

Società Italiana degli Economisti

Consiglio di Presidenza e Commissione Università, Ricerca, Valutazione

Verso una revisione delle procedure di valutazione

*Documento di discussione*¹

(4 giugno 2014)

1. Premessa

Non vi è dubbio che le procedure di valutazione della ricerca e delle strutture universitarie con la costituzione ed insediamento dell'Anvur (legge del 2006 ed insediamento nel 2011) siano state contrassegnate da discontinuità rispetto alle esperienze passate.

L'approccio adottato, pur con le debite differenze fra aree disciplinari, è stato quello di procedere con metodi "nuovi" rispetto alle precedenti esperienze distinguendosi per l'impiego di metodologie che richiamano quelle in vigore in ambito internazionale.

In particolare l'approccio adottato nella gestione della VQR e delle ASN per le quali l'Anvur ha sue specifiche competenze indicate dalla normativa e dai regolamenti che definiscono le sue attività è stato indubbiamente contrassegnato dall'utilizzo di strumenti di valutazioni quantitativi, ed in particolare dall'utilizzo di metodologie "bibliometriche", che soprattutto per alcune delle aree disciplinari umanistiche e sociali, di cui l'area economica è componente integrante, è stata accompagnata da metodologie di "valutazione dei pari".

La Sie ritiene che il processo intrapreso rappresenti un miglioramento, benché imperfetto, rispetto alla situazione pre-esistente. Pur con tutti i limiti che in varie sedi ed in più occasioni sono stati segnalati, l'attuale esperienza rappresenta un esperimento importante per far emergere ed implementare meccanismi premianti basati sul merito, che cercano di rimediare alle carenze ed imperfezioni altrettanto evidenti dei precedenti modelli che si sono contraddistinti da evidenti eccessi di autoreferenzialità del sistema di valutazione. Per questa ragione il processo non può e non deve essere interrotto, semmai rafforzato mediante anche profonde revisioni delle metodologie adottate che correggano gli errori compiuti e rispondano positivamente alle criticità emerse.

¹ Il presente "Documento di discussione" è stato predisposto in preparazione dell'incontro seminariale delle aree disciplinari umanistiche e sociali sulla valutazione della ricerca scientifica (Roma, 6 giugno 2014) ed intende costituire un momento di riflessione e confronto entro l'area disciplinare economica e tra aree disciplinari umanistiche e sociali. La stesura di questo testo è stata curata dal Presidente della Sie, Prof. Aldo Montesano, e dal Coordinatore Curv – Commissione Università, Ricerca, Valutazione, Prof. Paolo Pini. Su una versione precedente sono stati raccolti, e qui integrati per quanto i tempi lo hanno reso possibile, alcuni commenti e suggerimenti pervenuti dai componenti del Consiglio di Presidenza e della Curv della Sie, e dai Presidenti delle Associazioni Scientifiche di Area Economica che aderiscono alla Consulta delle Associazioni. *La presente versione non impegna comunque in alcun modo le Associazioni Scientifiche di riferimento.*

2. Fattori di criticità del processo di valutazione

Proprio al fine di fornire un contributo in tale direzione, la Sie ritiene cruciale non esimersi dal mettere in evidenza limiti che hanno caratterizzato il processo di cambiamento intrapreso e formulate alcune ipotesi di revisione delle procedure di valutazione.

Proprio perché sul versante dell'approccio bibliometrico si sono concentrate valutazioni critiche circa un impiego non del tutto convincente e non esente da criticità significative della bibliometria ai fini della valutazione, si ritiene importante evidenziare le criticità emerse, affinché dall'esperienza realizzata si traggano insegnamenti per il futuro e si implementino le necessarie revisioni.

Da un lato, è stato osservato più volte e da più parti che metodi quantitativi, anche basati sulla bibliometria, non possono essere utilizzati per la valutazione individuale dei singoli ricercatori o dei singoli prodotti di ricerca.

Dall'altro lato, le criticità hanno riguardato un utilizzo non sempre adeguato di metodi quantitativi, spesso basati sull'uso della bibliometria, per stilare liste di riviste, ordinate o meno, per fasce, oppure semplicemente per distinguere ciò che è una rivista scientifica da ciò che non lo è, e tra le riviste ritenute scientifiche quali meritano l'assegnazione nelle fascia di eccellenza, per la loro rilevanza soprattutto internazionale oltre che di impatto sulla professione.

Si ricorda che ai fini della definizione del carattere scientifico delle pubblicazioni e delle riviste è stata avviata una consultazione pubblica **ANPrePS** su iniziativa del CUN ed in collaborazione con l'Anvur, ma che il processo per giungere, in modo condiviso nella professione, alla realizzazione dell'*Anagrafe nazionale nominativa dei professori e dei ricercatori e delle pubblicazioni scientifiche*, prevista dalla normativa sin dal gennaio 2009, non è ancora giunto a termine. La mancanza di tale strumento ritenuto da tutti cruciale sta tuttora rendendo problematico ogni processo di valutazione dell'attività di ricerca e delle strutture universitarie ed enti di ricerca a vari livelli. Si riafferma pertanto l'urgenza improcrastinabile di poter disporre di tale strumento.

La Sie aveva a suo tempo espresso varie perplessità in merito sia alle procedure VQR (collocazione univoca delle riviste in aree disciplinari, collocazione in fasce di merito, procedura *peer review*, procedure di normalizzazione per area disciplinare, indicatori bibliometrici appropriati, ruolo delle monografie e dei saggi in volumi collettanei nel processo di valutazione), sia a quelle ASN (classificazione riviste come scientifiche, classificazione riviste di eccellenza internazionale fascia A). Al contempo la Sie ha formulato proposte di tipo specifico e generale, anche in relazione al metodo di consultazione delle associazioni esistenti nell'aree disciplinari delle scienze economiche. Nell'ambito di tali proposte, la Sie ha elaborato uno specifico ordinamento per fasce di merito delle riviste scientifiche, che è giunto alla seconda versione, quale strumento complementare nelle procedure di valutazione.

La Commissione Università, Ricerca, Valutazione ha peraltro segnalato importanti criticità delle procedure nel campo del VQR2004-2010 e ASN2012, criticità tali da sollevare perplessità importanti sugli esiti della valutazione. Tra queste la Commissione ha segnalato in particolare:

a) la procedura che ha portato alla compilazione delle riviste scientifiche di fascia A per il VQR, quindi impiegata in ambito ASN;

- b) l'insufficiente considerazione di specificità ed abitudini citazionali che caratterizzano le discipline economiche al loro interno;
- c) l'assenza di qualsiasi rivista italiana nella fascia A che disincentiva gli studiosi a pubblicare su riviste italiane ed al contempo rischia di creare una *trappola di non eccellenza* di lungo periodo per le stesse;
- d) le procedure di valutazione non del tutto adeguate dei prodotti di ricerca diversi dagli articoli su rivista scientifica, quali monografie e saggi su volumi, che hanno penalizzato questi ultimi.

In particolare si segnalano criticità che riguardano i seguenti aspetti.

- 1) Procedure di costituzione dei Gruppi di esperti della valutazione (Gev) e dei Gruppi di lavoro Riviste e libri scientifici
- 2) Trasparenza e pubblicizzazione dei criteri di valutazione, dell'ordinamento delle liste per fasce, del valore delle mediane, con *timing* corretti (*ex ante*) rispetto alle scadenze che devono rispettare i soggetti valutati nella scelta dei prodotti di ricerca per il VQR e nella presentazione delle domande per l'ASN
- 3) Verificabilità *ex-ante* da parte dei soggetti valutati (strutture od individui) dei criteri utilizzati in sede di valutazione dagli organi preposti, in particolare sulla collocazione delle riviste in fasce di merito, se non perfettamente conforme a quello delle specifiche aree disciplinari di riferimento, oltre che trasparenza circa la motivazione anche di eventuali sensibili scostamenti.
- 4) Definizione di rivista scientifica e rivista di eccellenza scientifica internazionale
- 5) Utilizzo complementare/sostitutivo degli strumenti bibliometrici e del metodo *peer review* nella valutazione dei prodotti di ricerca
- 6) Peso e ruolo delle monografie e dei saggi in volumi collettanei nel processo di valutazione
- 7) Peso e ruolo dei prodotti di ricerca pubblicati su riviste nazionali o per case editrici nazionali nella valutazione
- 8) Procedure di normalizzazione per la valutazione dei prodotti di ricerca che fanno riferimento ad aree disciplinare che hanno differenti abitudini citazionali e quindi influenzano la misurazione di impatti quantitativi bibliometrici nelle discipline economiche
- 9) Procedure di valutazione di prodotti della ricerca firmati da più ricercatori verso prodotti con un solo autore, da cui deriva il rischio di "scambi opportunistici" di firme
- 10) Impiego di banche dati bibliometriche private, la cui trasparenza non appare certificata, orientate a seguire più politiche di marketing che al progredire ed alla diffusione dei risultati della ricerca
- 11) Classificazione delle riviste per fasce di merito in base a criteri bibliometrici
- 12) Aggiornamenti del calcolo delle mediane degli indicatori bibliometrici
- 13) Rilevanza e adeguatezza delle classifiche di ordinamento applicato alle strutture, Atenei e Dipartimenti, e nell'ambito delle aree disciplinari
- 14) Utilizzo degli indicatori previsti dalla normativa o pubblicati dall'Anvur derivanti dagli esiti della valutazione al fine di distribuzione di risorse finanziarie, punti organico, valutazione interne di tipo concorsuale, ecc.

Tali criticità derivano anche dal ruolo privilegiato assegnato alla valutazione bibliometrica in assenza di adeguate normalizzazioni rispetto a quella della “valutazione dei pari”, e da come quest’ultima è stata condotta.

3. Alcune ipotesi di revisione delle procedure di valutazione

La Sie valuta positivamente l’avvio di una riflessione dal parte delle istituzioni coinvolte a vari livelli nel processo di valutazione circa alcune delle criticità sopra indicate e vede con favore una revisione delle procedure sia in ambito ASN che VQR.

In particolare si richiamano i seguenti aspetti.

Le abilitazioni non sono valutazioni comparative, ma certificazione di standard necessari per ricoprire un ruolo accademico.

Non tutte le commissioni sembrano hanno proceduto in tal senso. In area economia le difformità sono state a volte sostanziali. Inoltre, l’appartenenza a settore disciplinare è stata usata a volte come discriminare per l’abilitazione dei candidati in modo difforme e non pienamente giustificato, dando adito a rischi di discriminazione per campo di ricerca, metodo ed approccio, se non anche individuale, aspetto che può avere indotto la pratica dei ricorsi.

Agli Atenei spetta invece il compito di selezionare tra gli abilitati, seguendo proprie politiche di reclutamento.

Gli Atenei vengono poi premiati per il successo conseguito nelle politiche di reclutamento, valutato mediante la produzione scientifica, l’efficacia ed efficienza didattica e di terza missione che pure gli Atenei devono soddisfare

Ciò implica che le procedure devono essere tali da individuare criteri soglia ex-ante trasparenti e non discrezionali su cui accertare la qualità dei candidati. Criteri che tagliano la popolazione nel 50% sopra e sotto risultano troppo selettivi, perché portano ad escludere una quota troppo ampia di ricercatori dallo status di abilitato.

Si potrebbe anche ipotizzare che questi criteri siano del tutto automatici, ossia non prevedere l’intervento di commissioni che valutino i singoli candidati, ma prevedere che la domanda ad un concorso di Ateneo per il reclutamento possa essere esercitata soltanto da chi possiede pubblicazioni nel numero e nella qualità in base ai criteri stabiliti dalla commissione per l’abilitazione nelle singole aree o settori.

Al contempo, la discrezionalità con cui sono intervenute le commissioni nel rispetto delle mediane appare forse eccessiva ed eterogenea tra le stesse, dando luogo a quote di abilitati troppo variegata tra aree e settori disciplinari, tanto da creare casi anche eclatanti per taluni di questi.

Probabilmente, un certo tasso di standardizzazione e regole comuni sarebbero auspicabili, se non tra aree disciplinari, almeno tra settori afferenti alla stessa area.

In particolare, nel processo di valutazione, l’impiego di **strumenti quantitativi-bibliometrici** ha vantaggi ma anche svantaggi, come è ben noto. I secondi sono però troppo poco sottolineati e considerati dagli utilizzatori.

Con l'impiego di metodi quantitativi, concentrarsi sulla fascia superiore, di eccellenza o presunta tale, penalizza anche la varietà dei campi di ricerca e degli approcci metodologici e linee di pensiero, in quanto il rischio è che vengano premiate aree più *mainstream*, che non necessariamente sono quelle più innovative, e consolidate, dove i ritorni in termini citazionali sono maggiori e di breve periodo, rischiando di penalizzare non solo le nicchie ma anche la frontiera della ricerca, chi si muove in ambiti poco esplorati. Rammentiamo anche che gli esiti della ricerca scientifica sono per definizione incerti ed inattesi, le ricadute non previste e non prevedibili *ex-ante* per cui la varietà della ricerca e dei soggetti che la realizzano è un valore aggiunto cruciale per conseguire ricadute anche sociali, economiche, tecniche, per non dire culturali. Da questo punto di vista premiare l'eccellenza misurata dalla diffusione quantitativa (magari citazionale) può avere l'effetto - neppure tanto collaterale - di penalizzare l'innovatività. *Standardization driven vs. innovativeness driven* possono essere due modelli non sempre complementari se mal disegnati.

E comunque, in aggiunta, la politica più efficace non è quella "elitaria" di far sopravvivere e moltiplicare sempre la specie superiore, quanto indurre le specie erroneamente considerate inferiori verso percorsi virtuosi che accrescono l'esito complessivo del sistema. In alti termini eleviamo la media non penalizzando/discriminando chi sta sotto a favore dell'eccellenza, ma incentivando anche chi sta sotto la media a perseguire il miglioramento.

Inoltre l'applicazione delle soglie da esiti che dipendono da ciò che viene definito come **prodotto scientifico di eccellenza, di rilevanza internazionale, e comunque di elevata qualità**.

Con riferimento alla tipologia **prodotto pubblicato su rivista**, la collocazione delle stesse nella fascia A è cruciale e critica al contempo. Nel campo economico, la varietà di approcci, metodi, campi di ricerca non è adeguatamente rispettata. Inoltre, la varietà della lingua di diffusione e del luogo editoriale è insufficiente, e porta a distorsioni non solo nella valutazione ma ricadute nelle politiche di ricerca e di diffusione della produzione scientifica.

In questo ambito non possiamo non segnalare il rischio che si possa produrre una distorsione a svantaggio di campi e temi di indagine che hanno una base informativa geograficamente contenuta rispetto a quella richiesta da standard di diffusione internazionale. Analisi *country o sub-country specific* risentono ad esempio proprio di questo rischio di distorsione, che andrebbe poi a penalizzare campi e temi di ricerca.

Una criticità specifica ma significativa riguarda in campo economico le riviste italiane, sia che esse siano pubblicate in lingua inglese, quindi con maggiore potenzialità di diffusione e di impatto internazionale, sia in lingua italiana. L'assenza di qualsiasi rivista italiana in fascia A determina anche in presenza di un *upgrading* qualitativo (percorso virtuoso ad esempio verso l'internazionalizzazione) delle stesse, un fenomeno di *lock-in* verso il basso di queste: i ricercatori non sono incentivati a pubblicare su riviste nazionali perché nessuna in fascia A anche se queste sono impegnate in *upgrading* qualitativo (che non viene però riconosciuto), quindi le riviste attrarranno solo lavori non di eccellenza e potranno ambire solo con molte difficoltà ad un *upgrading* quantitativo (passare da fasce inferiori a fasce superiori).

Inoltre, l'assenza di riviste italiane in fascia A nel settore economico si è accompagnata con la presenza di numerose riviste italiane in fascia A in altri settori e aree, senza che vi siano differenze qualitative tali da giustificare questa diversità di collocazione.

Infine, il criterio bibliometrico utilizzato per qualificare le riviste in fascia A esclude le riviste nuove, ne pregiudica l'affermazione e ostacola l'affermazione di nuovi settori di ricerca.

Con riferimento alla tipologia **prodotto pubblicato su volume** (saggio o monografia), la mancanza di fasce di riferimento di case editrici e collane contribuisce ad indurre a considerare tali prodotti di ricerca come marginali, come regola di base, anche quando non lo sono. Volumi o saggi anche

importanti dal punto di vista contenutistico, in quanto monografia o saggi in volumi collettanei o pubblicazioni organiche e strutturati derivanti da progetti di ricerca, rischiano di essere considerati marginali, solo in ragione di una valutazione non premiante del “contenitore” più che del “contenuto”, e nonostante le citazioni che possono ricevere. In questo ambito non occorre però trascurare neppure l’impatto che può avere la struttura fortemente oligopolistica del mercato editoriale internazionale, che andrebbe a distorcere grandemente la collocazione delle case editrici per fasce di qualità basata eccessivamente su eventuali indicatori quantitativi, quali le citazioni.

Il metodo della **valutazione tra i pari**, d’altronde, soffre di problemi noti - ma diversi da quelli dei metodi quantitativi bibliometrici, e secondo alcuni si deve confrontare con vincoli di bilancio sui costi (ma anche la bibliometria dovrebbe allora essere valutata sui costi!).

Qui la varietà deve essere assicurata dalla popolazione dei referee, trasparente e non distorta. Un percorso da seguire proprio perché attualmente è soggetto a limiti significativi su entrambi i fronti.

Tuttavia, nonostante i suoi specifici limiti, la “valutazione dei pari” rimane cruciale nel processo di valutazione. Il suo utilizzo non può essere sostituito nella valutazione dei singoli prodotti di ricerca da metodologie unicamente quantitative-bibliometriche. Si riafferma quindi la necessità di procedure trasparenti e rigorose di merito affidate, in particolare, alla “valutazione dei pari” ed alla lettura diretta delle singole pubblicazioni.

Se questo strumento non può essere abbandonato nell’ambito del VQR, ipotesi differenti possono essere esaminate nell’ambito della valutazione per le ASN se queste sono correttamente intese come cosa diversa dalla valutazione comparativa (come sopra ricordato), ed a maggior ragione se si va verso una procedura “a sportello”. In tal caso, la revisione nell’impiego di procedure quantitative verso metodi meno distortivi appare ancora più cruciale. Ad iniziare dall’impiego delle mediane come soglia di sufficienza, soprattutto perché si sta adottando un metodo che porta a valutare un prodotto sulla base del “contenitore” che rischia di prescindere dal suo “contenuto”.

Occorre anche rimarcare un aspetto centrale e prioritario, costituito dalla definizione di prodotto scientifico di ricerca, su cui le soglie si applicano. Qui si continua a rilevare la relativa approssimazione presente nella valutazione, in assenza di una **anagrafe della ricerca scientifica**. La sua costruzione dovrebbe essere la priorità, peraltro prevista dalla normativa. Gli strumenti attuali svolgono funzione di supplenza, peraltro altamente imperfetta.

Con riferimento alla revisione dei settori disciplinari, si osserva che possibilità e praticabilità di questo percorso si deve confrontare con il **timing** di tale revisione, timing normativo e timing sostanziale che richiede questo ultimo una apertura del confronto tra aree disciplinari particolarmente impegnativo. Ciò non implica che la questione non debba essere posta in agenda, ma con soluzioni da realizzare che per i tempi attesi e prevedibili non possono comportare il rischio di blocchi del processo di valutazione.

Si potrebbe prospettare un allargamento o **accorpamento nell’ambito di una stessa commissione** di abilitazione che racchiuda più settori disciplinari con l’aumento delle competenze disciplinari presenti in ciascuna commissione. In tal modo si eviterebbe il fenomeno di “strane” non abilitazioni certificate da dubbie declaratorie tipo “non appartenenza al settore”, oppure di multiple abilitazioni ma per fasce accademiche diverse. Situazioni queste che, secondo alcuni, vengono forse non correttamente sanate ex-post sul piano normativo consentendo in modo generalizzato ai candidati la partecipazione a concorsi per reclutamento diversi dall’area disciplinare della abilitazione.

Questa opzione potrebbe anche consentire di superare lo scoglio del *timing* implicito nella revisione dei settori disciplinari, il cui impiego potrebbe rimanere invece nella fase del reclutamento, a meno che non si intenda intervenire anche su tale fronte.

Infine, ci preme segnalare come a fronte dei cambiamenti intervenuti nel modello e nelle procedure di valutazione della ricerca scientifica, ma anche in quelle della didattica, si sia generata a vari livelli una accresciuta “burocratizzazione” che ha condotto ad una intensificazione delle componenti “formali” dei processi a scapito a volte di quelle “sostanziali”. La necessità di andare verso una semplificazione normativa e procedurale diviene essenziale, affinché i miglioramenti indispensabili ed auspicabili nella valutazione non rischino di essere frenati se non compromessi dall’appesantimento delle procedure. La percezione ed i riscontri di una “deriva burocratica” della valutazione richiede con urgenza interventi che vadano nella direzione opposta, perché ad una valutazione premiale fondata sul merito non è concesso rinunciare.