

A cura di
**Antonio Fasanella
e Fabrizio Martire**

**Valutazione
della ricerca e ricerca
sulla valutazione**
Riflessioni, analisi e proposte
per la VQR



FrancoAngeli

OPEN  ACCESS

IL RICCIO E LA VOLPE

Studi, ricerche e percorsi di sociologia

Il riccio e la volpe
Studi, ricerche e percorsi di sociologia

Collana diretta da Enzo Campelli

Comitato scientifico: Maria Stella Agnoli, Maria Carmela Agodi, Maurizio Bonolis, Antonio Fasanella, Giuseppe Giampaglia, Renato Grimaldi, Carmelo Lombardo, Alberto Marradi, Sergio Mauceri, Luigi Muzzetto, Ambrogio Santambrogio

Questa collana ospita, con la più pronunciata apertura tematica e nel pluralismo consapevole delle interpretazioni, indagini empiriche e riflessioni teoriche nell'ambito della sociologia generale.

La sua instestazione richiama un verso di Archiloco che, in uno dei frammenti sopravvissuti, afferma lapidariamente, e in realtà piuttosto oscuramente, che "la volpe sa molte cose, ma il riccio ne sa una grande". Isaiah Berlin, interpretando questa presunta differenza di saperi, scrive, in un saggio degli anni '50, che "esiste un grande divario tra coloro, da una parte, che riferiscono tutto a una visione centrale, a un sistema più o meno coerente e articolato, con regole che li guidano a capire, a pensare e a sentire – un principio ispiratore, unico e universale, il solo che può dare significato a tutto ciò che essi sono e dicono –, e coloro, dall'altra parte, che perseguono molti fini, spesso disgiunti e contraddittori, magari collegati soltanto genericamente, de facto, per qualche ragione psicologica o fisiologica, non unificati da un principio morale ed estetico".

In anni di mutamento sociale e culturale imprevedibilmente accelerato, di "sconfinamenti" e di ibridazioni, questa collana punta dunque a cogliere e documentare le intersezioni e le contrapposizioni, nelle dinamiche sociali, fra l'unitario e il molteplice, il disordinato e il sistemico, il conforme e l'eterogeneo, il caso e la regola: *il riccio e la volpe*, per l'appunto.

Abbandonata la pretesa inattuale di ogni sintesi semplice, difficilmente la sociologia potrebbe oggi sottrarsi a questo lavoro paziente di ricostruzione.

La molteplicità delle tematiche affrontate e la pluralità delle prospettive trovano, peraltro, una precisa composizione unitaria nella ferma e rigorosa opzione disciplinare che ispira la collana stessa, e cioè nella puntigliosa rivendicazione della sociologia come disciplina costantemente attenta all'integrazione tra teoria e ricerca, al rigore logico-metodologico delle procedure, al rispetto della fondamentale esigenza di pubblicità e controllabilità dell'indagine scientifica.

Sulla base di questi convincimenti di natura teorico-metodologica, e nel costante richiamo alla responsabilità sociale di ogni disciplina scientifica, la collana si propone di fornire a studiosi, a studenti e a operatori strumenti qualificati di riflessione e di intervento.



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:

http://www.francoangeli.it/come_publicare/publicare_19.asp

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.

A cura di
**Antonio Fasanella
e Fabrizio Martire**

**Valutazione
della ricerca e ricerca
sulla valutazione**
Riflessioni, analisi e proposte
per la VQR

FrancoAngeli

Questo volume è stato pubblicato con il contributo del Dipartimento di Comunicazione e Ricerca Sociale di Sapienza Università di Roma.

Copyright © 2020 by FrancoAngeli s.r.l., Milano, Italy.

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore ed è pubblicata in versione digitale con licenza *Creative Commons Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale* (CC-BY-NC-ND 4.0)

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

Introduzione , di <i>Antonio Fasanella</i> e <i>Fabrizio Martire</i>	pag.	7
1. Esperienze di valutazione della ricerca scientifica , di <i>Marco Palmieri</i>	»	13
1.1. Introduzione	»	13
1.2. Le prime esperienze di valutazione in Italia	»	17
1.3. La valutazione triennale della ricerca, la Vtr 2001-2003	»	21
1.4. La valutazione della qualità della ricerca, la Vqr 2004-2010	»	25
1.5. La valutazione della qualità della ricerca, la Vqr 2011-2014	»	28
1.6. L'Anvur	»	35
1.7. Una comparazione dei principali sistemi europei di valutazione della qualità della ricerca	»	37
2. La definizione del concetto di qualità della ricerca , di <i>Antonio Fasanella</i>	»	48
2.1. Introduzione	»	48
2.2. Determinatezza e uniformità d'uso del concetto di qualità	»	49
2.3. Un tentativo di chiarificazione concettuale	»	58
3. I Gev e i revisori , di <i>Lorenzo Barbanera</i>	»	65
3.1. I Gruppi di Esperti Valutatori	»	69

3.1.1. La composizione dei gruppi	pag.	69
3.1.2. L'importanza dell'eterogeneità dei gruppi	»	71
3.1.3. L'assegnazione dei prodotti agli Ev	»	74
3.1.4. Strategie di <i>accountability</i>	»	76
3.2. I revisori	»	79
3.2.1. Il processo di selezione	»	79
3.2.2. Il <i>matching</i> prodotto-revisore	»	83
3.2.3. L' <i>accountability</i> recitativa	»	85
3.3. Conclusioni	»	91
4. La scheda di valutazione dei prodotti scientifici, di <i>Federica Floridi</i>	»	93
4.1. Introduzione	»	93
4.2. La struttura della scheda Vqr 2011-2014	»	94
4.3. Gli oggetti multipli e le conseguenze della sotto-determinazione semantica	»	98
4.4. La manipolazione <i>ex post</i> delle classi di merito: una questione di qualità	»	102
4.5. I giudizi obbligatori: un'occasione mancata	»	108
5. Formulazione e sintesi dei giudizi, di <i>Federica Fusillo</i>	»	112
5.1. La classificazione della qualità della ricerca	»	112
5.2. La sintesi del giudizio del singolo revisore	»	117
5.3. L'attribuzione della classe di merito finale	»	129
5.4. Conclusioni	»	137
6. Fare ricerca sulla Vqr, di <i>Fabrizio Martire</i>	»	139
Riferimenti bibliografici	»	149
Curatori e autori	»	157

1. Esperienze di valutazione della ricerca scientifica

di Marco Palmieri

1.1. Introduzione

In Italia l'esigenza di valutare le università comincia ad essere avvertita alla fine degli anni Ottanta a seguito del riconoscimento della loro autonomia nei confronti degli organi centrali di governo. La legge n. 168 del 1989, che ha sancito l'autonomia didattica, scientifica, organizzativa delle università italiane, già prevista dalla Costituzione, ha messo in moto un processo di riforma del sistema universitario proseguito nei decenni successivi.

Punto di svolta è la legge n. 180 del 2008 in cui si legge che, «a decorrere dall'anno 2009, al fine di promuovere e sostenere l'incremento qualitativo delle attività delle università statali e di migliorare l'efficacia e l'efficienza nell'utilizzo delle risorse, (...) una quota non inferiore al 7 per cento del fondo di finanziamento ordinario con progressivi incrementi negli anni successivi, è ripartita prendendo in considerazione la qualità dell'offerta formativa e i risultati dei processi formativi e la qualità della ricerca scientifica». La legge n. 69 del 2013 è poi intervenuta sulle modalità di attribuzione della quota premiale del fondo di finanziamento ordinario all'università negli anni a venire.

Questo intervento legislativo cambia nel profondo il meccanismo di finanziamento statale: parte dell'ammontare dei fondi pubblici destinati all'università viene legato alle prestazioni rese e al valore prodotto dai singoli atenei, che gestiscono i finanziamenti ricevuti in piena autonomia. Nel nuovo sistema di redistribuzione dei fondi pubblici diventa centrale la dimensione della *performance* degli atenei da sottoporre a valutazione (Reale, 2013a).

La didattica, più in generale la formazione, e la ricerca sono tra gli aspetti che entrano nel processo di valutazione. Dall'analisi dei decreti ministeriali del Ministero dell'Istruzione, dell'Università e della Ricerca (Miur) emerge che: la quota premiale del fondo di finanziamento ordinario (Ffo) cresce costantemente dal 2009 al 2019; una parte sempre più consistente del Ffo è redistribuito tenendo conto degli esiti degli esercizi di valutazione della qualità della ricerca prodotta dal sistema universitario nazionale¹.

Nel 2009 la quota premiale è stabilita al 7% del Ffo, il 50% della quale è assegnata secondo i risultati della Vtr 2001-2003². Nel 2010 e nel 2011 la quota premiale è pari al 12% del Ffo, e nel 2012 tale quota sale al 13%; in generale nel triennio il 12% del fondo premiale è redistribuito tenendo conto della Vtr 2001-2003. Nel 2013 la quota premiale è al 13,5% del Ffo, il 66% della quale è assegnata tenendo conto degli esiti della Vqr 2004-2010. Nel 2014 il fondo premiale sale al 18% del Ffo, il 90% del quale è ripartito in base agli esiti della Vqr 2004-2010. La quota premiale per l'anno 2015, e anche per il 2016, è il 20% del Ffo; l'85% di tale importo è suddiviso secondo gli esiti della Vqr 2011-2014. Nel 2017 il Miur ha destinato ben il 22% del Ffo al fondo premiale, e l'85% di questo è stato erogato in funzione dell'ultima Vqr. La quota premiale sale al 24% nel 2018 e al 26% nel 2019; il 60% di questa è redistribuita secondo i risultati della Vqr 2011-2014.

In questo mutamento di prospettiva ha giocato un ruolo importante la diffusione delle idee del *new public management*, una corrente di pensiero che ha orientato le scelte del governo nella direzione di importare nel settore pubblico metodi di *management* aziendale (Reale e Pennisi, 2012). Secondo questo approccio, la valutazione dei risultati è uno strumento di regolazione del funzionamento dei sistemi pubblici, perché supporta un programma di incentivazione che ha il fine di orientare i comportamenti dei soggetti valutati. Le università, in

¹ Vista la delicatezza dell'argomento, chi scrive consiglia l'avvio di un programma di ricerca valutativa che analizzi l'impatto della redistribuzione premiale dei fondi sull'operato degli atenei italiani. Al riguardo vedi il capitolo 6.

² La Vtr 2001-2003 è stata la prima esperienza di valutazione della qualità della ricerca nelle università italiane (sulla Vtr 2001-2003 vedi il par. 1.3), poi seguita da altri due esercizi di valutazione, la Vqr 2004-2010 (sulla Vqr 2004-2010 vedi il par. 1.4) e la Vqr 2011-2014, che è per l'appunto oggetto di questo libro.

quanto istituzioni del settore pubblico, sono quindi chiamate a rendicontare le loro attività in vista della ripartizione degli stanziamenti statali.

Non solo. La legge n. 537 del 1993 istituisce l'autonomia finanziaria degli atenei, grazie alla quale cadono i vincoli di destinazione contabile delle università nei confronti dello Stato finanziatore: i fondi pubblici non sono più definiti dal Ministero in voci di spesa univoche, che ne limitano la destinazione d'uso, ma entrano a far parte di un unico *budget*, il fondo di finanziamento ordinario; ciascun ateneo gestisce autonomamente la quota parte che gli viene riconosciuta.

L'autonomia di cui godono le università le obbliga all'esercizio di responsabilità nella gestione delle risorse pubbliche assegnate. Ciò implica la necessità di rendere conto pubblicamente del loro operato e dei risultati conseguiti per favorire il controllo democratico sulle modalità di gestione della cosa pubblica (*accountability*). In questo senso, la valutazione è strettamente connessa alle responsabilità di cui sono investite le università in seguito al riconoscimento della loro autonomia finanziaria. Valutazione come controllo di affidabilità e come mezzo per favorire un esercizio responsabile dell'autonomia.

La valutazione si afferma come corollario indispensabile del riconoscimento alle università di autonomia di tipo sostanziale e non meramente procedurale. Essa deve, infatti, nelle intenzioni del legislatore, consentire alle università di acquisire consapevolezza del proprio ruolo e delle proprie funzioni, favorire l'auto-riflessione, contribuire a rendere evidenti le criticità che necessitano d'intervento, e consentire di conoscere i risultati prodotti e il loro utilizzo, e chi concorre alla loro produzione (Reale, 2013b, p. 148-149).

Si assiste al passaggio da una forma di regolamentazione tradizionale, in cui lo Stato stabilisce *ex ante* l'allocazione delle risorse ed è ultimamente responsabile della loro gestione, a una in cui la responsabilità è trasferita agli atenei, che devono però dimostrare di saper gestire l'autonomia sottoponendosi a una valutazione *ex post* dei loro risultati. Lo Stato abbandona progressivamente la figura di finanziatore che controlla il sistema agendo a monte sulle quote di risorse assegnate ai diversi atenei, per assumere progressivamente il ruolo di valutatore che a valle attribuisce le risorse premiali in base al buono o cattivo uso che fin lì ne è stato fatto.

Da questa prospettiva, la valutazione delle università serve a dare

prova della loro affidabilità a tutti coloro che, a vario titolo, sono portatori di interessi nei confronti delle istituzioni accademiche: non solo gli organi centrali di governo ma anche le imprese del territorio e, più in generale, la cittadinanza (Ribolzi, 2013).

Il modello di *accountability* configuratosi nel contesto italiano ha sollevato però le resistenze di molti docenti e ricercatori, per via di un sistema di valutazione che spingerebbe a conformare le attività di ricerca e didattica delle università a criteri di qualità imposti dall'esterno.

In termini più generali, usi della valutazione di tal fatta comportano la sostituzione del sistema di regolazione specifico al campo della scienza con un nuovo sistema di normazione eteronoma: il potere economico e politico verrebbe così a sottrarre spazio ad altre forme di potere più consolidate nell'accademia, quali il potere istituzionale e istituzionalizzato (occupazione di posti di potere) e quello specifico della scienza (fondato sul prestigio e l'autorevolezza). La qual cosa è connessa a una perdita di autonomia del campo scientifico (Chessa e Vargiu, 2011, p. 23).

La valutazione è stata così percepita dal mondo accademico non come una procedura per incrementare la responsabilità istituzionale dall'università italiana, ma come uno strumento per controllare la conformità degli atenei agli standard prestazionali richiesti dalle autorità governative (Palumbo e Pennisi, 2011). Parte della comunità accademica è dell'opinione che in questo modello di valutazione l'innalzamento della qualità delle università è una finalità perseguita solo indirettamente, tramite un sistema di allocazione delle risorse che mira a incoraggiare i comportamenti conformi rispetto agli standard. Ciò ha favorito condotte di tipo opportunistico che puntano a garantire il rispetto formale, e non sostanziale, dei requisiti richiesti. Se però la conformità innescata è solo formale, la valutazione perde la sua più importante funzione, ossia favorire il miglioramento di didattica e ricerca mediante un processo di apprendimento (Rebora, 2013).

Quando scopo della valutazione è controllare il settore o verificarne la conformità alle politiche del governo o assicurare la rendicontazione (di solito in riferimento ai contribuenti), e la metodologia comprende qualche forma di ispezione, sebbene nella forma apparentemente benevola della *peer review*, allora non c'è alcun segnale di democrazia. Tutto il sistema appare autocratico e all'interno delle università è percepito come autocratico (Harvey, 2008, p. 9).

Secondo Valentini (2013), piuttosto che un'attività burocratica di rendicontazione a un organo di controllo esterno, la valutazione può costituire un'insostituibile occasione di riflessione, un momento di condivisione dei risultati ottenuti che può favorire un processo di apprendimento da parte della comunità accademica. Da questa prospettiva gli esiti del processo valutativo rappresentano non tanto un supporto al sistema premiale di allocazione delle risorse quanto un bacino informativo cui gli atenei possono attingere per definire le loro politiche nei campi della didattica e della ricerca. La logica del miglioramento porta cioè ad attribuire alla valutazione la funzione di rendere espliciti operato e risultati dei singoli atenei per promuovere un confronto dialettico tra le diverse esperienze del mondo accademico.

Cipriani (2013) osserva che un sistema orientato al miglioramento della qualità degli atenei più che alla ricerca di prove della loro affidabilità favorisce un meccanismo di autoregolazione da parte della comunità accademica. Ciò può accadere se la valutazione non è percepita dagli attori come attività inquisitoria cui reagire difensivamente, ma come pratica riflessiva che spinge alla ricerca di un cambiamento. La collaborazione della comunità accademica alla costruzione dell'edificio valutativo diventa così la chiave per la diffusione di una cultura della valutazione capace di favorire un processo di responsabilizzazione di tutti gli attori coinvolti.

Per costruire una simile rappresentazione della valutazione è opportuno coinvolgere docenti e ricercatori nel processo di definizione dei criteri con cui valutare la loro attività, riflettendo sulle dimensioni della qualità della ricerca per esplicitare, formalizzare e proporre pratiche di valutazione adeguate ai diversi contesti scientifico-disciplinari.

Proprio questo è lo spirito che ha guidato le riflessioni raccolte in questo volume, dedicato all'ultimo esercizio di valutazione della qualità della ricerca, la Vqr 2011-2014.

1.2. Le prime esperienze di valutazione in Italia

La valutazione entra nel mondo accademico italiano negli anni Novanta, in seguito all'approvazione della legge n. 537 del 1993, successivamente riformata dalla legge n. 370 del 1999. Le prime novità importanti riguardano l'istituzione dei Nuclei di valutazione interna e

dell'Osservatorio per la valutazione del sistema universitario. I primi sono organismi interni agli atenei «con il compito di verificare, mediante analisi comparativa dei costi e dei rendimenti, la corretta gestione delle risorse pubbliche, la produttività della ricerca e della didattica, nonché l'imparzialità ed il buon andamento dell'azione amministrativa [...] La relazione annuale dei Nuclei di valutazione interna è trasmessa al Ministero dell'Università e della ricerca scientifica e tecnologica anche ai fini della successiva assegnazione delle risorse» (legge n. 537 del 1993).

La legge affida ai Nuclei il ruolo delicato di valutatori interni all'università, con il compito di valutare la produttività dell'attività degli atenei. Il Nucleo è lo strumento di valutazione del sistema universitario italiano, che verifica la coerenza tra gli obiettivi programmati, le risorse impiegate, e i risultati raggiunti da ciascun ateneo. Al Nucleo spetta dunque il compito oneroso di valutare l'intera gestione dell'ateneo (didattica, amministrativa, di ricerca), ivi compresa l'analisi e la valutazione della direzione politica degli atenei.

Questa innovazione legislativa fatica però ad andare a sistema. In un documento del 1997 che riporta le conclusioni di un lavoro di monitoraggio dell'attività dei Nuclei a quattro anni dalla loro istituzione (Osservatorio, 1997, Doc 5/97) si legge che: in diversi atenei la nomina dei membri dei Nuclei di valutazione interna risponde solo a un obbligo formale, poiché i Nuclei non redigono la relazione annuale prevista dalla legge; la durata dell'incarico dei membri, e il relativo compenso, non sono adeguati all'impegno richiesto per svolgere un compito così oneroso; i Nuclei sono costituiti esclusivamente da membri interni all'ateneo, scelti dal rettore tra i docenti interni (a scapito di esperti esterni) senza che la loro autonomia sia adeguatamente tutelata.

L'altra importante novità della legge n. 537 del 1993 è la nascita di un Osservatorio permanente, istituito poi nel 1996. L'Osservatorio è un organo tecnico con il compito «di valutare i risultati relativi all'efficienza e alla produttività delle attività di ricerca e di formazione; di valutare i piani di sviluppo e riequilibrio del sistema universitario». Il suo compito principale è redigere una relazione annuale sullo stato della valutazione del sistema universitario italiano, elaborata sulla base delle relazioni preparate annualmente dai Nuclei di valutazione interna.

Con l'Osservatorio si passa da un sistema di valutazione costruito come sommatoria di soggetti decentrati (i Nuclei), più o meno omologati e tenuti in rete da basi di dati confrontabili, a un sistema duale, che combina l'attività di valutazione delle singole università con quella del sistema universitario nel suo complesso (Rizzi e Silvestri, 2002, p. 6).

Il neonato Osservatorio non sostituisce né limita il lavoro dei Nuclei di valutazione, anzi ne rafforza l'autonomia e l'indipendenza dagli atenei, essendo un organo tecnico composto da studiosi autorevoli, che raccoglie ed elabora il lavoro dei singoli Nuclei di valutazione e relaziona al Miur sullo stato della valutazione dell'università italiana. Si assiste così alla nascita del primo organismo nazionale di valutazione del sistema universitario, cui sono assegnati anche numerosi altri compiti: la valutazione dei piani triennali di sviluppo, i pareri sulle nuove università e sui nuovi dottorati di ricerca, la valutazione sulla riforma dell'autonomia didattica, ecc. Inoltre, l'Osservatorio suggerisce al Ministero come ripartire, o meglio riequilibrare, le quote del Ffo da assegnare agli atenei secondo il criterio delle pari opportunità da garantire a tutti gli atenei, con particolare attenzione a quelli che operano in contesti con condizioni socioeconomiche svantaggiate (Matarazzo, 2018).

Il ruolo e i compiti dei Nuclei di valutazione interna e dell'organismo centrale di valutazione del sistema universitario sono stati poi ridefiniti dalla legge n. 370 del 1999 che, facendo tesoro dell'esperienza di sei anni, cerca di valorizzare i punti di forza e correggere le inevitabili criticità. In primo luogo, vengono confermati i compiti già assegnati ai Nuclei di valutazione dalla precedente riforma (redazione della relazione annuale di accompagnamento al consuntivo di ateneo e valutazione di congruità tra gli obiettivi dichiarati e i mezzi indicati nell'ambito della programmazione triennale), conferendo loro anche altre funzioni che prima erano di competenza dell'Osservatorio (rilevazione delle opinioni degli studenti e redazione del rapporto sulla valutazione della didattica, valutazione dei requisiti per l'attivazione dei corsi di dottorato, valutazione della riforma degli ordinamenti didattici). Inoltre, viene rafforzata l'autonomia dei Nuclei rispetto agli atenei che sono chiamati a valutare, i quali da quel momento in poi sono obbligati a fornire loro le informazioni richieste per la valutazione dell'ateneo. Sono previste anche sanzioni (nel meccanismo di attribuzione dei fondi) per le università che non mettono il proprio Nucleo di

valutazione interna in condizione di lavorare con professionalità e imparzialità.

Rimane però un limite sostanziale all'azione del Nucleo: l'impossibilità di imporre all'ateneo modifiche ai progetti sulla base dei risultati dell'attività valutativa; solitamente il Nucleo non viene coinvolto nemmeno nella stesura iniziale, neanche per dare valutazioni informali. Al contrario, far partecipare il Nucleo alla scrittura del progetto sarebbe il «modo di procedere più vantaggioso sia per l'ateneo, che aumenterebbe le probabilità di successo dei propri progetti poiché in essi sarebbero già contenute le indicazioni del Nucleo interno, sia per il Nucleo che riuscirebbe a svolgere un positivo ruolo di valutazione senza scontrarsi con gli organi dell'ateneo, sia per il Comitato e il Ministero che riceverebbero progetti migliori» (Rizzi e Silvestri, 2002, p. 14).

La legge n. 370 del 1999 prevede la sostituzione dell'Osservatorio per la valutazione del sistema universitario con il Comitato nazionale per la valutazione del sistema universitario (Cnvsu), cui sono affidati i compiti fino a quel momento svolti dall'Osservatorio. La novità principale sta nel rapporto più stretto che il Comitato è tenuto a instaurare con i Nuclei di valutazione, fissandone gli obiettivi di lavoro e le informazioni che devono comunicare annualmente al Comitato, e preparando un programma annuale di valutazione esterna delle università. Si assiste, dunque, al rafforzamento dell'architettura piramidale dell'intero sistema di valutazione, al cui apice c'è il Cnvsu, che dialoga direttamente con il Ministero e che è il punto di riferimento di tutti i Nuclei di valutazione interna. Inoltre, questo suo ruolo è rafforzato dall'assenza di un legame orizzontale, a rete, tra i Nuclei medesimi (Scarpitti, 2001).

Dal punto di vista legislativo, lo sviluppo del sistema di valutazione dell'università italiana corre parallelo a quello dei sistemi di controllo interni alla pubblica amministrazione, istituiti dalla legge Bassanini n. 59 del 1997 e relativi decreti attuativi, in particolare il decreto n. 286 del 1999 di «riordino e potenziamento dei meccanismi e strumenti di monitoraggio e valutazione dei costi, dei rendimenti e dei risultati dell'attività svolta dalle amministrazioni pubbliche». In tale legge, e in quelle che successivamente la modificano parzialmente, si istituisce la chiara distinzione tra controllo di gestione, demandato a un organo interno all'amministrazione per controllare l'economicità, efficienza ed efficacia della gestione (nel sistema universitario tale ruolo è svolto dai Nuclei di

valutazione interna), e controllo sulla gestione, affidato a un organo esterno come la Corte dei Conti, cui è demandata «l'analisi, preventiva e successiva, delle conseguenze e/o degli eventuali scostamenti tra le missioni affidate dalle norme, gli obiettivi operativi prescelti, le scelte operative effettuate e le risorse umane, finanziarie e materiali assegnate, nonché l'identificazione degli eventuali fattori ostativi, delle eventuali responsabilità per la mancata o parziale attuazione, dei possibili rimedi».

1.3. La valutazione triennale della ricerca, la Vtr 2001-2003

La Vtr 2001-2003 è la prima vera esperienza di valutazione della qualità della ricerca prodotta dal sistema universitario italiano, approvata dal Miur, che con il d.m. n. 2206 del 2003 ha affidato al Comitato di indirizzo per la valutazione della ricerca (Civr) l'attuazione e l'organizzazione dell'intero esercizio di valutazione. Il Civr è istituito dal Miur con d.m. n. 204 del 1998, in cui si legge che il Comitato «è composto da non più di 7 membri, anche stranieri, di comprovata qualificazione ed esperienza, scelti in una pluralità di ambiti metodologici e disciplinari. Il comitato opera per il sostegno alla qualità e alla migliore utilizzazione della ricerca scientifica e tecnologica nazionale, secondo autonome determinazioni con il compito di indicare i criteri generali per le attività di valutazione dei risultati della ricerca».

Diversi sono i compiti assegnati al Civr in occasione dell'organizzazione della Vtr: a) la scrittura delle linee guida; b) la definizione dei criteri per la composizione dei panel di area e la proposta al Miur dei componenti da nominare; c) la valutazione dei dati trasmessi dalle strutture; d) la promozione di incontri periodici con i componenti dei panel; e) la valutazione delle relazioni e dei rapporti inviati dai panel; f) l'organizzazione di eventuali audizioni con le parti interessate; g) la relazione finale per singola struttura; h) relazioni periodiche di valutazione del Sistema Nazionale della Ricerca; i) la pianificazione dell'utilizzazione delle risorse ad esso destinate.

Alla Vtr sono invitati a partecipare le università e gli enti di ricerca pubblici e anche privati convenzionati con il Ministero.

Due sono le principali linee guida decise dal Civr per la Vtr 2001-2003: la volontarietà delle strutture di ricerca di aderire al programma di

valutazione³, l'attenzione alla produzione scientifica di alta qualità pubblicata dal 2001 al 2003. Riguardo a quest'ultimo aspetto, il Civr ha deciso di concentrare la valutazione su pochi prodotti considerati l'eccellenza scientifica. A ciascun ateneo, infatti, è stato chiesto di selezionare un numero esiguo di prodotti, pari al 50% del numero medio nel triennio 2001-2003 dei docenti e ricercatori in esso operanti. Per le strutture non vi è stato alcun obbligo di presentare prodotti in tutte le aree scientifiche; di conseguenza ciascun ateneo ha selezionato i prodotti nelle aree in cui godeva di maggior prestigio presso la comunità accademica nazionale e internazionale. Proprio per questo la Vtr non può essere considerata un esercizio di valutazione complessiva dell'intero sistema universitario, bensì un tentativo di valutare gli atenei per la loro capacità di contribuire all'eccellenza della produzione scientifica nazionale (Reale, 2013a).

I tipi di prodotto sottoposti a valutazione sono stati: libri e capitoli di libro, articoli su rivista, brevetti, progetti, composizioni, disegni, performance, mostre, esposizioni, manufatti e opere d'arte. Il compito più oneroso per le strutture è stata la selezione dei prodotti da inviare a valutazione. La procedura più adottata ha previsto una prima selezione interna al dipartimento; i prodotti scelti sono stati poi ulteriormente selezionati dai Car, Comitati di area nominati dall'ateneo, e infine inviati ai Cat, Comitati di ateneo (costituiti dai presidenti dei Car); dopo essere stata certificata dal Nucleo di valutazione, la lista dei prodotti veniva inviata ai panel di area competenti.

Ciascun panel delle 20 aree scientifico-disciplinari è stato composto da un minimo di 5 a un massimo di 9 esperti. I membri dei panel di area sono stati scelti in totale autonomia dal Civr, il quale attraverso una *call for experts* ha avuto cura di selezionare studiosi di comprovata esperienza nel settore scientifico di riferimento, con un elevato profilo accademico nazionale e internazionale (Civr, 2006).

Ai membri del panel di area competente è spettato il compito di nominare i due revisori di ciascun prodotto, che è stato poi valutato secondo cinque criteri: qualità, rilevanza, originalità e innovazione, internazionalizzazione⁴.

³ Nonostante ciò, tutte le strutture di ricerca pubbliche e private convenzionate con il Ministero hanno dato il proprio assenso.

⁴ Nel bando scritto dal Civr sulla Vtr 2001-2003 si legge che i revisori esprimono un giudizio di merito sui prodotti scientifici che sono chiamati a valutare, secondo i criteri della: qualità, rilevanza, originalità/innovazione, internazionalizzazione. Il bando rimanda poi al documento

Per la selezione dei revisori si è proceduto alla creazione di un database di esperti valutatori, con la partecipazione di tutta la comunità accademica che ha individuato i migliori esperti di ogni area scientifica. Nel complesso i diversi panel di area hanno fatto ricorso a 6.661 revisori che hanno valutato 17.329 prodotti. La procedura di valutazione è stata la *peer blind review*: ciascun prodotto è stato letto da due revisori anonimi e giudicato come eccellente, buono, accettabile, limitato. Nel caso di giudizio eccessivamente discordante tra i due revisori, il Civr ha previsto la possibilità di richiedere la valutazione di un terzo esperto.

Inoltre, i prodotti venivano ponderati in funzione delle affiliazioni dei loro autori; si è infatti voluto dare un peso maggiore al prodotto eccellente scritto da più autori afferenti, tutti o quasi, alla stessa struttura che ha presentato il prodotto, rispetto al prodotto eccellente scritto da più autori afferenti, per la maggior parte, a strutture scientifiche diverse dall'ateneo che lo ha presentato. Solo nel primo caso si può parlare di un'eccellenza che si trasferisce dal prodotto all'ateneo o centro di ricerca, poiché gli autori del prodotto valutato eccellente erano, al momento della pubblicazione, in carica presso la struttura che ha inviato il prodotto a valutazione; nel secondo caso no (Civr 2003a).

Nelle linee guida della Vtr 2001-2003 si legge che questo esercizio di valutazione si pone anche l'obiettivo di valutare la capacità delle strutture di sostenere e incentivare la ricerca nel medio periodo. Quattro sono stati gli indicatori utilizzati: numero di ricercatori in formazione (numero di dottorandi, assegnisti e borsisti di ricerca); il grado di mobilità internazionale in entrata e in uscita dei ricercatori di ruolo; l'ammontare delle risorse per la ricerca derivanti da fondi esterni all'ateneo (progetti Prin, fondi europei; ecc.); l'ammontare delle risorse interne destinate all'attività di ricerca.

Le strutture hanno poi ricevuto una valutazione sulla capacità di perseguire gli obiettivi di terza missione. Come indicatori sono stati

redatto dal Civr (2003b), contenente le linee guida per la valutazione della ricerca, per una trattazione più approfondita dei criteri di valutazione. Purtroppo, a un'attenta lettura del documento in questione, non si trova una chiara definizione dei quattro indicatori. Si legge infatti: «la valutazione della qualità scientifica e della rilevanza dei risultati si fonda sulla *peer review* e sull'applicazione di indicatori oggettivi, tra i quali, nei settori pertinenti, sono inclusi gli indici bibliometrici (in particolare, impact factor e citation analysis)». In nota si ricorda come «Il giudizio deve comprendere anche i riferimenti a originalità, innovazione e internazionalizzazione».

considerati il numero di brevetti depositati e/o attivi, e i costi e ricavi derivanti da ciascun brevetto⁵.

Il 26 gennaio del 2006 sono stati diffusi tutti i giudizi sintetici dei due revisori di ciascun prodotto, e la *ranking list*, cioè la classifica delle strutture in base al giudizio medio dei propri prodotti inviati a valutazione⁶. I panel hanno redatto anche un documento con i punti di forza e le criticità di ciascuna area scientifica (Civr, 2006). Ogni ateneo ha potuto così controllare la valutazione dei prodotti inviati, la propria posizione in classifica e le aree scientifiche in cui ha ricevuto le valutazioni migliori (Bressan, 2007).

La Vtr 2001-2003 è considerata un punto di riferimento per i successivi esercizi di valutazione per via di alcune pratiche di qualità, tra cui il metodo di valutazione dei pari, che ha evitato il ricorso a sterili coefficienti bibliometrici (specialmente nelle aree umanistiche dove questi sono poco diffusi), e l'efficiente sistema informatico che ha supportato gli atenei in tutte le fasi di selezione dei prodotti (al riguardo vedi Frabboni e Sacchetta, 2005). D'altronde, non sono mancate le critiche, tra cui la libertà concessa agli atenei di scegliere in quali aree scientifiche inviare i prodotti meglio valutabili, tralasciando le aree in cui gli atenei erano meno attrezzati per una rigorosa valutazione esterna, e quella di selezionare più prodotti pubblicati da un solo ricercatore, particolarmente brillante, trascurando i docenti meno attivi. Questa strategia ha reso la Vtr un esercizio di valutazione della produzione scientifica di eccellenza e non di valutazione complessiva della ricerca del sistema universitario italiano.

In realtà l'impatto più consistente prodotto da questa valutazione è stato di tipo culturale, perché ha reso evidente che era non solo concretamente possibile ma anche opportuno un esercizio di valutazione della ricerca nelle università italiane, obiettivo considerato fino allora non fattibile e non desiderabile da larga parte del corpo accademico. In diversi settori disciplinari le comunità accademiche si sono trovate a dover riflettere su un esito che rendeva conoscibile a tutti la qualità attribuita alle proprie università rispetto ad altre sedi accademiche, innestando un dibattito che è stato molto formativo rispetto all'importanza che lo strumento valutativo

⁵ Questo indicatore è stato aspramente criticato, vista la mancanza di dati ufficiali sui costi e ricavi di ciascun ateneo in merito ai propri brevetti, spin-off, ecc.

⁶ Le strutture sono state differenziate in base alle proprie dimensioni: piccole, se hanno inviato da 1 a 9 prodotti; medie, quando hanno presentato da 10 a 24 prodotti; grandi da 25-74; mega, oltre i 75 prodotti.

può avere per l'organizzazione e il governo interno dell'università. Sotto questo profilo, la Vtr ha rappresentato un punto di non ritorno, consolidando in via definitiva l'idea che la valutazione è attività possibile e ineludibile, anche quando si tratti di valutazione della ricerca (Reale, 2013b, p. 152).

1.4. La valutazione della qualità della ricerca, la Vqr 2004-2010

La Vqr 2004-2010 è il secondo esercizio di valutazione nazionale della qualità della ricerca prodotta dal sistema universitario italiano. Con il d.m. n. 17 del 2011 il Miur incarica l'Anvur, la neonata Agenzia nazionale di valutazione del sistema universitario e della ricerca, della scrittura del bando della Vqr 2004-2010 e dell'organizzazione dell'intero esercizio di valutazione.

Con la legge n. 286 del 2006 l'Anvur sostituisce il Civr e il Cnvsu ereditandone il ruolo e le funzioni, con molteplici aggiunte che prevedono l'analisi e la valutazione complessiva dell'intero sistema universitario italiano⁷. Soprattutto, l'Anvur è l'organismo deputato all'organizzazione del secondo esercizio di valutazione nazionale della qualità della ricerca prodotta dal sistema universitario italiano, e, come detto nel par. 1.1, gli esiti di tale valutazione hanno conseguenze sulla distribuzione premiale dei fondi.

Rispetto all'esercizio di valutazione che l'ha preceduta, la Vqr 2004-2010 ha presentato numerose novità. In primo luogo, cambia l'arco temporale in cui sono stati pubblicati i prodotti sottoposti a valutazione: si passa da tre a sette anni, per rispondere alle critiche di gran parte della comunità accademica che in seguito alla Vtr aveva obiettato come tre anni fossero troppo pochi per raccogliere un paniere di prodotti sufficientemente rappresentativo della qualità espressa da un ateneo (Minelli Reborà e Turri, 2008). Inoltre, cambia l'unità della valutazione, non più il solo ateneo bensì l'ateneo e suoi dipartimenti considerati come l'unità di funzionamento essenziale per il modello organizzativo dell'università; un cambiamento che sposta l'occhio della valutazione dai livelli più alti alle strutture intermedie di *governance* dell'università.

I tipi di prodotto sottoposti a valutazione sono gli stessi della Vtr: libri e capitoli di libro, atti di congresso, articoli su rivista, brevetti,

⁷ Vedi il par. 1.6 per un approfondimento sull'Anvur.

composizioni, disegni, performance, mostre, esposizioni, manufatti e opere d'arte, banche dati e *software*, carte tematiche.

A ciascun docente di ruolo è stato chiesto di inviare 3 prodotti pubblicati dal 2004 al 2010; ciò ha comportato un'impennata del numero di prodotti inviati a valutazione: si passa dai 17.329 della Vtr ai 179.280 della Vqr. Tutti i docenti hanno dovuto selezionare e inviare i propri prodotti, nessuno escluso, impedendo quindi all'ateneo di sceglierne una quota maggiore tra i docenti reputati più brillanti. Il docente che non ha inviato alcun prodotto è stato considerato "non attivo" e chi ne ha inviati meno di 3 "parzialmente attivo"; i docenti non attivi o parzialmente attivi hanno causato una penalizzazione nel giudizio finale del proprio dipartimento e ateneo.

Il concetto di "ricercatore attivo" introduce un parametro che valorizza quanti effettivamente contribuiscono all'attività di ricerca, e deprime chi non contribuisce, creando una differenziazione di status all'interno del corpo accademico. Il numero di pubblicazioni da sottomettere diventa indirettamente un parametro minimo di produttività per i singoli ricercatori: la Vqr propone, infatti, un numero di lavori di alta qualità che è lecito attendersi da ciascun ricercatore o docente di qualunque disciplina o settore (Reale, 2013b, p. 153).

Il bando Vqr chiarisce che i docenti devono fare ricorso allo strumento informatico messo a disposizione dal Cineca per indicare, in ordine di priorità, i prodotti di ricerca dai quali l'università di appartenenza sceglierà poi quelli da inviare a valutazione; alla struttura spetta l'importante compito di dirimere i conflitti di attribuzione in caso di coautoraggio, poiché un prodotto può essere attribuito a un solo autore⁸. Rispetto alla Vtr 2001-2003 i docenti assumono un ruolo più rilevante nel processo complessivo, poiché l'università può solo scegliere i prodotti da inviare ai Gev tra quelli presenti nella lista e selezionati *in primis* dai ricercatori (Rebora e Turri, 2010).

Per ogni area scientifica⁹ l'Anvur nomina i Gev, i gruppi di esperti

⁸ I prodotti inviati due volte dal medesimo ateneo sono valutati una volta sola, anche se scritti da due o più autori afferenti allo stesso ateneo; se invece un prodotto è stato scritto da più autori di diverse università, quello è valutato più volte.

⁹ Rispetto alla Vtr, nella Vqr è cambiato il numero delle aree scientifiche: non più 20 bensì 14. Rimane invariata la possibilità che, su richiesta dei Gev, siano istituite sub-aree per le aree caratterizzate da particolare eterogeneità disciplinare.

valutatori, che a loro volta nominano i due revisori di ciascun prodotto. I Gev di area ricevono il prodotto in formato elettronico, corredato da alcuni metadati bibliografici, tra i quali il settore scientifico di riferimento. I prodotti afferenti ad alcune aree disciplinari sono stati valutati attraverso specifici indici bibliometrici¹⁰; si tratta della novità più rilevante rispetto alla Vtr, che invece si è basata esclusivamente sulla *peer review*. Ai Gev è stato affidato il compito di decidere quali prodotti vanno valutati con l'uno o l'altro approccio, in base al tipo di prodotto e all'area scientifica dello stesso¹¹, con l'unico limite di mantenere prevalente il numero di prodotti valutati con *peer review*.

Il giudizio dei revisori è stato articolato sui criteri della rilevanza, originalità, internazionalizzazione. Nel bando dell'Anvur si legge che per rilevanza «è da intendersi il valore aggiunto per l'avanzamento della conoscenza nel settore e per la scienza in generale, anche in termini di congruità, efficacia, tempestività e durata delle ricadute», per originalità «il contributo all'avanzamento di nuove conoscenze o nuove acquisizioni nel settore di riferimento», per internazionalizzazione «il posizionamento nello scenario internazionale, in termini di rilevanza, competitività, diffusione editoriale, e apprezzamento della comunità scientifica» (Anvur, 2011, p. 7).

Lo strumento di lavoro dei revisori è la scheda di valutazione, composta da tre domande chiuse, ognuna delle quali rileva uno dei criteri di valutazione della qualità del prodotto. La somma dei punteggi equivalenti ai giudizi espressi dai revisori su ciascuno dei tre indicatori costituisce il giudizio sintetico di qualità, che assegna il prodotto a una delle classi di merito¹². Rispetto alla Vtr, cambia il numero delle classi e il punteggio di ciascun prodotto che in esse è classificato. Nella Vtr il giudizio era articolato su quattro livelli: eccellente (1), il prodotto si colloca tipicamente nel 20% superiore della scala di valore condivisa dalla comunità scientifica internazionale; buono (0,8), il prodotto si

¹⁰ In particolare, si fa riferimento all'analisi del numero di citazioni del prodotto e al fattore di impatto della rivista, ove applicabile.

¹¹ L'indicazione data ai Gev dall'Anvur è di prendere in considerazione la valutazione bibliometrica solo per i tipi di prodotto (come articoli su rivista o proceedings) per i quali esistono già basi di dati adatti all'utilizzo di tali metriche. Inoltre i Gev hanno dovuto tenere in debito conto il settore scientifico disciplinare del prodotto; è risaputo che nelle discipline scientifiche il ricorso alla bibliometria è pratica da tempo diffusa; nelle scienze umanistiche non lo è ancora.

¹² Alla fine della compilazione della scheda, al revisore appariva la classe di merito finale di assegnazione che poteva eventualmente modificare, cambiando i punteggi dati ai tre indicatori.

colloca nel segmento 60% - 80%; accettabile (0,6), il prodotto si colloca nel segmento 40% - 60%; limitato (0,2) il prodotto si colloca nel 40% inferiore. Nella Vqr, invece, le classi di merito sono sei: eccellente (1), la pubblicazione si colloca nel 20% superiore della scala di valore condivisa dalla comunità scientifica internazionale; buono (0,8), la pubblicazione si colloca nel segmento 60% - 80%; accettabile (0,5), la pubblicazione si colloca nel segmento 50% - 60%; limitato (0), la pubblicazione si colloca nel 50% inferiore; la pubblicazione non è valutabile (-1); la pubblicazione non è stata presentata (-0,5).

Dall'analisi delle nuove classi di merito e dei relativi punteggi si evince quanto nella Vqr, rispetto alla Vtr, sia stata incrementata la selettività nel giudicare un prodotto di qualità. In primo luogo, i prodotti che ricadono nella fascia peggio valutata non ottengono alcun punteggio; nella Vtr avevano un peso di 0,2. Inoltre, i prodotti accettabili sono valutati 0,5 e non più 0,6 ma soprattutto aumenta la distanza, in termini di valutazione, tra questi e i prodotti considerati buoni ed eccellenti. Infine, sono istituite due classi di merito che danneggiano gravemente i ricercatori inattivi e gli atenei che, per errore o incuria, hanno inviato prodotti non valutabili (Rebora e Turri, 2010).

Nel complesso la Vqr 2004-2010 si è rivelata essere un esercizio di valutazione di enorme portata, che ha coinvolto 95 università e 38 enti di ricerca, 61.822 ricercatori, 436 esperti valutatori (d'ora in poi Ev) che hanno composto i Gev, 14.770 revisori di cui 4.620 in servizio presso strutture straniere. Inoltre, per la prima volta la connessione tra l'esito della valutazione e l'allocazione dei fondi pubblici all'università diventa uno dei cardini dell'intero esercizio di valutazione, a tal punto che le relazioni redatte dall'Anvur e inviate al Miur sono state poi utilizzate per ripartire quote considerevoli del finanziamento pubblico alla ricerca (Rebora, 2010).

1.5. La valutazione della qualità della ricerca, la Vqr 2011-2014

La Vqr 2011-2014 è il terzo esercizio di valutazione del sistema universitario italiano, istituito dal Miur che emana il d.m. n. 458 del 2015 contenente le linee guida. L'Anvur è l'agenzia incaricata di organizzare l'intero esercizio di valutazione e di scrivere il bando che sarà poi pubblicato il 30 luglio 2015.

La Vqr 2011-2014 è caratterizzata da alcuni elementi di discontinuità rispetto alla Vqr precedente. La novità principale sta nell'ampiezza dell'arco temporale di 4 anni che viene considerato utile per selezionare i prodotti pubblicati da sottoporre a valutazione; tale periodo era di 7 anni nella precedente Vqr e di 3 nella prima Vtr. Si ha quindi la sensazione che questo nuovo *range* costituisca un parametro stabile nei prossimi esercizi di valutazione.

I tipi di prodotto sottoposti a valutazione rimangono sostanzialmente gli stessi: libri e capitoli di libro, articoli su rivista, brevetti, composizioni, disegni, performance, mostre, esposizioni, manufatti e opere d'arte, banche dati e *software*, carte tematiche (Anvur, 2015). Questa volta vengono però esclusi i manuali e testi meramente didattici, le recensioni di un singolo lavoro che non abbiano analisi critica della letteratura sull'argomento, brevi voci enciclopediche o di dizionario senza carattere di originalità, brevi note a sentenza di tipo redazionale senza carattere di originalità o meramente ricognitive, brevi schede di catalogo prive di contributi scientifici autonomi.

Il numero di prodotti da inviare a valutazione passa da 3 a 2 per ciascun docente universitario; cambia anche il numero delle aree scientifiche, che passano da 14 a 16¹³. La nomina dei Gev di ciascuna area rimane prerogativa dell'Anvur; come nella precedente Vqr, compito principale dei Gev è nominare i revisori dei prodotti: due per ciascuno di essi¹⁴.

Come nella precedente Vqr, anche per questo esercizio di valutazione lo strumento di lavoro dei revisori è la scheda di valutazione. In parte cambiano i criteri di valutazione della qualità dei prodotti: originalità, rigore metodologico, impatto nella comunità scientifica internazionale di riferimento. A seguito dei giudizi formulati dai revisori su queste tre dimensioni, ogni pubblicazione è attribuita a una delle seguenti classi di merito: eccellente, elevato, discreto, accettabile, limitato, non valutabile¹⁵.

Dall'analisi delle nuove classi di merito e i relativi punteggi si evince quanto nella seconda Vqr sia stata incrementata la selettività nel giudicare un prodotto di qualità eccellente: sono considerati eccellenti

¹³ Vedi la Tab. 1 e i successivi commenti sulle ragioni di tale ampliamento.

¹⁴ Sui criteri di composizione dei Gev e di selezione dei revisori dei prodotti vedi il cap. 3.

¹⁵ La scheda di valutazione dei revisori, i criteri di giudizio della qualità del prodotto e le classi di merito cui il prodotto è attribuito sono descritti analiticamente nel capitolo 4.

i prodotti che idealmente si trovano nel primo decile della produzione scientifica e non più nel primo 20%. In compenso i punteggi vengono distribuiti più equamente tra le classi cui sono attribuiti i prodotti di qualità intermedia. Nella prima Vqr i prodotti di qualità buona, “elevata” nella seconda Vqr, avevano un punteggio di 0,8; nella seconda Vqr di 0,7. Nella Vqr 2004-2010 i prodotti accettabili, “discreti” nella Vqr 2011-2014, avevano 0,5; nella seconda 0,4. Al resto dei prodotti di qualità inferiore è assegnato un punteggio di 0,1 e non più 0. Cambia anche il punteggio (0) assegnato ai prodotti attribuiti alla classe “non valutabile”, che quindi non contribuiscono più a deprimere il punteggio generale dell’istituzione di appartenenza (nella prima Vqr ricevevano un punteggio di -1).

Non mancano gli elementi di continuità tra i due ultimi esercizi di valutazione. Le tecniche di valutazione rimangono invariate. I prodotti sono stati valutati sia con informazioni di natura bibliometrica (utilizzando soprattutto l’*impact factor* della rivista su cui è pubblicato il prodotto e il numero di citazioni ricevute dal prodotto¹⁶), sia con l’*informed peer review*¹⁷, ricorrendo alle informazioni associate a ciascun prodotto inviato a valutazione (i metadati bibliografici del prodotto, identificativo Orcid dell’autore o dei coautori, l’area e il settore concorsuale e il settore scientifico disciplinare, l’*abstract* del prodotto, la lingua del prodotto, l’indicazione che il prodotto viene da un’area scientifica meglio valutabile con *peer review* o con indici bibliometrici).

Nella Tab. 1 sono riportati i dati relativi ai prodotti valutati con *peer review* o con bibliometria nelle diverse aree scientifiche. Da essa risulta che i prodotti di area 8a, 10, 11a, 12 e 14 sono stati valutati esclusivamente tramite la *peer review*.

Questa tabella mostra anche come la disaggregazione dell’area 8 e dell’area 11 in quattro distinte aree sia stata quanto mai opportuna. Nella Vqr 2004-2010 architettura e ingegneria civile costituivano l’area 8; scienze storiche, filosofiche, psicologiche e pedagogiche erano parte

¹⁶ Se i due indici bibliometrici hanno valutato il prodotto in modo diametralmente opposto (ad esempio, un prodotto con elevato numero di citazioni pubblicato su una rivista con impatto molto basso o viceversa), lo stesso è stato classificato IR e sottoposto a valutazione con *peer review*.

¹⁷ Per comporre l’albo dei revisori, incaricati di valutare i prodotti con *peer review*, l’Anvur ha selezionato alcuni nomi dal database Reprise del Miur, con criteri di merito scientifico (indice H di Hirsch, numero di citazioni, produzione scientifica recente), e altri non inclusi in Reprise, valutati con i medesimi criteri. Solo il Gev 12 ha pubblicato un modulo di autocandidatura per chi avesse voluto dare il proprio contributo come revisore. Il database finale è di 14.000 nomi.

dell'area 11. Vista l'eccessiva eterogeneità tra le discipline presenti all'interno di ciascuna delle due aree, nella Vqr 2011-2014 si è deciso di scorporare architettura (la nuova area 8a), valutata esclusivamente con *peer review*, da ingegneria civile (ora area 8b), valutata principalmente con indici bibliometrici, e le scienze storiche, filosofiche e pedagogiche (la nuova area 11a), valutate con giudizio dei pari, dalle scienze psicologiche (ora area 11b), valutate soprattutto bibliometricamente.

Tab. 1 - Numerosità e percentuale di prodotti valutati con peer review o con bibliometria nelle diverse aree scientifiche

<i>Area</i>	<i>Prodotti conferiti</i>	<i>Prodotti sottoposti alla peer review</i>	<i>% sui prodotti conferiti</i>	<i>Prodotti con valutazione bibliometrica</i>	<i>% sui prodotti conferiti</i>	<i>Prodotti con valutazione IR</i>	<i>% sui prodotti conferiti</i>
1	6.062	2.356	38,9	3.680	60,7	1.074	17,7
2	10.588	1.291	12,2	9.287	87,7	890	8,4
3	6.897	1.394	20,3	5.464	79,2	1.049	15,2
4	4.430	1.257	28,4	3.121	70,5	840	19,0
5	10.986	2.183	19,9	8.672	78,9	1.758	16
6	16.693	3.731	22,4	12.722	76,2	2.266	13,6
7	7.541	2.463	32,7	5.040	66,8	1.343	17,8
8a	3.456	3.433	99,3	5	0,1	0	0
8b	2.832	996	35,2	1.830	64,6	543	19,2
9	11.564	3.346	28,9	8.142	70,4	1.887	16,3
10	8.744	8.720	99,7	5	0,1	3	0
11a	6.123	5.956	97,3	151	2,5	37	0,6
11b	2.276	868	38,1	1.380	60,6	397	17,4
12	8.488	8.431	99,3	0	0	0	0
13	8.385	2.662	31,7	5.531	65,9	15	0
14	2.971	2.953	99,4	9	0,3	1	0,3
To- tale	118.036	52.040	44,1	65.039	55,1	12.103	10,3

Fonte: Anvur (2017)

La valutazione dei prodotti di area 2, 3, 4, 5, 6, 7, 8b, 9 e 11b è stata in larga misura bibliometrica e ha riguardato gli articoli pubblicati su riviste indicizzate nei database di Wos¹⁸ e Scopus¹⁹. L'Anvur ha acquistato da Thomson-Reuters ed Elsevier le informazioni bibliometriche di tali archivi per gli anni 2011-2014. Nel momento dell'invio dei prodotti a valutazione, alle università è stato chiesto di specificare la base dati (WoS oppure Scopus) adatta alla valutazione bibliometrica di ciascun prodotto,

¹⁸ login.webofknowledge.com

¹⁹ www.scopus.com

e l'indicatore di impatto preferito (IF5Y, Article Influence Score per WoS e IPP, SJR per Scopus). Oltre alle basi di dati WoS e Scopus, il Gev 1 ha utilizzato MathSciNet dell'American Mathematical Society, limitatamente all'indicatore d'impatto della rivista. Il Gev 13 ha impiegato un algoritmo bibliometrico che valorizza i prodotti con un numero di citazioni significativamente alto e quelli pubblicati da autori "prestigiosi" (tale indicatore è la città dove l'editore del prodotto ha sede).

Come nella prima Vqr, anche nella seconda Vqr sono stati oggetto di valutazione le università e i dipartimenti e non i singoli docenti. Per ciascuna università e dipartimento è stato stilato un profilo di qualità complessivo che è la combinazione del profilo di qualità dei prodotti di ricerca inviati a valutazione (che pesa per il 95%) e della capacità dell'istituzione di attrarre finanziamenti internazionali e statali (per il restante 5%). Per stabilire il profilo di qualità dei prodotti di ricerca di ciascuna università si è calcolato: il rapporto tra la somma delle valutazioni attribuite ai prodotti dell'ateneo nell'area scientifica di riferimento e la valutazione complessiva di area (indice *IrasI*); il rapporto tra il voto medio attribuito ai prodotti dell'università nell'area di riferimento e il voto medio ricevuto da tutti i prodotti dell'area (indice *R*); il rapporto tra la frazione di prodotti valutati eccellenti dell'istituzione nell'area di riferimento e la frazione di prodotti eccellenti dell'area (indice *X*).

Nella Vqr 2004-2010, al contrario, il profilo di qualità dei prodotti dell'ateneo era stabilito sostanzialmente da un solo indice, esito della somma dei punteggi conseguiti in base alle valutazioni dei prodotti inviati. In compenso questo indice pesava solo per il 50% della valutazione complessiva dell'istituzione; per il restante 50% contribuivano altri indici sintetici legati alla capacità dell'ateneo di attrarre finanziamenti pubblici e privati e sostenere adeguatamente la formazione e la mobilità dei propri ricercatori.

Come la precedente Vqr, anche questa si è rivelata essere un esercizio di valutazione di enorme portata, che ha coinvolto 96 università e 18 enti di ricerca, 60.455 addetti alla ricerca, 436 Ev che hanno composto i Gev (sulla distribuzione dei Gev tra le diverse aree scientifiche vedi la Tab. 2), 16.969 revisori di cui 3.423 in servizio presso strutture straniere²⁰.

²⁰ Nell'area 1 e nell'area 9 i revisori stranieri sono circa il 60% del totale, mentre nelle altre aree prevalgono i revisori italiani.

Tab. 2 - Numerosità dei Gev nelle diverse aree scientifiche

Area scientifica	Numerosità dei Gev
1 - Scienze matematiche e informatiche	22
2 - Scienze fisiche	33
3 - Scienze chimiche	23
4 - Scienze della terra	15
5 - Scienze biologiche	33
6 - Scienze mediche	62
7 - Scienze agrarie e veterinarie	23
8a – Architettura	16
8b - Ingegneria civile	11
9 - Ingegneria industriale e dell'informazione	38
10 - Scienze dell'antichità, filologico-letterarie e storico-artistiche	43
11a - Scienze storiche, filosofiche e pedagogiche	29
11b - Scienze psicologiche	6
12 - Scienze giuridiche	39
13 - Scienze economiche e statistiche	31
14 - Scienze politiche e sociali	12
Totale	436

Fonte: Anvur (2017)

I prodotti inviati a valutazione sono 118.036, di cui il 73,5% sono articoli su rivista, il 19,9% sono monografie/contributi su libro/cura-tele, il 5,8% contributi e *abstract* in atti di convegno (0,8% altro). Il 76,6% dei prodotti inviati è in lingua inglese; tale percentuale supera il 90% nelle aree bibliometriche; nelle aree non bibliometriche delle scienze umane, giuridiche e sociali prevale invece la lingua italiana.

Nel complesso il sistema universitario italiano, e in generale il mondo della ricerca, è stato valutato positivamente (vedi Tab. 3). Il 32,6% dei prodotti scientifici è valutato dai revisori eccellente (A) perché raggiunge i massimi livelli in termini di originalità e rigore metodologico, e ha conseguito o è presumibile che consegua un forte impatto nazionale e internazionale. Non basta. Il 30,8% dei prodotti è considerato di qualità elevata (B) avendo buoni livelli in termini di originalità e rigore metodologico, con un impatto significativo nella comunità scientifica nazionale e internazionale. Insomma, quasi i due

terzi della produzione scientifica italiana sono di alta qualità. Solamente l'11,6% è considerato accettabile (D) e il 3,5% limitato (E).

Tab. 3 - Numerosità e percentuale di prodotti attribuiti alle diverse classi di valutazione Vqr

	A	B	C	D	E	F	Totale prodotti inviati
Numero di prodotti inviati	38.435	36.394	24.433	13.639	4.178	957	118.036
% di prodotti inviati	32,6	30,8	20,7	11,6	3,5	0,8	100

Fonte: Anvur (2017)

Ovviamente non tutte le aree scientifiche contribuiscono egualmente a tale risultato (vedi Tab. 4).

Le scienze fisiche e le scienze chimiche sono le due aree scientifiche che hanno la quota maggiore di prodotti valutati come eccellenti. Seguono poi le scienze mediche, ingegneria industriale e dell'informazione, le scienze matematiche e informatiche, le scienze biologiche e ingegneria civile. Architettura, le scienze politiche e sociali e quelle giuridiche contribuiscono poco alla produzione scientifica di qualità eccellente.

Tab. 4 - Percentuale di prodotti per area attribuiti alle diverse classi di valutazione Vqr

	Area scientifica	A	B	C	D	E	F
1	(Scienze matematiche e informatiche)	38,4	28,0	18,2	10,8	4,2	0,4
2	(Scienze fisiche)	62,2	21,6	10,4	4,7	0,9	0,1
3	(Scienze chimiche)	49,2	32,0	12,9	4,6	0,8	0,6
4	(Scienze della terra)	27,9	29,7	21,6	14,2	5,5	1,2
5	(Scienze biologiche)	37,3	31,5	19,0	9,3	1,8	1,2
6	(Scienze mediche)	39,5	25,8	17,8	11,9	3,6	1,4
7	(Scienze agrarie e veterinarie)	28,4	31,5	19,4	14,9	5,4	0,5
8a	(Architettura)	8,6	34,2	35,9	16,0	4,8	0,5
8b	(Ingegneria civile)	37,6	29,3	17,7	12,6	2,5	0,2
9	(Ingegneria industriale e dell'informazione)	38,6	27,6	18,2	12,3	2,7	0,7
10	(Scienze dell'antichità, filologico-letterarie e storico-artistiche)	18,1	46,2	25,4	8,7	1,4	0,2
11a	(Scienze storiche, filosofiche e pedagogiche)	16,1	42,4	29,2	10,2	1,8	0,3
11b	(Scienze psicologiche)	30,8	23,4	19,1	18,7	6,8	1,2
12	(Scienze giuridiche)	7,8	41,2	35,9	12,2	2,2	0,7
13	(Scienze economiche e statistiche)	24,6	22,9	17,9	19,5	12,7	2,3
14	(Scienze politiche e sociali)	8,3	32,5	34,5	20,0	4,4	0,3

Fonte: Anvur (2017)

1.6. L'Anvur

Come anticipato nel par. 1.4, l'Anvur è istituita con la legge n. 286 del 2006 nel cui testo si legge: «Al fine di razionalizzare il sistema di valutazione della qualità delle attività delle università e degli enti di ricerca pubblici e privati destinatari di finanziamenti pubblici, nonché dell'efficienza ed efficacia dei programmi statali di finanziamento e di incentivazione delle attività di ricerca e di innovazione, è costituita l'Agenzia nazionale di valutazione del sistema universitario e della ricerca (Anvur), con personalità giuridica di diritto pubblico, che svolge le seguenti attribuzioni: a) valutazione esterna della qualità delle attività delle università e degli enti di ricerca pubblici e privati destinatari di finanziamenti pubblici, sulla base di un programma annuale approvato dal Ministro dell'università e della ricerca; b) indirizzo, coordinamento e vigilanza delle attività di valutazione demandate ai nuclei di valutazione interna degli atenei e degli enti di ricerca; c) valutazione dell'efficienza e dell'efficacia dei programmi statali di finanziamento e di incentivazione delle attività di ricerca e di innovazione».

Nonostante la centralità del ruolo assegnato all'Anvur nel nuovo sistema di valutazione dell'università, i lavori dell'agenzia sono stati avviati solo quattro anni più tardi, per via dei molti ritardi sull'emanazione del regolamento attuativo. Mi riferisco al decreto n. 76 del 2010 del Presidente della Repubblica, «Regolamento concernente la struttura ed il funzionamento dell'Agenzia nazionale di valutazione del sistema universitario e della ricerca (Anvur)», che insieme alla legge n. 240 del 2010 di riforma dell'università ne incrementa ulteriormente le competenze. Oggi l'Anvur, oltre a organizzare gli esercizi nazionali di valutazione della qualità della ricerca, accredita e valuta i dottorati di ricerca, i corsi di studio e le sedi universitarie; ha compiti in materia di valutazione della didattica e di terza missione; accredita gli enti privati che richiedono il riconoscimento dello Stato; monitora la *performance* amministrativa delle università e degli enti pubblici di ricerca; valuta le riviste scientifiche per le aree non bibliometriche ai fini dell'Abilitazione Scientifica Nazionale (Asn), e redige il Rapporto sullo stato del sistema universitario e della ricerca (Fantoni, 2015).

Dal confronto tra le mansioni attribuite all'Anvur dalla legge n. 286 del 2006 e quelle che l'agenzia attualmente svolge, si osserva un in-

cremento considerevole degli impegni a essa affidati. Il primo regolamento disciplinava il ruolo dell'Anvur (condurre la valutazione esterna del sistema universitario italiano, indirizzare e coordinare l'attività dei Nuclei di valutazione interna, valutare i programmi di finanziamento pubblico agli atenei) in quanto organo tecnico che avrebbe dovuto valutare principalmente l'attività di ricerca delle università. Ad oggi, l'Anvur è incaricata di valutare non solo la qualità della ricerca, ma anche la qualità della didattica, della formazione, degli obiettivi di terza missione, e dei servizi amministrativi che le università pubbliche erogano ai cittadini utenti.

L'accorpamento del Civr e Cnvsu nell'Anvur non è stata solo un'operazione di trasferimento di funzioni amministrative da un ente pubblico a un altro. L'Anvur ha assunto le funzioni della sola e unica agenzia nazionale di valutazione del sistema universitario e della ricerca, con cui tutti gli operatori del sistema universitario italiano devono necessariamente confrontarsi. Una centralità che al contempo richiede, per essere legittimata, piena autonomia nei confronti del suo principale portatore di interesse, il Miur, che stabilisce l'entità dei finanziamenti destinati ai singoli atenei e dipartimenti in base agli esiti degli esercizi di valutazione organizzati dall'Anvur.

L'autonomia dell'agenzia rispetto all'operato del Ministero è sancita dalla legge. All'articolo 1 comma 3 del decreto n. 76 del 2010 del Presidente della Repubblica, si legge che: «L'Agenzia ha personalità giuridica di diritto pubblico (...). È dotata di autonomia organizzativa, amministrativa e contabile, anche in deroga alle disposizioni sulla contabilità generale dello Stato». Peraltro, questa autonomia è una delle ragioni per cui l'Anvur ha ricevuto l'accreditamento internazionale da parte dell'ENQA (*European Association for Quality Assurance in Higher Education*)²¹.

Anche il meccanismo di nomina dei sette componenti del consiglio direttivo è trasparente e rafforza l'autonomia dell'agenzia dalla politica. All'articolo 8 comma 3 del suddetto decreto del Presidente della Repubblica si legge che: «I componenti del Consiglio direttivo sono nominati con decreto del Presidente della Repubblica, su proposta del

²¹ Come si legge negli *Standards and guidelines for quality assurance in the European Higher Education Area*: «Agencies should be independent and act autonomously. They should have full responsibility for their operations and the outcomes of those operations without third party influences» (Enqa, 2015, p. 22).

Ministro, sentite le competenti Commissioni parlamentari (...). Ai fini della proposta, il Ministro sceglie i componenti in un elenco composto da non meno di dieci e non più di quindici persone definito da un comitato di selezione appositamente costituito con decreto del Ministro. Il comitato di selezione è composto da cinque membri di alta qualificazione, designati, uno ciascuno, dal Ministro, dal Segretario generale dell'OCSE e dai Presidenti dell'Accademia dei Lincei, dell'*European Research Council* e del Consiglio nazionale degli studenti».

Inoltre, il legislatore ha ritenuto opportuno istituire un Comitato consultivo, nominato dal presidente dell'Anvur su proposta del Consiglio direttivo, con la funzione di dare «pareri e formulare proposte al Consiglio direttivo, in particolare sui programmi di attività e sui documenti riguardanti la scelta dei criteri e dei metodi di valutazione», come si legge nell'articolo 11 del decreto n. 76 del 2010 del Presidente della Repubblica. I candidati a divenire i diciotto membri del Comitato consultivo sono indicati dal Consiglio Universitario Nazionale (Cun), dalla Conferenza dei Rettori delle Università Italiane (Cruì), dal Consiglio Nazionale degli Studenti Universitari (Cnsu), dalla Conferenza dei Presidenti degli enti pubblici di ricerca, dall'Accademia dei Lincei, dal Consiglio nazionale dell'economia e del lavoro (Cnel), dalla Conferenza unificata Stato-Regioni, dall'*European Research Council*, dall'*European University Association*, dalla *National Union of Students in Europe*, dal Convegno permanente dei Direttori Amministrativi e dirigenti delle Università italiane (CoDau), e dal Segretario generale dell'OCSE. Il Comitato consultivo nasce dall'esigenza di istituzionalizzare un organo, interno all'agenzia ma composto da personalità del mondo dell'università e della ricerca, con cui l'Anvur si confronta per ricevere consigli e proposte in merito alla sua azione valutativa.

1.7. Una comparazione dei principali sistemi europei di valutazione della qualità della ricerca

In Europa il primo esercizio di valutazione nazionale della qualità della ricerca (il Rae, *Research Assesment Exercise*) si è tenuto nel 1986 in Inghilterra. In 5 occasioni²² il Rae ha valutato la qualità della

²² Nel 1989, 1992, 1996, 2001, 2008 (Palumbo, 2013).

produzione scientifica in tutti i paesi della Gran Bretagna. Nel 2014 il Rae è stato sostituito dal Ref (*Research Excellence Framework*), il cui impianto è stato confermato per l'immediato futuro: nel 2021 si svolgerà la seconda edizione del Ref.

In Gran Bretagna l'obiettivo dichiarato degli esercizi di valutazione della ricerca è sempre stato diversificare l'allocazione delle risorse pubbliche destinate alle università in base alla qualità della ricerca prodotta da ciascuna di esse. Tali attività, di gestione dell'esercizio di valutazione e di redistribuzione delle risorse, sono condotte da un'agenzia in ciascuna nazione britannica, coordinate dalla neonata agenzia Ukri (*United Kingdom Research and Innovation*). In Inghilterra se ne occupa il Rec (*Research England Council*), in Scozia il Sfc (*Scottish Funding Council*), in Galles l'Hefcw (*Higher Education Funding Council for Wales*), e il Dfe in Irlanda del Nord (*Department for the Economy in Northern Ireland*)²³. L'esercizio di valutazione incide su una quota parziale dell'intero ammontare di risorse destinate alle università pubbliche; sono i cosiddetti *recurrent fundings*, che vengono assegnati in base agli esiti degli esercizi di valutazione.

Sono valutati i dipartimenti e le università. Ciascun dipartimento fornisce informazioni riguardanti il personale attivo nella ricerca, i prodotti di ricerca (4 prodotti per ciascun ricercatore, valutati con *peer review*), i corsi di dottorato, le fonti di finanziamento alla ricerca, e anche una descrizione accurata dell'ambiente di ricerca; non è valutata la didattica. La valutazione è condotta dai membri dei panel operanti nelle diverse aree scientifiche²⁴; oggetto della valutazione sono la produzione scientifica, l'ambiente di ricerca, e una serie di indicatori di prestigio della struttura. Nella valutazione finale, il giudizio sulla produzione scientifica pesa per il 70%, l'ambiente di ricerca per il 20% e gli indicatori di prestigio per il restante 10%. Ciascun dipartimento ha la libertà di selezionare i ricercatori da sottoporre a valutazione; soli-

²³ Sono tutti enti pubblici i cui Boards sono di nomina politica.

²⁴ Nel Rae erano previsti 15 panel di esperti e relativi sotto-panel (da 3 a 8 per ogni panel), per un totale di 67 unità di valutazione scientifiche, ognuna per una differente area scientifico-disciplinare. Nel Raf le unità scientifiche di valutazione sono diventate 36. I panel principali coordinano i propri sotto-panel. I presidenti dei panel principali sono nominati dal ministro competente; i membri dei panel e dei sotto-panel sono esperti accademici di riconosciuta competenza nel proprio ambito scientifico, nominati dai direttori delle agenzie (Otley, 2010).

tamente «il dipartimento decide di presentare pochi ricercatori di qualità eccellente» (Perotti, 2008, p. 95). Il giudizio finale è articolato su 4 livelli: 1) qualità leader a livello mondiale, 2) qualità di eccellenza internazionale, 3) qualità riconosciuta a livello internazionale, 4) qualità riconosciuta a livello nazionale.

Nel 2006 il governo inglese ha avviato uno studio per riformare il Rae, specialmente a causa dei costi elevati e sempre crescenti. Obiettivo dichiarato del governo era passare da un sistema di valutazione basato sulla *peer review*, e per questo finanziariamente oneroso, a una valutazione bibliometrica dei prodotti scientifici. In un documento del Ministero delle Finanze si legge: «*The Government has firm presumption that after 2008 RAE the system for assessing research quality and allocating quality-related research funding to Universities from the Department for Education and Skills will be mainly metrics-based*» (HM Treasury, 2006, p. 3).

Nel 2008 si è tenuto l'ultimo Rae, e nel 2007 e 2010 si sono svolte due consultazioni per testare la possibilità di abbandonare *in toto* il sistema di valutazione basato sul giudizio dei pari con uno interamente bibliometrico; è stato condotto anche un esercizio pilota. Al termine di tale simulazione, si è deciso che la bibliometria avrebbe avuto solamente il ruolo di coadiuvare il giudizio dei pari, in quanto *informed peer review*.

Le novità principali del passaggio dal Rae al Ref sono: la diminuzione del numero di panel (36 anziché 67) e il peso assegnato, in sede di valutazione, all'impatto socioeconomico, quindi non accademico, dell'attività di ricerca. I valutatori Ref tengono in forte considerazione i cosiddetti obiettivi di terza missione, cioè la capacità dell'ateneo di fare ricerca che abbia ricadute a breve e medio termine sulle condizioni economiche e sociali dei cittadini e delle imprese del territorio in cui opera; lo studio di caso è lo strumento cui ci si affida per valutarne l'impatto. Sul giudizio finale alla struttura, l'impatto socioeconomico della ricerca pesa per il 20%, l'ambiente di ricerca per il 15% e la qualità della produzione scientifica per il 65% (Ref, 2010)

Nonostante queste novità, il programma di valutazione britannico ha mantenuto intatte le sue peculiarità. È infatti un programma intrinsecamente meritocratico perché *performance-based*, poiché premia la qualità e riconosce risorse aggiuntive a chi ha dimostrato di esserne

all'altezza. È valutata solo la ricerca, non i programmi di insegnamento; la valutazione è rigorosamente *ex post* da parte di un panel di esperti con l'uso combinato di *peer review* e indicatori bibliometrici; la valutazione è esterna all'ateneo e nazionale; l'esercizio di valutazione è finalizzato alla redistribuzione dei finanziamenti pubblici; la connessione tra i risultati dell'attività di valutazione e la ripartizione dei fondi pubblici è diretta (Hicks, 2011).

Ovviamente un sistema di valutazione totalmente *performance-based*, come quello britannico, non presenta solo vantaggi ma anche diverse controindicazioni. In primo luogo, una valutazione meritocratica, che si affida soprattutto al giudizio dei pari e solo parzialmente alla bibliometria, è molto costosa, in termini di risorse economiche e umane; richiede quindi finanziamenti sostanziosi e tempi lunghi, oltre a una capillare collaborazione di tutto il mondo accademico. Inoltre, se le università vengono valutate per la qualità della loro ricerca, il pericolo maggiore è che investano meno nell'attività didattica, diventando dei veri e propri centri di ricerca e non più l'istituzione cui è demandato il compito di trasmettere il sapere ai cittadini di domani. È probabile anche che le università concentrino le proprie risorse nelle aree scientifiche meglio spendibili in fase di valutazione (vedi le scienze ingegneristiche e mediche), deprimendo ulteriormente altre discipline (ad esempio quelle umanistiche e artistiche) già gravemente sottofinanziate (Geuna e Martin, 2003).

In Olanda l'attività di valutazione della qualità della ricerca universitaria è affidata a tre organismi: la *Royal Netherlands Academy of Arts and Sciences* (KnaW), la *Netherlands Organization for Scientific Research* (Nwo), l'*Association of Universities in Netherlands* (Vsnu). La KnaW raggruppa fino a 220 studiosi con elevato profilo scientifico, che svolgono attività di consulenza nei confronti del governo centrale in materia di politiche riguardanti il sistema nazionale della ricerca; il compito principale dell'organismo è finanziare 19 grandi istituti di ricerca in base agli esiti dell'esercizio di valutazione. La Nwo è un'agenzia governativa composta da un presidente e quattro membri nominati direttamente dalla corona, che rimangono in carica per 5 anni. Ci sono otto sezioni scientifiche, in ognuna delle quali sono valutati i progetti e le richieste di finanziamento che vengono sottoposti direttamente dai singoli ricercatori e dai gruppi di ricerca. Inoltre, 9 istituti di ricerca sono valutati e finanziati direttamente da questa agenzia. Infine, c'è la

Vsnu, un'associazione di rappresentanza del mondo universitario, composta dalle maggiori università del paese. Il suo ruolo è promuovere gli interessi dell'accademia olandese nei confronti del governo centrale, negoziando le politiche riguardanti il mondo dell'università. Nel 2002 al Knaw, al Nwo e al Vsnu il governo ha affidato l'onere di definire i criteri e le procedure con cui avviare il successivo esercizio di valutazione; la partnership tra accademia e politica ha dato alla luce il Sep (*Standard Evaluation Protocol*) 2009-2015 (Palumbo, 2013).

Il Sep valuta ogni sei anni gli istituti di ricerca attraverso una procedura di valutazione integrata: interna ed esterna all'università. La valutazione interna è una sorta di autoriflessione critica, condotta dallo stesso ateneo sottoposto a valutazione, sulla qualità del proprio lavoro di ricerca, che include la preparazione di una dettagliata documentazione sulla propria produzione scientifica, sulle strutture e sui programmi di ricerca cui l'università partecipa; è inviata ai valutatori quattro settimane prima della loro visita *in loco*. La valutazione esterna è implementata tramite una visita all'istituto di un gruppo di Ev, inviati dal Ministero dell'Università e della Ricerca, che intervistano il direttore del dipartimento, alcuni ricercatori e componenti dello staff (anche i dottorandi) per controllare le informazioni contenute nel documento di valutazione interna e per acquisirne di nuove utili alla valutazione.

Gli Ev sono nominati dai *Boards* delle tre agenzie, tra i nominativi suggeriti dagli istituti valutati; i valutatori sono soprattutto stranieri, così da evitare conflitti di interesse, e la numerosità dei singoli gruppi può variare molto in base alle esigenze delle agenzie. Il rapporto finale di valutazione è inviato ai *Boards* delle tre agenzie, che esprimono un giudizio sui risultati dell'attività di valutazione; tutti i documenti sono pubblicamente consultabili. Nel periodo che intercorre tra due valutazioni esterne, ciascun istituto è tenuto a riflettere sugli esiti della precedente valutazione per poter essere meglio valutato nella successiva occasione. Ciascun valutatore dà un giudizio finale su una scala di 5 livelli (eccellente, molto buono, buono, soddisfacente, insoddisfacente) che è poi discussa collegialmente per trovare un accordo comune (Knew e Nwo e Vsnu, 2009).

Il sistema olandese non può quindi definirsi *performance-based* perché la valutazione non è condotta esclusivamente *ex-post* da valutatori esterni ma anche *ex ante* da Nuclei interni all'università. Inoltre, oggetto di valutazione non è la sola produzione scientifica, valutata

con indicatori bibliometrici o con *peer review*. Sul giudizio finale pesa molto la qualità dei programmi di ricerca (intesa come originalità delle idee e dell'approccio alla ricerca, come reputazione e profilo scientifico dei partecipanti al programma di ricerca), la qualità dei programmi di formazione dottorale (intesa come l'ammontare delle risorse destinate ai progetti di formazione dei giovani ricercatori e come integrazione dell'attività dottorale con i progetti di ricerca dell'istituto), l'impatto dell'attività di ricerca sulla comunità che vive nel territorio in cui opera l'ateneo (intesa come la capacità della ricerca di migliorarne le condizioni socioeconomiche), la qualità delle strutture di ricerca e la pianificazione degli investimenti futuri.

Le università, affrancate dagli oneri imposti da una valutazione *performance-based*, sono dunque libere di orientare le proprie politiche accademiche in base alla *mission* che si sono date, ad esempio investendo in progetti di ricerca e formazione a lungo termine, anche in settori disciplinari nuovi o in aree scientifiche che al momento risultano essere poco attrattive per il mondo industriale. È un sistema di valutazione che promuove la pluralità degli approcci alla ricerca, incoraggia l'integrazione tra ricerca e alta formazione, facilita l'indipendenza dell'ateneo rispetto alle esigenze ondivaghe e passeggiere dei mercati. Inoltre, il Sep non ha l'obiettivo di ripartire le risorse pubbliche in base alla *performance* scientifica dell'università; lo scopo è promuovere una cultura di auto-valutazione interna alle strutture, che, attraverso un meccanismo di confronto e apprendimento critico, incoraggi i ricercatori a contribuire attivamente alla costruzione di progetti di ricerca di qualità che accrescano il prestigio del dipartimento e dell'università per cui lavorano (Rip e Van der Meulen, 1995).

In Francia il sistema di valutazione della qualità della ricerca è molto simile a quello in vigore nei Paesi Bassi. Qui si è deciso di accorpate le funzioni di valutazione del sistema universitario, prima svolte da diversi organismi, in un'unica agenzia di valutazione, l'Aeres (*Agence d'évaluation de la recherche et de l'enseignement supérieur*), oggi Hceres (*Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur*). L'agenzia è stata istituita nel 2007, ed è un'autorità amministrativa indipendente – ciò è sancito nel suo statuto – con lo scopo di valutare il sistema universitario con trasparenza, equità e affidabilità. È composta da un consiglio direttivo di 25 membri con un alto profilo nel mondo scientifico, nominati dall'agenzia per quattro anni, i quali hanno

potere di indirizzo sulle politiche di valutazione. Per ciascuna delle tre sezioni scientifiche (*Sciences humaines et sociales*, *Sciences dures*, *Sciences de la vie*) sono nominati un centinaio di delegati scientifici, responsabili delle procedure di valutazione nella propria sezione. Il ruolo chiave è occupato dalle migliaia di esperti, che, ciascuno nel dominio scientifico di competenza, monitorano e valutano l'adeguatezza del metodo di valutazione rispetto al proprio ambito disciplinare; gli esperti sono formati al sistema di valutazione dell'agenzia e il loro *feedback* è cruciale per il miglioramento costante di tutte le procedure di valutazione (Mesr, 2007).

La differenza principale tra il sistema di valutazione francese e gli altri due finora analizzati è l'unità di valutazione: non solo le università e i dipartimenti, anche i singoli team di ricerca. Lo scopo del programma è individuare i migliori gruppi di ricerca – ogni anno ne sono valutati circa 700 – all'interno di ciascuna area scientifica, assegnare loro un punteggio e inserirli in una classifica nazionale.

La valutazione del sistema universitario francese si fonda su due pilastri: l'autovalutazione, attraverso la redazione di un rapporto di valutazione interna all'università, e la valutazione esterna, condotta dai valutatori dell'agenzia, che integra la prima. Il periodo di valutazione dura solitamente nove mesi²⁵ ed è articolato in tre fasi: preparazione, visita, restituzione. Nella prima fase il Nucleo di esperti valuta il rapporto di valutazione interno redatto dal singolo team di ricerca; nella fase successiva, i valutatori fanno visita agli evaluandi per raccogliere le informazioni di controllo rispetto a quanto dichiarato nel primo documento; infine, viene redatto il documento dell'attività di valutazione con le osservazioni dei valutatori. In questo lasso di tempo si avvia un dialogo costante tra l'agenzia e la struttura valutata, al punto che il rapporto finale non viene reso pubblico se prima se non si è data possibilità all'università di commentarlo e discuterlo ove si trovasse in disaccordo con le riflessioni dei valutatori (Mesr, 2009).

Il Hceres valuta soprattutto la ricerca, sia a livello universitario, dipartimentale e dei singoli gruppi di lavoro. L'esito del rapporto di valutazione tiene conto: della qualità della produzione scientifica tramite *peer review* e indici bibliometrici, della partecipazione a progetti e programmi

²⁵ Ciascuna struttura è valutata ogni quattro anni; ogni anno sono valutate le università che ricadono in una determinata area geografica.

di ricerca nazionali e internazionali, dei finanziamenti ottenuti da agenzie nazionali e internazionali, dell'apertura a nuovi ambiti di ricerca interdisciplinari (la cosiddetta ricerca di frontiera), degli investimenti nella diffusione della cultura scientifica (European Commission, 2010).

Affinché la produzione scientifica del quadriennio di riferimento sia considerata apprezzabile, un ricercatore deve avere pubblicato almeno due articoli o monografie di rango A, cioè che sono state referate da un comitato scientifico. Per articolare ancor di più la valutazione della produzione scientifica, le riviste scientifiche sono state classificate in tre classi (A, B e C), suddivise in base all'ampiezza della produzione scientifica in esse pubblicata. I gruppi meglio valutati e le università loro affiliate ricevono finanziamenti aggiuntivi dal Ministero dell'Istruzione e della Ricerca (Palumbo, 2013).

In Spagna la valutazione della ricerca universitaria nazionale è condotta dall'Anep (*Agencia Nacional de Evaluacion y Prospectiva*) e dal Cneai (*Comision Nacional Evaluadora de la Actividad Investigadora*). Questi due organismi non svolgono un vero e proprio esercizio di valutazione nazionale congiunto; ciascuno lavora con propri tempi e procedure.

L'Anep è l'agenzia del Ministero della scienza e dell'innovazione, nata con lo scopo di finanziare i progetti dei ricercatori che ne fanno richiesta. Per ciascuna area un team di 4-8 coordinatori nominati dal Sottosegretario di Stato alla ricerca su proposta del direttore dell'agenzia, supportati da 120 assistenti, seleziona da un database di 20.000 nomi i valutatori che giudicano in modo anonimo i progetti di ricerca. La valutazione è condotta *ex ante* tramite *peer review* da Ev nelle diverse aree disciplinari. Il progetto, per la cui realizzazione il ricercatore chiede il finanziamento, è inviato all'Anep e assegnato a una delle 26 aree tematiche; per ogni area tematica vi è un gruppo di coordinamento che seleziona i valutatori competenti nell'area di afferenza del progetto. Il progetto è valutato secondo le sue qualità intrinseche (significatività, realizzabilità, originalità e interdisciplinarietà, chiarezza degli obiettivi, valenza dei risultati attesi) e la qualità della produzione scientifica dei suoi proponenti. Sulla scorta del giudizio di pari, si predispose la relazione finale e si comunica al ricercatore l'esito della valutazione (Palumbo, 2013).

Il Cneai, organo del Ministero, valuta la produzione scientifica di

quegli studiosi che decidono su base volontaria di sottoporsi a valutazione per ottenere uno stipendio più alto; l'incremento è a carico delle Regioni, e le commissioni sono nominate dal Ministero. La valutazione è dunque *ex post*. Il ricercatore chiede di sottoporre a valutazione la propria produzione scientifica; non c'è una selezione dei prodotti da parte del ricercatore, poiché la valutazione verte sull'intero *corpus* pubblicato negli ultimi sei anni. I prodotti sono assegnati a una delle 15 aree disciplinari previste; per ciascuna area è nominato un gruppo di coordinamento che rimane in carica per due anni. I valutatori, selezionati dal gruppo di coordinamento, dispongono di una serie di informazioni di carattere bibliometrico (grado di prestigio dell'editore o della rivista su cui si è pubblicato, il numero di citazioni ricevute, ecc.). Nelle scienze umane, dove indicatori di questo tipo sono poco utili, i valutatori si affidano alla *peer review* e ad altre informazioni, ad esempio la rigosità delle procedure di revisione cui è stato sottoposto il prodotto poi pubblicato (Jimenez-Contreras, 2010).

La Germania è certamente il paese europeo con il sistema di valutazione più complesso, perché in esso coesistono modelli di valutazione differenti. Gli attori coinvolti nel processo di valutazione sono diversi; ciascuno con proprie procedure di valutazione, precipue tempistiche, finalità e referenti.

Il Dfg è l'ente nazionale di diritto privato, costituito da tutte le università e centri di ricerca tedeschi, che si occupa di distribuire il 70% dei fondi pubblici destinati alle università, stanziati dallo stato centrale e dai *Länder*. Il Dfg finanzia i progetti individuali dei ricercatori che ne fanno richiesta, le borse di studio ai dottorandi e giovani ricercatori per i loro studi di specializzazione, i finanziamenti a gruppi di ricerca (da 5 a 25 componenti) impegnati in progetti complessi che si protraggono anche per diversi anni. Il Dfg nomina i panel dei revisori dei progetti di cui sono chiamati a giudicare *ex ante* la qualità tecnico-scientifica (l'idea progettuale, l'entità del finanziamento richiesto, la rilevanza dei risultati attesi), il *know-how* di cui dispongono i proponenti e la coerenza del loro *curriculum vitae* e delle loro esperienze rispetto al progetto presentato. La valutazione è praticamente a ciclo continuo, e solitamente la procedura si conclude in breve tempo.

Oltre alla valutazione del Dfg, ciascun *Land*²⁶ gestisce autonomamente l'attività di valutazione esterna dei singoli atenei. Ad esempio, lo stato del Baden-Württemberg ha costituito un'agenzia *ad hoc*, la Evalag, cui ha affidato la valutazione delle attività scientifiche delle sue università. Il panel di Ev, solitamente ogni cinque anni, valuta l'intero ateneo, grazie a una serie di informazioni di cui dispone (qualità e numerosità della produzione scientifica, ammontare dei fondi raccolti su progetti internazionali, parametri relativi all'attività didattica). La procedura di valutazione è *ex post* con *informed peer review*. La produzione scientifica è valutata in base al numero e prestigio delle pubblicazioni degli addetti alla ricerca, all'impatto del prodotto (il numero di citazioni ricevute), al posizionamento nel *journal ranking*²⁷ della rivista in cui il prodotto è stato pubblicato. L'ultimo livello di valutazione è interno a ciascuna università, la cui finalità è redistribuire i fondi tra le singole facoltà, dipartimenti e ricercatori. Ciascuna università decide autonomamente quali procedure adottare e come costituire i Nuclei di valutazione interna. In alcuni atenei queste valutazioni sono informali; le quote di finanziamento sono negoziate in base a valutazioni scientifiche e didattiche (come il numero di studenti, le ore di lezione, numero di docenti). In altre università la negoziazione lascia spazio a procedure standardizzate e altamente formalizzate, specialmente nell'assegnazione dei fondi ai singoli ricercatori, considerando il profilo di *performance* scientifica del singolo studioso. L'attività di valutazione interna è solitamente annuale e posteriore all'assegnazione statale dei fondi (Palumbo, 2013).

La Svezia dal 2009 ha invece adottato un modello integralmente *performance-based*, in cui si valuta solo la produzione scientifica.

²⁶ L'attuale sistema di valutazione è centrato sull'autonomia dei Länder, principio ulteriormente rafforzato dall'ultima riforma costituzionale del 2006. Attualmente il governo federale contribuisce per il 55% allo stanziamento dei fondi per le università pubbliche; i Länder contribuiscono per il restante 45%.

²⁷ Il *journal ranking* è la classificazione delle riviste scientifiche nazionali e internazionali in base al loro prestigio presso la comunità accademica tedesca. Vista l'insufficienza di indici bibliometrici in grado di cogliere la qualità della maggior parte dei prodotti scientifici dei ricercatori tedeschi (Schmitz, 2008), l'Accademia Tedesca di Economia Aziendale ha avviato un lavoro di classificazione delle riviste scientifiche in base alla qualità del processo di revisione cui sono sottoposti gli articoli inviati a valutazione. Grazie a due survey nel 2002 e nel 2008, i ricercatori e professori hanno contribuito a differenziare le riviste scientifiche in 5 classi molto selettive; basti pensare che le riviste top ranking presenti nella prima classe (A+) sono appena 12.

L'attività di valutazione poggia su due pilastri: uso diffuso della bibliometria e ampliamento al 25% della quota premiale di fondi pubblici. Tale modello pesa poco sul bilancio dello Stato centrale perché la valutazione è esclusivamente bibliometrica e non prevede in nessun caso il ricorso al giudizio dei pari. Lo scopo è dare priorità alle realtà universitarie che sono state in grado di attirare in Svezia i ricercatori che fanno ricerca industriale di qualità. Di conseguenza, alcune discipline scientifiche hanno acquisito, più di altre, prestigio nel mondo accademico e riconoscibilità presso l'opinione pubblica. Molto finanziate e apprezzate sono le scienze ingegneristiche, mediche, le nano e bio-tecnologie, in grado di rispondere al meglio alle esigenze industriali del mercato e di contribuire al benessere generale del paese; le scienze umanistiche e sociali – ma anche altre scienze dure la cui ricerca non fornisce applicazioni immediatamente spendibili sul mercato – sono sempre più marginalizzate nella redistribuzione dei fondi pubblici e anche nella rappresentazione che l'opinione pubblica ha del rapporto tra scienza e società (European Commission, 2010).

Riferimenti bibliografici

- Agnoli M.S. (1994), *Concetti e pratica nella ricerca sociale*, FrancoAngeli, Milano.
- 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: 1-17.
- Anvur (2011), *Valutazione della qualità della ricerca 2004-2010 (Vqr 2004-2010). Bando di partecipazione*, www.anvur.it
- Anvur (2015), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Bando di partecipazione*, www.anvur.it
- Anvur (2015a), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). La selezione dei componenti Gev*, www.anvur.it
- Anvur (2015b), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Linee guida dei gruppi di esperti della valutazione (Gev)*, www.anvur.it
- Anvur (2017), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale Anvur. Parte Prima: Statistiche e risultati di compendio*, www.anvur.it
- Anvur (2017a), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Politiche e Sociali (Gev 14)*, www.anvur.it
- Anvur (2017b), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Politiche e Sociali (Gev 14). Appendice A*, www.anvur.it
- Anvur (2017c), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Politiche e Sociali (Gev 14). Appendice B*, www.anvur.it
- Anvur (2017d), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Politiche e Sociali (Gev 14). Appendice C*, www.anvur.it
- Anvur (2017e), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Criteri per la valutazione dei prodotti di ricerca. Appendice A*, www.anvur.it
- Anvur (2017f), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area*

- Medica (Gev 6). Appendice*, www.anvur.it
- Anvur (2017g), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Agrarie e Veterinarie (Gev 7). Appendice* www.anvur.it
- Anvur (2017h), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze dell'antichità, filologico-letterarie e storico-artistiche (Gev 10). Appendice*, www.anvur.it
- Anvur (2017i), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Scienze Psicologiche (Gev 11b). Appendice*, www.anvur.it
- Anvur (2017l), *Valutazione della Qualità della Ricerca 2011-2014 (Vqr 2011-2014). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area Giuridica (Gev 12). Appendice C*, www.anvur.it
- Belleri D. (2014), *Semantic Under-determinacy and Communication*, Palgrave Macmillan, London.
- Bence V., Oppenheim C. (2005), "The evolution of the UK's Research Assessment Exercise: publications, performance and perceptions", *Journal of Educational Administration and History*, 37, 2: 137-155.
- Bentley R., Blackburn R. (1990), "Changes in Academic Research Performance Over Time: A Study of Institutional Accumulative Advantage", *Research in Higher Education*, 31,4: 327-353.
- Bonaccorsi A. (2012), *Potenzialità e limiti della analisi bibliometrica nelle aree umanistiche e sociali. Verso un programma di lavoro*, www.anvur.it
- Bornmann L., Daniel H. D. (2008), "What do citation counts measure? A review of studies on citing behaviour", *Journal of Documentation*, 64, 1: 45-80.
- Boudon R., Lazarsfeld P.F. (1965), *Introduction*, in Boudon R., Lazarsfeld P.F., eds, *Méthodes de la sociologie: I. Le vocabolaire des sciences sociales*, Mouton & Co., Paris (trad. it.: Cavazzani A.S., a cura di, *L'analisi empirica nelle scienze sociali. Volume I: Dai concetti agli indici empirici*, Il Mulino, Bologna, 1969).
- Bressan M. (2007), "Presentazione dei risultati dell'esercizio di valutazione della ricerca 2001-2005", Modena, 14 settembre.
- Campelli E. (2001), "Tohu va-vohu. Note non tecniche sul problema della classificazione", *Sociologia e Ricerca Sociale*, 64: 10-21.
- Campelli E. (2011), "A proposito di riviste. Valutazione e aritmetica, «protezionismo» e «liberismo»", *Sociologia e Ricerca Sociale*, 32, 95: 5-12.
- Cantril H. (1965), *The pattern of human concerns*, Rutgers University Press, New Brunswick.
- Chessa S., Vargiu A. (2011), "Valutazione universitaria e mutamenti istituzionali in Europa", *Studi di Sociologia*, 49: 3-34.
- Cipriani R. (2013), "È scoppiata la valutazione. Una proposta: il criterio della non prevalenza", *Sociologia e Ricerca Sociale*, 100: 11-16.
- Civr (2003a), *Valutazione triennale della ricerca. Bando di partecipazione all'esercizio 2001-2003*, www.civr.it
- Civr (2003b), *Linee guida per la valutazione della ricerca*, www.civr.it

- Civr (2006), *Vtr 2001-2003. Risultati delle valutazioni di panel di area*, www.civr.it
- Coats R., Bordon M., van Leeuwen T. N., van Raan A. (2009), "Scaling rules in the science system: influence of field specific citation characteristics on the impact of individual researcher", *Journal of the American Society for Information Science and Technology*, 60, 4: 740-753.
- Collins R. (1994), *Four Sociological Traditions*, Oxford University Press, Oxford (trad. it.: *Quattro tradizioni sociologiche: manuale introduttivo di storia della sociologia*, Zanichelli, Bologna, 1996).
- Corbetta P. (2014), *Metodologia e tecniche della ricerca sociale*, Seconda Edizione, Il Mulino, Bologna.
- Crane D. (1965), "Scientists at major and minor university. A study of productivity and recognition", *American Sociological Review*, 30, 5: 699-714.
- Crane D. (1967), "The gatekeepers of science. Some factors affecting the selection of articles in scientific journals", *American Sociologist*, 2, 4: 195-201.
- Decreto Ministeriale n. 458 del 27/06/2015, *Linee guida valutazione qualità della ricerca (VQR) 2011-2014*.
- Di Benedetto A. (2015), "Un'analisi del concetto di qualità della ricerca nella Vqr", *Sociologia e Ricerca Sociale*, 108: 95-112.
- Di Franco G. (1989), *Qualità della vita: dai modelli alle ricerche empiriche* in Vergati S., a cura di, *Dimensioni sociali e territoriali della qualità della vita*, La Goliardica, Roma.
- Enqa (2015), *Standards and guidelines for quality assurance in the European Higher Education Area*, Brussels, Belgium. <https://enqa.eu/>
- European Commission (2010), *Assessing Europe's University Based-Research*, <https://ec.europa.eu/info/research-and-innovation>
- Faggiano M.P. (2012), *Gli usi della tipologia nella ricerca sociale empirica*, FrancoAngeli, Milano.
- Fantoni S. (2015), "Il Sistema di valutazione Anvur", *Scuola Democratica*, 3: 695-703.
- Fasanella A., Di Benedetto A. (2014), "Luci e ombre nella Vqr 2004-2010. Un focus sulla scheda di valutazione peer nell'Area 14", *Sociologia e Ricerca Sociale*, 104: 59-84.
- Fasanella A., Di Benedetto A. (2015), "La valutazione dei pari nelle scienze sociali e politiche. La lezione della Vqr 2004-2010", *Sociologia e Politiche Sociali*, 18, 2: 44-72.
- Fasanella A., Martire F. (2017), "Considerazioni metodologiche sulla Vqr 2011-2014 e possibili sviluppi della valutazione", *Sociologia e Ricerca Sociale*, 114: 89-116.
- Fideli R. (2001), "La costruzione di un indice tipologico: criteri semantici, numerici ed empirici", *Sociologia e Ricerca Sociale*, 64: 124-137.
- Frabboni B., Sacchetta P. (2005), "Vtr 2001-2003: valutazione triennale della ricerca. Struttura del sistema informatico gestito via web", *Cineca Magazine*, 53: 5.
- Garfield E. (1979), "Is citation analysis a legitimate evaluation tool?", *Scientometrics*, 1, 4: 359-375.

- Geuna A., Martin B.R. (2003), "University Research Evaluation and Funding: an International Comparison", *Minerva*, 41: 277-304.
- Glänzel W. (2008), "Seven myths in bibliometrics: about facts and fiction in quantitative science studies", in Kretschmer, H., Havemann F. (eds., 2008), *Proceedings of WIS 2008*, Berlin, Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting.
- Harvey L. (2008), "Democratizing Quality", Budapest, 20 novembre 2008.
- Hceres (2014), *Criteria for the evaluation of research unit: the HCERES standards*, <https://www.hceres.fr/en>.
- Hempel C. G. (1952), *Fundamentals of Concept Formation in Empirical Science*, Chicago, University of Chicago Press (tr. it., *La formazione dei concetti e delle teorie nella scienza empirica*, Feltrinelli, Milano, 1961).
- Hicks D. (2011), "Performance based university research funding system", *Research Policy*, 41: 251-261.
- Hicks D., Wouters P., Waltman L., de Rijcke S., Rafols I. (2015), "Bibliometrics: The Leiden Manifesto for research metrics", *Nature*, 520, 7548: 429-431.
- H.M. Treasury (2006), *Science and Innovation Investment Framework 2004-2014: Next Steps*, www.hm-treasury.gov.uk
- Jimenez-Contreras (2010), *Como Utilizar los indicadores bibliometricos para la evaluacion de la actividad investigadora, la solicitud de sexenios y acreditacion profesores*, www.uc3.ugr.es
- Knaw, Vsnu and Now (2009), *Standard Evaluation Protocol 2009-2015: Protocol for Research Assessment in The Netherlands*, www.knaw.nl
- Kuhn T.S. (1969), *The Structure of Scientific Revolution*, Poscritto, The University of Chicago (trad. it.: Carugo A., a cura di, *La struttura delle rivoluzioni scientifiche*, Poscritto, Giulio Einaudi Editore, Torino, 2009).
- Lazarsfeld P.F., Barton A.H. (1951), *Qualitative Measurement in the Social Sciences: Classifications, Typologies and Indices*, in Lerner D. and Lasswell H.D., eds., *The Policy Science*, Stanford University Press, Stanford (trad. it.: *Classificazioni, tipologie e indici*, in Lazarsfeld P. F., *Metodologia e ricerca sociologica*, a cura di Capocchi V., il Mulino, Bologna, 1967).
- Lazarsfeld P.F. (1966), *Concept formation and measurement in the behavioral sciences: some historical observation*, in Di Renzo G. J., ed., *Concept, Theory, and Explanation in the behavioral sciences*, Random House, New York (trad. it., *Formazione e misurazione dei concetti nelle scienze del comportamento*, in Lombardo C., a cura di, *Saggi storici e metodologici*, Eucos, Roma, 2001).
- Lee C.J., Sugimoto C.R., Zhang G., Cronin B. (2013), "Bias in Peer Review", *Journal of the American Society for Information Science and Technology*, 66, 1: 2-17.
- Liani S., Martire F. (2017), *Pretest: Un approccio cognitivo*, FrancoAngeli, Milano.
- Mahoney, M.J. (1977), "Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System", *Cognitive therapy and research*, 1, 2: 161-175.
- Marradi A. (1984), *Concetti e metodi per la ricerca sociale*, La Giuntina, Firenze.
- Marradi A. (1990), "Fedeltà di un dato, affidabilità di una definizione operativa", *Rassegna Italiana di Sociologia*, 31, 1: 55-96.

- Marradi A. (1994), “Referenti, pensiero e linguaggio: una questione rilevante per gli indicatori”, *Sociologia e Ricerca Sociale*, 43: 137-207.
- Marradi A. (2007), *Metodologia delle scienze sociali*, Il Mulino, Bologna.
- Martini A., Cais G. (2000), *Controllo (di gestione) e valutazione (delle politiche): un (ennesimo ma non ultimo) tentativo di sistematizzazione concettuale*, in Palumbo M., a cura di, *Valutazione 2000. Esperienze e riflessioni, Primo Annuario dell’Associazione Italiana di Valutazione*, FrancoAngeli, Milano.
- Matarazzo F. (2018), “Il percorso politico e parlamentare della valutazione nelle università. Una storia lunga quarant’anni”, *Articolo* 33, 3: 17-48.
- Mattioli F., Anzera G., Toschi L. (2014), *Teoria e ricerca nell’analisi delle reti sociali*, Euroma, Roma.
- Mauceri S. (2003), *Per la qualità del dato nella ricerca sociale. Strategie di progettazione e conduzione dell’intervista con questionario*, FrancoAngeli, Milano.
- Mauceri S. (2008), “Ri-scoprire l’analisi dei casi devianti. Una strategia metodologica di supporto dei processi teorico-interpretativi nella ricerca sociale di tipo standard”, *Sociologia e Ricerca Sociale*, 87: 109-157.
- Merton, R.K. (1968a), “The Matthew Effect in Science”, *Science*, 159, 3810: 56-63.
- Merton, R.K. (1968b), *Social Theory and Social Structure*, Glencoe, Ill (trad. it.: *Teoria e struttura sociale*, Il Mulino, Bologna, 2000).
- Mesr (2007), *Les clés de la réforme des universités*, www.nouvelleuniversite.gouv.fr
- Mesr (2009), *L’état de l’enseignement supérieur et de la recherche en France – 29 indicateurs. Direction de l’évaluation, de la prospective et de la performance*, Imprimerie Moderne de l’Est.
- Minelli E., Reborà G., Turri M. (2008), *The Structure and Significance of the Italian Research Assessment Exercise (VTR)*, in Mazza C., Quattrone P. and Riccaboni A., eds., *European Universities in Transition. Issues, Models and Cases*. Cheltenham (UK), Edward Elgar.
- Morcellini M. (2015), “Per un’università sostenibile e moderna. Uno sguardo sociologico sulla valutazione e sull’Anvur”, *Rassegna Italiana di Valutazione*, 63: 68-82.
- Nobile S. (1997), *La credibilità dell’analisi del contenuto*, FrancoAngeli, Milano.
- Nobile S. (2008), *La chiusura del cerchio. La costruzione degli indici nella ricerca sociale*, Bonanno Editore, Acireale-Roma.
- Osservatorio per la valutazione del sistema universitario (1997), *Ruolo, organizzazione e attività dei Nuclei di valutazione interna delle università, relazione presentata all’Incontro nazionale sulla valutazione del sistema universitario*, 19 settembre 1997, Doc 5/97.
- Otley D. (2010), “Research Assessment in the UK: An Overview of 1992-2008”, *Australian Accounting Review*, 20: 3-13.
- Palumbo M. (2003), “La valutazione partecipata e i suoi esiti”, *Rassegna italiana di valutazione*, 7, 25: 71-88.
- Palumbo M., Torriggiani C., a cura di (2009), *La partecipazione fra ricerca e valutazione*, FrancoAngeli, Milano.

- Palumbo M., Pennisi C. (2011), “Le ragioni delle regole per la valutazione dell’università: per un’etica della pratica accademica”, *Studi di Sociologia*, 49: 35-50.
- Palumbo M., Pennisi C. (2014), “La valutazione senza governo”, *Rassegna Italiana di Valutazione*, 59: 7-33.
- Palumbo R. (2013), *La valutazione periodica della ricerca nelle discipline economico-aziendali*, FrancoAngeli, Milano.
- Penfield T., Baker M. J., Scoble R., Wykes M. C. (2014), “Assessment, evaluations, and definitions of research impact: A review”, *Research evaluation*, 23, 1: 21-32.
- Perotti R. (2008), *L’Università truccata*, Einaudi, Torino.
- Pित्रone M.C. (2009), *Sondaggi e interviste. Lo studio dell’opinione pubblica nella ricerca sociale*, FrancoAngeli, Milano.
- Pित्रone M.C., Pavsic R. (2003), *Come conoscere opinioni e atteggiamenti*, Bonanno, Catania.
- Polanyi M. (1966), *The Tacit Dimension*, New York, Anchor Books (tr. it. *La conoscenza inespresa*, Armando, Roma, 1979)
- Poli S. (2008), *Lo studio degli atteggiamenti nella ricerca sociale: dalle definizioni alle tecniche*, in Bichi R., a cura di, *La distanza sociale. Vecchie e nuove scale di misurazione*, FrancoAngeli, Milano.
- Reale E. (2013a), *La valutazione della ricerca pubblica. Un’analisi della valutazione triennale della ricerca*, FrancoAngeli, Milano.
- Reale E. (2013b), “La valutazione della ricerca e il cambiamento dell’Università”, *Sociologia e Ricerca Sociale*, 100: 148-159.
- Reale E., Pennisi C. (2012), “Valutare nella crisi: effetti sull’Università e la ricerca”, *Rassegna Italiana di Valutazione*, 14: 7-14.
- Realfonzo R., Perone G. (2016), *Qualità degli Atenei e contesto socio-economico. La sperequazione nell’allocazione delle risorse tra le Università italiane*, in Ragozini G., a cura di, *Lo scenario universitario. Mercato del lavoro e sbocchi occupazionali*, FrancoAngeli, Milano.
- Rebora G. (2010), *Tra inferno e paradiso. Gli atenei italiani alla prova della valutazione*, Scripta Web, Napoli.
- Rebora G. (2013), *Nessuno mi può giudicare: l’università e la valutazione*, Guerini, Milano.
- Rebora G., Turri M. (2010), “Lo sviluppo dei sistemi di valutazione della ricerca: un’analisi critica dell’esperienza italiana”, Paper presentato al Convegno AID-DEA 2010 “Pubblico & non profit per un mercato responsabile e solidale” tenutosi dal 21/10 al 22/10/2010 presso l’Università Bocconi di Milano.
- REF (2010), *REF, Impact Pilot Exercise: Findings of the Expert Panels*, <http://www.ref.ac.uk>
- REF (2019), *Panel criteria and working methods*, <https://www.ref.ac.uk/>
- Ribolzi L. (2013), “Valutare l’università: una sfida non solo per l’Anvur”, *Sociologia e ricerca sociale*, 100: 23-32.
- Rip A., van der Meulen B.J.R. (1995), “The Patchwork of the Dutch Evaluation System”, *Research Evaluation*, 5: 45–53.
- Rizzi D., Silvestri P. (2002), “La valutazione del sistema universitario italiano: una

- storia recente”, *Nota di lavoro dell'Università Ca' Foscari di Venezia*, 1: 1-23.
- Scarpitti L. (2001), *La valutazione nel sistema universitario italiano*, in Stame N., a cura di, *Valutazione 2001. Lo sviluppo della valutazione in Italia*, FrancoAngeli, Milano.
- Schroter S., Black N., Evans S., Carpenter J., Godlee F., Smith R. (2004), “Effects of Training on Quality of Peer Review: Randomised Controlled Trial”, *British Medical Journal*, 328, 7441: 673.
- Schmitz C. (2008), *Messung Der Forschungsleistung in der Betriebswirtschaftslehre auf Basis der ISI-Zitationsindizes. Eine kritische Analyse anhand konzeptioneller Überlegungen und empirischer Befunde*, Eul Verlag.
- Searle J.R. (1979), *Expression and Meaning. Studies in the Theory of Speech Acts*, Cambridge, University Press.
- Searle J.R. (1980), *The Background of Meaning*, in Searle J.R., Kiefer F., Bierwisch M., eds., *Speech Act Theory and Pragmatic*, Reidel Publishing Company, Dordrecht.
- Seglen P.O. (1997), “Citations and journal impact factors: questionable indicators of research quality”, *Allergy*, 52, 11: 1050-1056.
- Stame N. (2016), *Valutazione pluralista*, FrancoAngeli, Milano.
- Travis C. (1975), *Saying and Understanding. A Generative Theory of Illocutions*, Basil Blackwell, Oxford.
- Travis C. (1981), *The True and False. The Domain of the Pragmatic*, John Benjamins Publishing Company, Amsterdam-Philadelphia.
- Turri M. (2012), “Linee di evoluzione della valutazione nei sistemi universitari europei”, *Liuc Papers 259, Serie Economia e Impresa*, 67: 1-15.
- Tusini S. (2006), *La ricerca come relazione: l'intervista nelle scienze sociali*, FrancoAngeli, Milano.
- Valentini E. (2013), “Ritorno al passato? Il cortocircuito riforme/valutazione nel campo delle scienze umanistiche e politiche-sociali”, *Sociologia e Ricerca Sociale*, 100: 72-90.
- Wennerås, C., Wold, A. (1999), “Nepotism and Sexism in Peer Review”, *Nature*, 387, 6631: 341-343.
- Wright Mills C. (1959), *The Sociological Imagination*, Oxford University Press, Oxford (trad. it.: *L'immaginazione sociologica*, Il Saggiatore, Milano, 1962).
- Xie Y. (2014), “«Undemocracy»: Inequalities in Science”, *Science*, 344, 6186: 809-810.

Curatori e autori

Antonio Fasanella è professore ordinario di Storia e metodo delle scienze sociali, Metodologia della ricerca sociale, Teorie e pratiche della valutazione presso il Dipartimento di Comunicazione e ricerca sociale della Sapienza Università di Roma. Ha pubblicato numerosi saggi, articoli e volumi su temi metodologici e di valutazione sociale.

Fabrizio Martire, professore associato presso il Dipartimento di Comunicazione e ricerca sociale dell'Università di Roma La Sapienza, insegna materie metodologiche e si occupa di temi connessi alla valutazione, alla metodologia della ricerca sociale e alla storia della sociologia.

Lorenzo Barbanera è dottore di ricerca in Comunicazione, ricerca sociale e marketing – con curriculum in Metodologia delle scienze sociali – presso l'Università La Sapienza di Roma. Ha inoltre collaborato con l'istituto Censis, nella divisione Economia e Territorio. Fra i suoi interessi di ricerca vi sono la valutazione delle politiche formative in ambito universitario e la realizzazione di strumenti di rilevazione per indagini standard.

Federica Floridi è dottore di ricerca in Comunicazione, ricerca, innovazione, curriculum in Metodologia delle scienze sociali, presso Sapienza Università di Roma. Collabora con enti pubblici e privati di valutazione ed è stata responsabile per gli aspetti tecnico-metodologici di ricerche valutative nell'ambito di progetti e programmi della coo-

perazione internazionale del Ministero degli Affari esteri. I suoi principali interessi di ricerca sono la valutazione delle politiche pubbliche e sociali, i processi di scolarizzazione, la valutazione dell'università e della ricerca.

Federica Fusillo ha conseguito il dottorato di ricerca in Comunicazione, ricerca, innovazione, curriculum in Metodologia delle scienze sociali, presso Sapienza Università di Roma; ha conseguito la laurea magistrale presso lo stesso Ateneo con una tesi sull'Abilitazione Scientifica Nazionale. L'interesse verso la valutazione delle politiche formative e, in particolare, del sistema universitario italiano è proseguito nel corso degli anni con la partecipazione a indagini e progetti di ricerca sulla valutazione della qualità della ricerca, sulla produttività scientifica dei docenti universitari, sulle carriere accademiche e sui criteri di valutazione Asn.

Marco Palmieri è attualmente assegnista di ricerca per il Dipartimento di Studi politici e sociali dell'Università di Salerno. È docente a contratto di Metodologia della ricerca sociale presso il Dipartimento di Scienze umane dell'Università dell'Aquila. È anche docente al Master interuniversitario "Sociologia: Teoria, Metodologia, Ricerca" (So.Te.Me.Ri.), organizzato da Università Sapienza, Università di Tor Vergata e Università di Roma 3. Tra i suoi principali interessi di ricerca vi è la costruzione di strumenti standardizzati per lo studio di atteggiamenti, opinioni e valori.

A cura di Antonio Fasanella e Fabrizio Martire

Valutazione della ricerca e ricerca sulla valutazione

Il volume propone una riflessione critica sul tema della valutazione della qualità della ricerca condotta entro una prospettiva orientata in senso progressivo. Non si intende negare l'utilità o, se si vuole, la necessità di sottoporre a valutazione gli esiti della ricerca, ma si ribadisce altresì l'indispensabilità di un controllo pubblico dei processi di valutazione, perseguendo inflessibilmente un principio di *accountability*, al fine di superare la falsa contrapposizione tra approcci *judgemental* e approcci standardizzati alla valutazione.

Se l'unica scelta praticabile è tra una valutazione basata esclusivamente su conoscenza tacita non ricostruibile, su risorse squisitamente soggettive come l'intuizione, e una valutazione ispezionabile/replicabile perché basata su indicatori quantificabili, la tentazione della bibliometria e dell'uso incondizionato dei ranking delle riviste diventa irresistibile. Tutta la trattazione qui svolta sostiene fermamente la possibilità e l'opportunità di una soluzione alternativa: un sistema di valutazione incentrato sulla *peer review* e che rispetti, in ogni suo passaggio, il principio dell'*accountability*. Il volume prende atto positivamente degli sforzi che sono stati condotti in sede Vqr in questo senso, ma non può non riconoscere, d'altra parte, che l'attuale Vqr, nonostante tali sforzi, è ancora contraddistinta da elementi di opacità, di ambiguità, di arbitrio, di scarsa ispezionabilità delle procedure, che la rendono suscettibile di essere perfezionata, anche sulla base delle proposte qui avanzate.

Antonio Fasanella è professore ordinario di Storia e metodo delle scienze sociali, Metodologia della ricerca sociale, Teorie e pratiche della valutazione presso il Dipartimento di Comunicazione e ricerca sociale della Sapienza Università di Roma. Ha pubblicato numerosi saggi, articoli e volumi su temi metodologici e di valutazione sociale.

Fabrizio Martire, professore associato presso il Dipartimento di Comunicazione e ricerca sociale della Sapienza Università di Roma, insegna materie metodologiche e si occupa di temi connessi alla valutazione, alla metodologia della ricerca sociale e alla storia della sociologia.