

# VQR 2015-2019: criticità e proposte per le aree non-bibliometriche

*Commissione Presidio alla Qualità, Area Ricerca e Terza Missione  
del Dipartimento Economia e Management, Università di Ferrara<sup>1</sup>  
(approvato 31 marzo 2019)*

*Consiglio di Dipartimento Economia e Management, Università di Ferrara  
(approvato 7 maggio 2019)*

## **PREMESSA**

Il documento approvato dal Consiglio del Dipartimento di Economia e Management del 10 Settembre 2018 ha proposto una riflessione sull'adeguatezza dell'attuale sistema di Valutazione della Qualità della Ricerca. Quel documento fu il prodotto di un percorso di discussione, realizzato attraverso incontri e seminari con esperti del settore, e occasioni di dibattito quali un'assemblea di Dipartimento e alcuni Consigli di Dipartimento.

Il documento stabilisce, tra l'altro, la necessità di una proposta alternativa all'attuale sistema VQR, guidata da principi generali finalizzati ad evitare le distorsioni generate dall'attuale sistema, e a creare migliori condizioni per il raggiungimento del più importante obiettivo di un qualsiasi sistema di valutazione, che è quello di orientare i comportamenti verso direzioni e finalità desiderate. L'obiettivo specifico è migliorare la qualità della produzione scientifica, declinabile in (i) contenuti innovativi, sia incrementali sia radicali, e (ii) elevata rilevanza, in termini di impatti e diffusione della ricerca, (iii) dialogo tra approcci e paradigmi, tra loro anche alternativi, valorizzando l'approccio collaborativo alla ricerca e premiando l'ottica inter/multi/trans-disciplinare.

## **OBIETTIVI**

Il documento propone alcuni elementi 'correttivi', che mirano concretamente a realizzare quanto auspicato durante il percorso di riflessione sopra descritto. Anche se tutto il sistema VQR richiederebbe una profonda revisione, in questa sede ci si focalizza principalmente sull'area non-bibliometrica, con un riferimento specifico all'area 13. Nel seguito elenchiamo e commentiamo brevemente gli obiettivi perseguiti, che non sono proposti secondo un particolare ordine di importanza. Il documento si colloca all'interno del dibattito attuale sulla valutazione della ricerca, i suoi obiettivi, e possibili riforme del sistema ANVUR/VQR<sup>2</sup>.

### **1. Semplificare le procedure del processo valutativo**

- Lo scopo è modificare un esercizio che, allo stato attuale, assorbe una quantità di risorse significative, sia economiche che cognitive. Si ritiene sia possibile ottenere una maggiore efficacia del processo valutativo senza necessariamente aumentarne la complessità.

---

<sup>1</sup> Il presente documento è stato elaborato nel marzo 2019 dalla Commissione Presidio alla Qualità del DEM, Area Ricerca e Terza Missione (componenti i professori Enrico Bracci, Giovanni Masino, Massimiliano Mazzanti, Antonio Musolesi, Paolo Pini) e presentato nel Consiglio di Dipartimento in data 2 aprile 2019. Approvato dal Consiglio di Dipartimento in data 7 maggio 2019.

<sup>2</sup> Si veda, ad esempio, il documento CUN sulla valutazione della ricerca, e in particolare sulla VQR, approvato nella seduta del 21 marzo 2019, e recenti.

## 2. Mettere in atto metodologie di auto-valutazione che evitino o quanto meno riducano le documentate distorsioni dell'attuale sistema

- Le distorsioni di cui si parla, documentate in diverse pubblicazioni attraverso specifiche indagini, che sono ormai parte del vissuto quotidiano di molti (se non tutti) accademici, riguardano numerosi elementi. In particolare:
  - i. Il sistema attuale non sembra in grado di valutare con pertinenza la qualità della produzione scientifica, in quanto si limita ad osservare la collocazione editoriale delle pubblicazioni anziché valutarne il contenuto, peraltro utilizzando metodi di classificazione delle collocazioni discutibili e discriminando apertamente alcune modalità di pubblicazione.
  - ii. Il sistema attuale focalizza *sostanzialmente* l'oggetto della valutazione, e l'uso dei risultati, sul singolo individuo, trascurando (anzi, spesso penalizzando, direttamente o indirettamente) i processi di ricerca collaborativa, che sono spesso forieri di nuova conoscenza di frontiera e innovativa in senso radicale. È oggi carente l'attenzione all'aspetto progettuale ex ante, che dovrebbe essere uno dei principali obiettivi, dal punto di vista dell'orientamento dei comportamenti, del processo valutativo: la costruzione di progetti di ricerca dipartimentali e tra dipartimenti, valutati sulla base di una varietà di indicatori (originalità del progetto, CV del team, pubblicazioni individuali, impatti, etc..). Si nota come, pur con una enfasi su aspetti di 'competizione', che non dovrebbero caratterizzare primariamente i processi di ricerca<sup>3</sup>, il bando 'Dipartimenti di Eccellenza' mostrava la positiva peculiarità di orientare i Dipartimenti alla creazione di progettualità innovative, connotate da aspetti legati alla ricerca, didattica, terza missione, e loro integrazione, per lo sviluppo ulteriore delle strutture<sup>4</sup>.
  - iii. Il sistema attuale trascura completamente, nel processo valutativo, le condizioni di contesto nel quale gli studiosi si trovano ad operare. Il che significa, nel concreto, che vengono trascurati la qualità delle condizioni organizzative (intese nel senso più ampio) che le strutture e i soggetti istituzionali (dipartimenti, centri di ricerca e, in ultima analisi, gli Atenei) sono in grado di realizzare al fine di favorire una ricerca di qualità. Valutare la qualità della ricerca dovrebbe significare, anzitutto, valutare le condizioni di contesto che consentono ai ricercatori di sviluppare ricerca, appunto, di qualità.
  - iv. Il sistema attuale genera sempre più incentivi verso orientamenti di ricerca di tipo incrementale, in vari modi. Da un lato, premiando soprattutto la quantità, piuttosto che la qualità. Questo spinge i ricercatori verso ricerche di respiro meno ampio, che generano produzioni e pubblicazioni più frequenti, e che implicano un minor sforzo progettuale e concettuale. D'altro lato, premiando primariamente la ricerca disciplinare, che è certamente molto importante e non va in alcun modo trascurata. Al tempo stesso, la

---

<sup>3</sup> Si veda, tra le altre, la recente proposta dell'Università di Ghent: "Ghent University is deliberately choosing to step out of the rat race between individuals, departments and universities. We no longer wish to participate in the ranking of people" (<https://www.ugent.be/en/news-events/ghent-university-talent-rat-race-transformation-career-evaluation-model.htm>).

<sup>4</sup> Si ritiene utile proseguire, migliorando il quadro di riferimento normativo, su esperienze di valutazione di progetti. A complemento, si nota come andrebbe riportato il PRIN al centro del sistema, garantendo finanziamenti adeguati e soprattutto certezza e continuità dei bandi (elementi assenti da più di un decennio).

penalizzazione della ricerca interdisciplinare, che può portare a progressi di conoscenza molto innovativi<sup>5</sup>, è un fatto evidente. Il sistema attuale non spinge i ricercatori ad “osare”, a spingersi oltre la frontiera esistente, ma a puntare su ciò che sta ampiamente dentro i confini di ciò che è già consolidato. In aggiunta, la ricerca inter/multi disciplinare è rilevante sui livelli di progettazione, soprattutto internazionale. Il sistema VQR attuale genera effetti negativi sulle possibilità di costruzione di proposte di ricerca sul livello nazionale ed internazionale.

- v. Infine, ignorando o addirittura penalizzando il *processo* di ricerca in quanto tale, focalizzandosi invece soltanto sugli output. La ricerca (specialmente la ricerca di frontiera e di qualità) è un processo intrinsecamente e altamente incerto, e una valutazione che miri a generare comportamenti virtuosi dovrebbe tenerne conto; dovrebbe premiare, per esempio, anche la progettualità di qualità, cioè gli sforzi (purché genuini, competenti e ben progettati) orientati a produrre ricerca innovativa, e dovrebbe farlo in un modo che sia almeno in parte indipendente dei loro esiti finali. Tutte le organizzazioni più innovative sono quelle che non ostacolano il coraggio di osare, ma lo premiano, anche quando gli esiti non sono quelli sperati. Occorre quindi anche ‘ripensare’ il significato di fallimento, de-enfatizzare gli elementi negativi, analizzando le motivazioni del fallimento del ricercatore/innovatore, e concentrando sforzi e finanziamenti anche su questa criticità. Penalizzare solamente il fallimento, senza comprenderne la cause, conduce a esacerbare i disincentivi all’innovazione radicale.

Ciascuno dei punti critici sopra elencati (ed altri che non citiamo per brevità) evidenziano un sistema che genera comportamenti contrari alla diffusione di una ricerca che sia intrinsecamente (e non solo simbolicamente, retoricamente, talvolta persino fittiziamente) di qualità. Questo è particolarmente dannoso se si considerano i giovani ricercatori, i quali, per ovvie e legittime ragioni legate alle ambizioni di carriera, sono particolarmente sensibili e attenti a conformarsi a ciò che il sistema legittima e premia. Sono distorsioni gravissime, che pregiudicano il presente e, ancor più, il futuro della ricerca italiana. Per questo, occorre adottare un sistema di auto-valutazione che cerchi di evitare o ridurre queste distorsioni, e che produca effettivamente gli esiti che dichiara di voler perseguire, cioè di orientare i comportamenti verso la produzione di ricerca altamente qualitativa, nel rispetto dei principi irrinunciabili (ma, di fatto, anch’essi gravemente incrinati dal sistema attuale) della libertà di ricerca individuale e della pari dignità dei diversi approcci concettuali, metodologici ed epistemologici.

### **3. Sostituire procedure valutative ibride che creano confusione tra le aree bibliometriche e non bibliometriche**

Vi sono differenze significative tra le varie aree, discipline, tradizioni, approcci metodologici ed epistemologici. Un buon sistema di valutazione dovrebbe riconoscere la ricchezza di tale varietà, ed incoraggiarla attraverso appropriate modalità valutative. La distinzione più semplice è quella tra aree bibliometriche e aree non bibliometriche. Non è certamente la distinzione più fine che si possa immaginare. Al contrario, è piuttosto grossolana, ma per lo meno fornisce un punto di partenza per poter affermare la necessità di metodologie valutative che siano il più possibile specifiche e contestualizzate rispetto al tipo di ricerca che si intende valutare. Dunque, l’ibridazione tra aree bibliometriche e aree non-bibliometriche è un ottimo esempio di ciò che si dovrebbe evitare. Ciò che qui si propone è l’esatto contrario, cioè

---

<sup>5</sup> Un’enfasi totalmente disciplinare caratterizza la VQR, l’esperienza dei Dipartimenti di Eccellenza, l’assetto delle Lauree magistrali. L’unico spazio di sviluppi progettuali non solamente disciplinari è il Dottorato di ricerca.

distinguere nettamente l'approccio valutativo quanto meno tra queste due grandi aree, nella consapevolezza che distinzioni anche più fini, che tengano conto di ulteriori elementi di specificità tra le diverse discipline, sarebbero necessarie. La proposta più semplice, al momento, è quella di utilizzare il metodo della *peer review* per le aree non bibliometriche. Nel seguito del documento verranno proposte modalità e soluzioni per ridurre o evitare i problemi che si possono prefigurare in relazione all'uso di questo metodo, e al tempo stesso coglierne i numerosi vantaggi.

#### **4. Riservare ai GEV di area scientifica disciplinare l'implementazione e il monitoraggio delle procedure, e quindi l'analisi della auto-valutazione compiuta dalla comunità scientifica**

Questo punto è un corollario importante del punto precedente. Se è vero che esistono specificità disciplinari che vanno salvaguardate (perché questo arricchisce il processo di ricerca complessivo e lo rende più fertile) e rispetto alle quali le metodologie di valutazione devono adeguarsi (e non il contrario: cioè non devono essere le discipline ad adeguarsi a una metodologia di valutazione monolitica, uniformante e omogeneizzatrice), allora occorre che siano i GEV di area scientifica-disciplinare a rendersi 'protagonisti' del processo valutativo. Peraltro, questo deve essere fatto in un modo che al tempo stesso sia cogente rispetto alle specificità delle discipline, ma anche aperto e, anzi, favorevole agli incroci interdisciplinari. Dunque, servirebbe un continuo dialogo, caratterizzato anche da momenti pubblici di discussione, da questo punto di vista, tra i diversi GEV.

#### **5. Distinguere nettamente la valutazione delle strutture (Atenei, Dipartimenti, enti di ricerca), dalla valutazione dei singoli individui**

Questo punto deriva da una riflessione sulle debolezze del sistema attuale che abbiamo già menzionato nel punto 1, ma che è cruciale e dunque vale la pena ribadire. Un sistema che miri a favorire la crescita della qualità della propria produzione scientifica deve occuparsi non solo della qualità della produzione a livello individuale, ma anche (soprattutto) delle condizioni contestuali (organizzazione, risorse, policy, cultura, etc..) nelle quali gli individui operano, in quanto tali condizioni possono influenzare in modo molto significativo, talvolta decisivo, i comportamenti dei soggetti protagonisti della ricerca (individui e gruppi). La ovvia conseguenza di ciò è che i due processi valutativi (di individui e strutture) dovrebbero essere chiaramente distinti, anche nel modo in cui i loro esiti sono utilizzati, e considerati ugualmente importanti, in relazione, tuttavia, a obiettivi totalmente distinti. Da un lato, gli individui vengono già valutati in numerosi momenti della loro carriera (ASN, reclutamento, progressioni di carriera, componenti premiali della retribuzione, etc). Lo scopo di tutti questi passaggi (per l'appunto, valutativi) è fare in modo che gli individui si dotino di capacità e conoscenze necessarie per acquisire la professionalità richiesta nelle attività di ricerca. D'altro lato, la VQR dovrebbe focalizzare pienamente l'attenzione sulle strutture, per valutare la loro capacità di creare le condizioni affinché i ricercatori al loro interno siano in grado di produrre ricerca di qualità. La qualità della ricerca dipende infatti da due elementi: le competenze e la professionalità degli individui, e l'adeguatezza dei contesti nei quali gli individui agiscono. Occorrono dunque due processi valutativi distinti, diversi per modalità e finalità. Questo si ricollega all'esigenza di legare maggiormente il processo valutativo al suo fine primario, la valutazione di progettualità.

**6. Riaffermare che se la valutazione della qualità della ricerca è delegata all'Anvur, la distribuzione delle risorse anche nelle componenti premiali dell'FFO è effettuata al MIUR**

Considerata la nota anomalia dell'ANVUR, che non è una vera agenzia di valutazione indipendente, occorre chiarire il perimetro di competenza ANVUR e MIUR all'interno del sistema di valutazione e allocazione delle risorse. Lo status attuale dell'Agenzia appare come un "ibrido" tra agenzia di valutazione indipendente ed autonoma, ed agenzia che valuta, direziona, prescrive comportamenti, ed ancora agenzia a supporto del decisore politico ma anche sotto il controllo del decisore politico (MIUR). Questo ruolo ha portato l'Agenzia a svolgere funzioni che non sono sue proprie, supplendo o sostituendo il decisore politico nelle sue funzioni. Separare chiaramente questi ruoli e funzioni appare urgente, come indicano anche recenti documenti ministeriali.

**7. L'esercizio di valutazione ha l'obiettivo di monitorare la qualità della produzione scientifica, segnalare fattori di forza e fattori di debolezza, individuare criticità entro le aree scientifiche, suggerire opzioni per ridurre tale criticità; deve dunque essere improntato alla analisi e non alla prescrizione**

Questo punto sottolinea lo spirito generale che dovrebbe guidare un processo valutativo, in un contesto ad elevata incertezza e ad elevata utilità pubblica quale la ricerca. Dovrebbe essere dunque uno strumento di supporto e di analisi, non un modo per creare vincoli e obblighi (anche se impliciti o informali), non per imporre modalità di ricerca o, ancor peggio, per predeterminare il valore della produzione scientifica a prescindere dai suoi contenuti. Queste sono caratteristiche che invece, in modo diretto e indiretto, connotano il sistema attuale, e che dunque vanno superate.

## **PROBLEMATICHE E SOLUZIONI PER LE AREE NON-BIBLIOMETRICHE**

In questa sezione, proponiamo un approfondimento specifico sulle aree non-bibliometriche. Due principi generali, già sottolineati sopra, vanno ribaditi. Da un lato, sono le strutture ad essere oggetto di valutazione, e non gli individui (nella forma e nella sostanza). D'altro lato, è necessario fare in modo che vi sia una logica generale comune entro le aree non-bibliometriche sia al fine di valorizzare quanto meno la specificità di queste aree rispetto alle bibliometriche, sia di facilitare il dialogo interdisciplinare. Le specificità disciplinari sono salvaguardate nel processo valutativo di ogni prodotto di ricerca in base ai suoi contenuti. Tuttavia, se vi fossero parametri valutativi radicalmente diversi tra discipline appartenenti alla stessa area non-bibliometrica, si creerebbero notevoli difficoltà di dialogo tra le aree e non si incoraggierebbero le collaborazioni di ricerca tra discipline diverse.

Alla luce di quanto affermato sinora, e nella consapevolezza che il miglioramento dell'attuale sistema richiederebbe un processo di trasformazione molto radicale, proponiamo le seguenti soluzioni che, in modo relativamente semplice e fattibile da un punto di vista operativo, e tenendo anche conto dell'esigenza di riduzione della complessità, possano condurre verso le direzioni che sono state delineate nei paragrafi precedenti.

### **1. Numero delle pubblicazioni / prodotti di ricerca da valutare**

In primis, la riduzione del numero delle pubblicazioni è volta a ridurre la complessità del processo. Da un lato, aiuterebbe a focalizzare l'attenzione sulla qualità della ricerca anziché su aspetti meramente quantitativi. Dall'altro, aiuterebbe a contenere i costi dell'esercizio di valutazione. Infine, aiuterebbe anche al miglioramento della pertinenza (dunque l'efficacia) delle valutazioni: quando il numero di prodotti da valutare è ridotto, i valutatori possono dedicare

più attenzione e tempo ad ogni singolo prodotto da valutare. L'attività di *peer review* è usuale e quotidiano elemento di partecipazione del ricercatore allo sviluppo della conoscenza.

Lo schema VQR attuale prevede che per ogni struttura (ad esempio i dipartimenti) vengano individuati per la valutazione due prodotti di ricerca da parte di ogni ricercatore afferente.

Il cambiamento proposto è il seguente:

- il numero di prodotti da valutare coincide con il numero di strutturati in servizio in ogni struttura
- le posizioni da considerare come strutturati sono RTDA, RTDB, PA, PO, in servizio alla data del bando di avvio dell'esercizio di valutazione<sup>6</sup>

Esempio: Se un dipartimento è composto da 50 afferenti strutturati, in totale deve selezionare 50 prodotti

## **2. Selezione delle pubblicazioni / prodotti di ricerca**

Lo schema attuale prevede che in ogni struttura dipartimentale ai singoli strutturati si associno 2 prodotti di ricerca, da inviare per la valutazione. Questo approccio focalizza il processo valutativo sull'individuo. Per le ragioni già espresse, l'unità di analisi dovrebbe invece essere – non solo formalmente ma anche sostanzialmente - la struttura, e non l'individuo. E' bene evidenziare la differenza tra focus formale e sostanziale. E' vero che, formalmente, già oggi l'unità di analisi della valutazione sono le strutture. Ma, sostanzialmente (cioè se si considera il metodo utilizzato, in cui l'unità di osservazione è l'individuo, e i vari modi in cui sono usati i risultati della valutazione sono utilizzati, per esempio nel caso dei Collegi di Dottorato, almeno fino al 2018) il focus valutativo è chiaramente, di fatto, l'individuo.

Il cambiamento proposto è il seguente:

- Una volta definito il numero di prodotti di ricerca che ogni struttura deve sottoporre alla valutazione in base al numero degli strutturati (vedi sopra)
- Sono le strutture a selezionare i prodotti da inviare scegliendoli come più adatti tra tutti quelli a loro disposizione
- Ciò deve avvenire indipendentemente dal fatto che siano stati realizzati da uno o dall'altro ricercatore

Ad esempio: se un Dipartimento con 50 afferenti strutturati deve selezionare 50 prodotti, la scelta è fatta dal Dipartimento secondo criteri che saranno determinati dal Dipartimento stesso (criteri che potrebbero riguardare, per esempio, la rappresentatività dei prodotti selezionati rispetto alle aree di ricerca presenti nella struttura, il collegamento dei prodotti selezionati con iniziative o progetti di ricerca esistenti, o altri criteri che, liberamente, la struttura ritiene di utilizzare).

Ribadiamo quanto già affermato sopra nel punto 5 del paragrafo relativo agli obiettivi: gli individui vengono valutati in altri momenti della loro progressione di carriera (ASN, reclutamento, progressioni di carriera, componenti premiali della retribuzione, etc.). L'esercizio VQR non ha nulla a che fare con questo, e dovrebbe invece essere uno

---

<sup>6</sup> Pur non potendo considerare l'assegnista di ricerca uno strutturato, si nota a complemento come andrebbe risolto il 'problema italiano della figura giuridica dell'assegnista. Gli assegnisti supportano ampiamente il sistema didattico e di ricerca, e spesso se non sempre non hanno accesso formale e riconoscimento a possibilità di finanziamento e ruoli istituzionali.

strumento di analisi della capacità delle strutture di creare le condizioni affinché i ricercatori afferenti siano in grado di produrre ricerca di qualità.

Un corollario di tutto ciò è che la eventuale criticità rappresentata dagli “strutturati che non pubblicano” non deve rientrare entro la valutazione delle strutture. Sono altre le procedure che devono affrontare questa problematica, a livello decentrato nei singoli Atenei, e non rinviata all’esercizio VQR, perché occorre mantenere nettamente distinta la valutazione delle strutture da quella dei singoli individui.

Tuttavia, per evitare usi non corretti del principio generale che ispira questa proposta, si suggerisce un correttivo, per cui i Dipartimenti debbano selezionare i prodotti da valutare coinvolgendo almeno il 50% dei propri strutturati. In questo modo, si evita il possibile verificarsi di casi limite, chiaramente indesiderabili, in cui solo una piccola parte degli strutturati sia coinvolta sistematicamente nel processo.

### **3. Individuazione dei settori tematici di ricerca**

La singole strutture definiscono per ciascuna pubblicazione / prodotto di ricerca specifiche aree di ricerca, avvalendosi della classificazione settori ERC (European Research Council) al secondo livello di dettaglio. Nel settore delle scienze sociali e umanistiche, vi sono 6 macrosettori di primo livello, e 71 settori di secondo livello. Ogni pubblicazione / prodotto di ricerca deve essere associato a 3 distinti settori ERC di secondo livello, in ordine di rilevanza.

Questa scelta ha ricadute nella fase di valutazione peer-review dei singoli prodotti (come indicato successivamente) e consente una classificazione e gestione della eventuale natura inter-multi-trans disciplinare, in coerenza con le principali linee progettuali europee e nazionali. L’attuale assetto della VQR disincentiva la ‘curiosità’ e collaborazione fra discipline, data l’assenza di valutazione positiva, ad esempio, nel caso di pubblicazioni di fascia A in aree disciplinari diverse da quelle di appartenenza.

### **4. Tipologie di pubblicazioni / prodotti di ricerca da selezionare**

Nelle aree non-bibliometriche, sono individuate 3 tipologie di base delle pubblicazioni /prodotti di ricerca da valutare:

- articoli pubblicati su rivista che possiede la classificazione ISSN (*International Standard Serial Number*)
- saggi pubblicati su un volume collettaneo che possiede classificazione ISBN (*International Standard Book Number*)
- monografie pubblicate come volume che possiede classificazione ISBN (*International Standard Book Number*)

A queste tre tipologie di base, si possono aggiungere “altre” specifiche tipologie, in base alle esigenze delle aree disciplinari. Queste altre eventuali tipologie vengono individuate dai corrispondenti GEV di area, consultando le associazioni scientifiche.

### **5. Aree disciplinari dei singoli dipartimenti e loro composizione in base alle aree disciplinari degli strutturati**

Ogni singolo Dipartimento farà riferimento prioritariamente ad una singola area scientifica disciplinare, come definite dalla seguente consueta classificazione:

- Area 10 - Scienze dell’antichità, filologico-letterarie e storico-artistiche
- Area 11a – Scienze storiche, filosofiche e pedagogiche
- Area 11b – Scienze psicologiche
- Area 12 – Scienze giuridiche

- Area 13 – Scienze economiche e statistiche
- Area 14 – Scienze politiche e sociali

I corrispondenti GEV saranno quelli di riferimento nel processo di valutazione.

Per i dipartimenti che vedono, tra i loro strutturati, studiosi afferenti a diverse aree scientifiche disciplinari, per i quali la struttura ha selezionato le pubblicazioni / prodotti di ricerca da valutare, deve valere il principio che la selezione dei prodotti sia congrua con quanto stabilito nel punto 4 precedente in termini di tipologia di pubblicazione / prodotto da valutare, senza operare alcuna forzatura influenzata dal riferimento del dipartimento ad una specifica area scientifica disciplinare e quindi al corrispondenti GEV.

Al contempo, per ogni pubblicazione / prodotto di ricerca selezionato, si applica il principio stabilito nel punto 3 precedente, ovvero il fatto che l'individuazione dei settori ERC deve essere effettuata in base alle caratteristiche della pubblicazione / prodotto di ricerca, senza operare alcuna forzatura influenzata dal riferimento del dipartimento ad una specifica area scientifica disciplinare e quindi al corrispondente GEV.

## **6. Valutazione delle pubblicazioni / prodotti di ricerca via peer-review**

Nell'esercizio di valutazione per le aree non-bibliometriche, è necessario abbandonare totalmente l'impiego della bibliometria, sia che essa venga utilizzata direttamente o indirettamente.

Indicatori bibliometrici (impact factor, h-index, citazioni, etc..) non possono essere impiegati nelle procedure di valutazione delle aree non-bibliometriche, anche qualora tali indicatori siano impiegati come strumenti di secondo o terzo livello per inferire circa la qualità delle pubblicazioni / prodotti di ricerca.

Ciò implica che riferimenti a liste di riviste per classi di merito o liste di case editrici o di collane di case editrici per classe di merito non è consentito in quanto distortivo rispetto alle finalità della valutazione.

Pertanto va superata la situazione attuale, per la quale alcune aree scientifiche disciplinari utilizzano nella valutazione della qualità della ricerca liste di riviste o di case editrici o di collane di case editrici, oppure un ibrido tra liste e peer-review per tipologie di pubblicazioni / prodotti di ricerca, mentre altre aree scientifiche disciplinari adottano prioritariamente procedure peer-review.

Peraltro, in alcune aree scientifiche disciplinari che usano liste di riviste per classe di merito, alcuni ricorsi ai tribunali amministrativi hanno condotto non solo alla cancellazione della collocazione di specifiche riviste nella classe di merito individuata dal GEV di area non-bibliometrica, ma anche alla invalidazione delle intere liste di riviste per classe di merito. Il riferimento qui è alla sentenza del TAR del Lazio del 2017, passata in giudicato, che ha fondamentalmente “delegittimato” nelle sue motivazioni a favore del ricorrente la costruzione delle liste di riviste utilizzate dal GEV13 in due consecutivi esercizi di valutazione.

In sintesi, è necessario stabilire la totale abolizione dell'utilizzo delle liste delle riviste o delle case editrici, o di collane delle case editrici, nel prossimo esercizio di valutazione.

In sostituzione di queste procedure, deve essere previsto l'impiego generalizzato della peer-review per tutte le tipologie di pubblicazioni / prodotti di ricerca. La peer-review è l'unico strumento in grado di realizzare una valutazione della qualità dei prodotti scientifici basata interamente ed esclusivamente sui contenuti dei prodotti, dunque sul loro grado di innovatività e sulla loro contribuzione al progresso scientifico e alla acquisizione di nuova conoscenza.

Questa procedura deve valere per tutte le aree disciplinari non-bibliometriche, in modo che vi sia uniformità e coerenza tra aree, secondo le modalità di seguito indicate.

## **7. Procedure per la individuazione dei valutatori nel modello peer-review**

La valutazione circa la qualità della produzione scientifica delle strutture viene realizzata dalla comunità scientifica relativa alle differenti aree scientifiche disciplinari, nella loro interezza, e non dai singoli GEV di area scientifico disciplinare, oppure da una selezione di valutatori individuati dai GEV mediante costruzione di liste di valutatori accreditati.

*E' la comunità scientifica che responsabilmente, in modo rigoroso, con regole condivise, si riappropria della valutazione della sua produzione scientifica.*

I componenti della comunità scientifica, universitaria e appartenente agli enti di ricerca, ovvero afferente alle strutture che sono oggetto di valutazione, deve essere *responsabilmente coinvolto* dell'esercizio di valutazione, secondo la modalità di coinvolgimento descritta qui di seguito.

Ogni ricercatore afferente alle strutture oggetto di valutazione è tenuto ad individuare tre sue specifiche aree di ricerca, avvalendosi della classificazione "settori ERC" (European Research Council) al secondo livello di dettaglio. Nel settore delle scienze sociali e umanistiche vi sono 6 macrosettori di primo livello, e 71 settori di secondo livello, per cui devono essere selezionati 3 tra i 71 settori potenziali. Sulla base di queste scelte, vengono quindi costruite liste di valutatori per ciascun settore di secondo livello di dettaglio.

## **8. Abbinamento tra valutatori e pubblicazioni / prodotti di ricerca nel modello peer-review**

Il processo di valutazione tramite peer-review deve essere orientato (in aggiunta all'obiettivo primario, che è una valutazione accurata, responsabile e basata sui contenuti specifici di ogni prodotto di ricerca) anche alla riduzione della complessità. Oltre a ciò, deve minimizzare il rischio di arbitrarietà nella scelta dei valutatori, perché questa può essere foriera di distorsioni dannose. Per perseguire questi obiettivi, si propone la seguente procedura:

- L'abbinamento tra valutatore e oggetto da valutare (pubblicazione / prodotto di ricerca) deve avvenire in modo del tutto casuale, mediante sorteggio, abbinando i settori ERC del valutatore con i settori ERC dell'oggetto da valutare;
- Ogni pubblicazione / prodotto di ricerca deve essere valutato da tre distinti valutatori, ognuno dei quali scaturisce tramite sorteggio nell'ambito dei tre settori ERC: l'abbinamento valutatore-oggetto da valutare per ognuno dei 3 ERC;
- Ogni valutatore non può valutare più di 5 pubblicazioni / prodotti di ricerca;
- L'arco temporale entro il quale la valutazione deve essere realizzata è fissata in 3 mesi;
- Il valutatore può rinunciare alla valutazione di pubblicazioni / prodotti di ricerca a lui assegnati entro la prima settimana dopo l'assegnazione del prodotto; in tal caso si procede immediatamente ad assegnare l'oggetto da valutare ad altro valutatore, secondo analoga procedura;
- Non è consentito assegnare al valutatore pubblicazioni / prodotti di ricerca realizzati da ricercatori che afferiscono alla stessa struttura del valutatore, sia a livello dipartimentale che di Ateneo, o di ente di ricerca.

## **9. Procedura e criteri di valutazione**

In coerenza con la premessa e le finalità di cui al punto precedente, la peer-review deve essere realizzata lungo le linee guida che seguono.

- Il valutatore è chiamato a fornire un punteggio su una scala prestabilita, identica per tutte le aree non-bibliometriche;
- Sono individuate 5 item oggetto di valutazione che sono:
  - a) rilevanza del lavoro nella letteratura / ambito di ricerca (*field*) di riferimento
  - b) originalità scientifica
  - c) coerenza interna
  - d) adeguatezza e originalità delle metodologie utilizzate
  - e) robustezza e rilevanza dei risultati conseguiti.
- Per ognuno di questo item il valutatore esprime un punteggio su scala identica tra aree, fissato nell'intervallo 0-10;
- Il punteggio complessivo assegnato al lavoro da un valutatore è semplicemente la somma dei punteggi assegnati sui 5 item oggetto di valutazione, per cui nell'intervallo 0-50;
- Il punteggio complessivo assegnato al lavoro dall'insieme dei 3 valutatori è semplicemente la somma dei punteggi assegnati da ciascuno dei valutatori, per cui nell'intervallo 0-150

Naturalmente, i punteggi possono essere normalizzati in una scala 0-10, o altra scala ritenuta adatta, per facilitarne la leggibilità.

#### **10. Indicatore sintetico di Classe della pubblicazione / prodotto di ricerca**

Al termine di tale procedura non rimane che ricollocare il risultato della valutazioni di ogni pubblicazione / prodotto di ricerca in un numero limitato di classi di merito. La ragione della necessità di questa operazione è duplice. Da un lato, serve a semplificare e facilitare ulteriormente la leggibilità dei risultati. D'altro lato, serve a ridurre distorsioni possibili legate al fatto che piccole differenze di punteggio possono emergere anche per pubblicazioni di pari valore. Queste differenze sono inevitabili anche in presenza delle migliori condizioni possibili (imparzialità, competenza, dedizione e attenzione dei valutatori), semplicemente perché la valutazione è, per sua natura, un processo umano che, come tale, inevitabilmente riflette almeno in parte la diversità dei punti di vista e dei criteri valutativi di ciascuno. Per evitare che a minori differenze di punteggio siano attribuiti significati eccessivi e non giustificati in sede di lettura e interpretazione dei risultati, appare opportuna una operazione di "accorpamento" dei punteggi in classi omogenee.

Le classi di merito della VQR 2014-2017 erano le seguenti 5:

1. ECCELLENTE A → peso: 1
2. ELEVATO B → peso: 0,7
3. DISCRETO C → peso: 0,4
4. ACCETTABILE D → peso: 0,1
5. LIMITATO E → peso: 0

A queste classi si aggiungevano le seguenti situazioni:

6. NON VALUTABILE - F → peso: 0, (la pubblicazione appartiene a tipologie escluse dalla VQR, senza alcun contenuto di ricerca o presenta allegati e/o documentazione inadeguati per la valutazione)

## 7. PRODOTTO MANCANTE rispetto al numero atteso → peso: 0

Secondo la proposta qui avanzata, per le ragioni indicato nei punti 1 e 2 di questo paragrafo, è improbabile che si presenti il caso (7) di cui sopra, che comunque può essere previsto.

Considerazione analoga si applica al caso (6), per le medesime ragioni; comunque anche esso può essere qui mantenuto.

Le classi rilevanti sono quindi le prime 5.

Dall'esito della valutazione, normalizzata in scala 0-10, ogni pubblicazione / prodotto di ricerca dovrebbe essere classificato nelle 5 Classi come segue:

1. ECCELLENTE A → peso: 1, per punteggio  $> 7$
2. ELEVATO B → peso: 0,7, per punteggio nell'intervallo  $> 4$  sino a  $=7$
3. DISCRETO C → peso: 0,4, per punteggio nell'intervallo  $> 1$  sino a  $=4$
4. ACCETTABILE D → peso: 0,1, per punteggio nell'intervallo  $> 0$  sino a  $=1$
5. LIMITATO E → peso: 0, per punteggio  $=0$  e casi (6) e (7) di cui sopra

## CONCLUSIONI

Le proposte correttive dell'attuale assetto sono finalizzate ad un miglioramento dell'efficacia della valutazione VQR rispetto all'obiettivo di generare una produzione scientifica più diversificata, di natura collaborativa tra le discipline, di maggiore qualità complessiva. I punti principali sono:

- Fornire più incentivi verso direzioni di ricerca diversificate e di natura più radicale, non focalizzata solo su tematiche e metodologie prevalenti in un dato momento storico
- 'Liberare' spazi alla ricerca inter / multi / trans disciplinare, in coerenza con gli approcci e scenari della progettualità e valutazione internazionale
- Ridurre l'enfasi sulla mera quantità dei prodotti da valutare
- Porre enfasi sul ruolo dello strumento di valutazione per analizzare e supportare il sistema della ricerca, anche per affrontarne le criticità
- Porre più enfasi sulla progettualità complessiva ed ampia del dipartimento come attore (posto che la VQR non è per nulla adeguata allo scopo e servono altri canali di finanziamento con enfasi su progettualità, ma il cambiamento di assetto può fornire incentivi diversi)
- Coinvolgere attivamente la comunità dei ricercatori nel processo di valutazione
- Contenere la complessità e i costi del processo valutativo

Sarebbe tuttavia necessario un ripensamento più radicale dell'assetto di valutazione, per andare veramente 'oltre' l'esperienza VQR, verso scenari dove la valutazione si pone a supporto di un sistema della ricerca caratterizzato da (i) integrazione tra la missione didattica e di ricerca (ii) nuova enfasi sulle progettualità di ricerca, connesse al lavoro di team, dei dipartimenti, degli atenei, e, non da ultimo, (iii) finanziamenti in crescita, almeno in linea con la media dei principali paesi europei.