società *spin-off* CNR, nel 2006. Nel 2023 è stato eletto Presidente dell'International Society for Scientometrics and Informetrics (ISSI).

Ciriaco Andrea D'Angelo è professore associato presso il Dipartimento di Ingegneria dell'Impresa dell'Università di Roma “Tor Vergata”. È stato visiting research fellow presso il Centre for Science and Technology Studies (CWTS) alla Leiden University, in Olanda. Co-fondatore del Laboratorio di Studi sulla Valutazione della Ricerca e di Research Value, uno *spin-off* del Consiglio Nazionale delle Ricerche, i suoi interessi scientifici riguardano la valutazione della ricerca e la realizzazione di sistemi di supporto alle decisioni di *policy maker* e *research manager*.

Valentina Carazzolo è assegnista di ricerca presso il CNR-IRCrES (sede di Roma). Ha conseguito il dottorato in Scienze Sociali Applicate presso l'Università di Roma La Sapienza. I suoi interessi di ricerca riguardano lo studio delle determinanti della *research agenda* dei ricercatori, le politiche della valutazione della ricerca, e le attività di terza missione e *public engagement* delle università.

Flavia Di Costa è ricercatrice a tempo determinato in *tenure-track* (RTT) presso l'Università Telematica Universitas Mercatorum. Ha conseguito il Dottorato di ricerca in Ingegneria Economico-Gestionale presso l'Università di Roma "Tor Vergata" ed è stata assegnista post-doc presso l'Università di Roma "La Sapienza", l'Istituto di Analisi dei Sistemi e Informatica (IASI) del CNR e l'Università di Roma "Tor Vergata". I suoi principali interessi di ricerca sono: Scientometria; Bibliometria; Valutazione della Ricerca; Politiche e Management della Ricerca; Scienza, tecnologia e innovazione.

Serena Fabrizio è ricercatrice presso il CNR-IRCrES, con cui collabora dal 2015 su progetti di ricerca europei e nazionali. In particolare nel team di ricerca del PRIN2017, e nel progetto RISIS dove ha gestito le attività di formazione, comunicazione e *community building*. I principali interessi di ricerca riguardano le politiche per l'università e la ricerca, l'analisi dei finanziamenti pubblici per R&S, lo studio (e valutazione) dell'impatto della ricerca in SSH e in particolare delle Infrastrutture di Ricerca. Attualmente è coinvolta nel progetto PNRR-FOSSR, su attività di *community building* e formazione.

Ugo Finardi è primo ricercatore presso il CNR-IRCrES (sede di Torino). I suoi principali interessi di ricerca sono legati agli studi sull'economia e il management dell'innovazione e delle nuove tecnologie per le imprese, a quelli sulle politiche e sulla valutazione dell'università e della ricerca, e a quelli sulle ricadute socioeconomiche della cybersicurezza. Nel corso del tempo i suoi lavori di ricerca hanno adottato sia metodologie qualitative di vario genere sia strumenti di analisi quantitativi.

Leonardo Grilli è professore ordinario di Statistica presso il Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti" (DiSIA) dell'Università di Firenze. È stato Presidente del corso di laurea magistrale in Statistica e *data science* ed è attualmente membro eletto del Consiglio Direttivo della Società Italiana di Statistica. È titolare di corsi di metodologia statistica e modelli statistici in corsi di laurea, dottorati di ricerca e corsi di specializzazione. Gli interessi di ricerca preminentri riguardano aspetti metodologici relativi a modelli di regressione multilivello e metodi per l'inferenza causale e la valutazione di efficacia, con applicazioni in ambito economico, sociale, demografico e medico.

Eliana Alessandra Minelli è professore associata presso la LIUC-Università Cattaneo, dove è titolare di corsi a livello Master e Dottorato. Ha conseguito il dottorato in Sociologia presso l'Università di Gent (Belgio). I suoi interessi di ricerca riguardano il cambiamento, il comportamento organizzativo e la gestione delle risorse umane, i sistemi di *governance* nell'Higher Education. In particolare ha condotto studi sull'impatto della globalizzazione e del cambiamento istituzionale sulle organizzazioni, con particolare riferimento al contesto dell'università. La sua esperienza accademica include, oltre all'insegnamento, anche la partecipazione a progetti di ricerca nazionali e internazionali (PRIN e Interreg). I suoi articoli sono pubblicati su riviste scientifiche internazionali.

Lucio Morettini è ricercatore presso il CNR-IRCrES, sede di Roma. La sua attività di ricerca si concentra principalmente sulle politiche universitarie e sulla valutazione della ricerca, in particolare sulle dinamiche di carriera dei lavoratori ad alta formazione e sull'impatto delle politiche per la ricerca e l'alta formazione sull'innovazione nelle imprese. Attualmente è referente dell'unità CNR – IRCRES del progetto PRIN – PNRR “POTION” sulla valutazione di impatto delle politiche di transizione europee applicate nelle regioni italiane.

Andrea Orazio Spinello è ricercatore presso il CNR-IRCrES, sede di Roma. La sua attività di ricerca si concentra principalmente sulle politiche e la valutazione dell'università e della ricerca, con un focus sugli investimenti pubblici in R&S e i cambiamenti nella pratica scientifica. È tra i responsabili del dataset EFIL, incluso nell'infrastruttura di ricerca europea RISIS per gli studi su ricerca e innovazione. Attualmente ricopre il ruolo di WP leader per le attività di formazione e *community building* del progetto PNRR-IR “FOSSR” ed è responsabile di unità per il CNR del progetto PRIN2022 “PLACES” sul *community engagement* delle università.

Maria Rucsandra Stan è assegnista di ricerca presso l'Università di Milano-Bicocca, si occupa di temi di innovazione ed *entrepreneurship*. Ha conseguito il dottorato in Management, Accounting e Finance presso la LIUC-Università Cattaneo. È stata visiting researcher presso il Dipartimento di Education della University of Oxford, partecipando al progetto “Research on Research: the research function and mission of higher education”. I suoi principali ambiti di ricerca sono gli studi sul ruolo del linguaggio nelle iniziative imprenditoriali, l'impatto della valutazione della ricerca sulla produzione di conoscenza, e i risvolti dell'intelligenza artificiale sulla performance organizzativa.

Antonio Zinilli è ricercatore presso il CNR-IRCrES, sede di Roma, ed è membro del comitato scientifico del dottorato di ricerca in Psicologia Sociale dell'Università Sapienza di Roma. È coordinatore della Scuola in “Data Science: tools and methods for analysing complex Science, Technology and Innovation (STI) systems”. La sua ricerca, basata su un approccio interdisciplinare, si concentra sull'analisi e la modellizzazione dei processi di diffusione della conoscenza e innovazione. Ha pubblicato e svolto attività di *referee* in diverse riviste internazionali ed è responsabile dell'unità CNR per l'infrastruttura europea RISIS e per il progetto PRIN “NEUTRALISE”, oltre a essere WP leader nel progetto PNRR-IR “FOSSR”.

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.



Management, finanza,
marketing, operations, HR
Psicologia e psicoterapia:
teorie e tecniche
Didattica, scienze
della formazione
Economia,
economia aziendale
Sociologia
Antropologia
Comunicazione e media
Medicina, sanità



Architettura, design,
territorio
Informatica, ingegneria
Scienze
Filosofia, letteratura,
linguistica, storia
Politica, diritto
Psicologia, benessere,
autoaiuto
Efficacia personale
Politiche
e servizi sociali

FrancoAngeli

La passione per le conoscenze

Copyright © 2025 by FrancoAngeli s.r.l., Milano, Italy. ISBN 9788835179092

FrancoAngeli

a strong international commitment

Our rich catalogue of publications includes hundreds of English-language monographs, as well as many journals that are published, partially or in whole, in English.

The **FrancoAngeli**, **FrancoAngeli Journals** and **FrancoAngeli Series** websites now offer a completely dual language interface, in Italian and English.

Since 2006, we have been making our content available in digital format, as one of the first partners and contributors to the **Torrossa** platform for the distribution of digital content to Italian and foreign academic institutions. **Torrossa** is a pan-European platform which currently provides access to nearly 400,000 e-books and more than 1,000 e-journals in many languages from academic publishers in Italy and Spain, and, more recently, French, German, Swiss, Belgian, Dutch, and English publishers. It regularly serves more than 3,000 libraries worldwide.

Ensuring international visibility and discoverability for our authors is of crucial importance to us.

FrancoAngeli



Quali cambiamenti è in grado di produrre la valutazione esterna sul modo di fare ricerca e sulle strategie individuali dei ricercatori e docenti universitari? Fino a che punto è possibile cogliere gli effetti della valutazione a livello individuale, utilizzando approcci misti che combinano dati quantitativi bibliometrici e non bibliometrici e dati qualitativi derivanti da interviste e storie di vita? Queste domande hanno guidato lo studio di un Progetto PRIN (2017NKWYFC), che ha esaminato gli effetti prodotti dalla VQR (Valutazione della Qualità della Ricerca) e dall'ASN (Abilitazione Scientifica Nazionale) sugli accademici italiani, i cui risultati sono contenuti in questo volume.

I diversi capitoli propongono al lettore evidenze su come la valutazione sia stata e continui a essere strumento per la politica dell'università in Italia per indirizzare i comportamenti verso obiettivi di qualità costruiti intorno a standard omogenei, spesso non adatti a conoscere il valore della ricerca scientifica. I risultati mostrano quanto diversi siano gli effetti nei vari ambiti disciplinari, come gli effetti stessi si producano in modo più accentuato sui giovani ricercatori e ricercatrici, in genere più vulnerabili ed esposti a pressioni esterne, incidendo su produttività, qualità scientifica, modi di disseminazione e agenda di ricerca.

Il quadro che emerge presenta una situazione eterogenea, molto legata all'impostazione esclusivamente sommativa della valutazione, la quale mostra limiti sempre più evidenti che meriterebbero di essere affrontati non con modifiche marginali degli strumenti esistenti ma con una trasformazione sostanziale dell'impostazione da dare alla valutazione che valorizzi gli aspetti di apprendimento, sia da parte del valutatore che da parte dei valutati, su come migliorare qualità, responsabilità e apertura della ricerca accademica e delle università.

Emanuela Reale, sociologa, è stata direttrice dell'Istituto di Ricerca sulla Crescita Economica Sostenibile-IRCrES CNR fino al 2023 e attualmente è Associato senior presso l'IRCrES del CNR. È stata PI e coordinatrice di numerosi progetti internazionali e nazionali, tra cui il più recente RISIS, l'infrastruttura europea per la ricerca scientifica, tecnologica e l'innovazione, il PRIN2017 sugli effetti della valutazione sul lavoro accademico e il progetto PNRR-IR “FOSSR”. È Presidente del Forum Europeo di Studi sulle Politiche di Ricerca e Innovazione-Eu-SPRI e co-Editor di *Research Evaluation*. Ha pubblicato oltre 150 lavori in sedi editoriali nazionali e internazionali ed è referee per numerose riviste internazionali che si occupano di politica della ricerca, valutazione della ricerca e studi sull'istruzione superiore.