

Società Italiana degli Economisti

Nota di discussione per una revisione delle procedure di valutazione¹

a cura di

Paolo Pini (Coordinatore Curv - Sie)

e

Marco Valente (Rappresentante CUN Area 13, per le discipline economiche)

(versione 22 giugno 2014)

Predisposta per i lavori del Consiglio di Presidenza Sie del 28 giugno 2014 e per la consultazione nell'ambito della Commissione Università, Ricerca, Valutazione - Sie

1. Premessa

Non vi è dubbio che le procedure di valutazione della ricerca, delle strutture universitarie e dei ricercatori siano state contrassegnate nelle ultime riforme da discontinuità rispetto alle esperienze passate.

Se da un lato il processo intrapreso ha rappresentato un miglioramento, benché imperfetto, rispetto alla situazione pre-esistente, dall'altro ha fatto emergere carenze e limiti che possono rischiare di fermare gli avanzamenti realizzati.

Il rischio è duplice. Da un lato quello di tornare a modelli di valutazione che si erano distinti per un eccessivo tasso di autoreferenzialità, rischio che deve essere evitato, dall'altro quello opposto di un percorso poco condiviso dalle comunità scientifiche e dagli stessi soggetti oggetto di valutazione, strutture e singoli ricercatori.

Al fine di far avanzare meccanismi premianti basati sul merito, che cercano appunto di rimediare alle carenze ed imperfezioni evidenti dei precedenti modelli, occorre invece intervenire anche con profonde revisioni ritenute necessarie delle metodologie adottate di recente, in modo da correggere gli errori compiuti e rispondere positivamente alle criticità emerse.

Il principio cui si ispirano le presenti note è comunque che il sistema accademico deve proseguire nel percorso che introduce un sistema di valutazione delle sue componenti (meglio se più di un sistema) per evitare che le necessarie scelte siano basate su criteri diversi da quelli di interesse della società. Proprio per garantire questo obiettivo è necessario tener conto che l'irriducibile incertezza dei frutti della ricerca, nonché la sua intrinseca immisurabilità, impone procedure con ampi margini di tolleranza, in modo da evitare che gli inevitabili errori si traducano in danni irrimediabili.

Queste note si propongono di evidenziare le criticità più rilevanti e fornire alcune opzioni di revisione.

¹ Questa nota trae spunto, utilizza ed estende le considerazioni contenute nel *Documento di discussione* del 4 giugno 2014 a firma di Aldo Montesano e Paolo Pini, preparato in occasione della riunione delle aree umaniste e sociali 10-14 del CUN dedicato alla revisione delle procedure di valutazione, 6 giugno 2014, Roma. La responsabilità del presente testo è da attribuirsi agli autori. Si ringraziano coloro che hanno discusso e commentato con gli autori gli argomenti e le tesi qui avanzate.

2. Fattori di criticità e ipotesi di revisione del processo di valutazione in ambito VQR

Tenuto conto che l'ultimo esercizio di valutazione della ricerca VQR 2004-2010 è stato contrassegnato per l'area 13 dall'ampio utilizzo di strumenti di valutazioni quantitativi, ed in particolare dall'utilizzo di metodologie bibliometriche, anche con un approccio sperimentale, e che questo esercizio ha mostrato evidenti limiti nella fase della sua realizzazione, si ritiene indispensabile avviare una fase di revisione delle metodologie adottate.

2.1 Anagrafe nazionale della ricerca

Va anzitutto segnalato che l'esercizio è stato svolto in assenza di una definizione del carattere scientifico dei prodotti di ricerca (pubblicazioni) e delle sedi di pubblicazioni (riviste o volumi). Questa mancanza ha comportato in particolare la costruzione di liste di riviste, distinte poi per fasce, sulla base di informazioni non verificate e non certificate tratte dagli autonomi inserimenti da parte dei ricercatori di prodotti di ricerca nella banca dati Cineca U_Gov. Tale procedura, come è noto, ha contribuito all'inserimento nelle suddette liste di riviste carenti di standard di scientificità adeguati.

Successivamente all'inizio delle procedure VQR, è stata realizzata una consultazione pubblica **ANPrePS** su iniziativa del CUN ed in collaborazione con l'ANVUR, da cui è seguita la stesura di un documento CUN sul carattere scientifico dei prodotti di ricerca e delle sedi di pubblicazione trasmesso poi al MIUR con l'intento di favorire, in modo condiviso nella professione, la realizzazione dell'*Anagrafe nazionale nominativa dei professori e dei ricercatori e delle pubblicazioni scientifiche*, prevista peraltro dalla normativa sin dal gennaio 2009.

Tale processo non è però giunto al termine e la costruzione dell'*Anagrafe* tuttora non è stata avviata.

La mancanza di tale strumento - ritenuto cruciale - rende 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 *Anagrafe nazionale*.

Le resistenze all'ANPrePS appaiono comunque forti. In alcuni casi vengono giustificate dal fatto che alcune componenti delle istituzioni pubbliche hanno sviluppato una loro anagrafe proprietaria i cui ritorni sarebbero a rischio qualora un nuovo sistema si affermasse o venisse perseguito. In aggiunta, vi è la preoccupazione che la nascita dell'anagrafe nazionale richieda un ennesimo impegno di tutti i soggetti coinvolti che rischia di essere percepito come ulteriore elemento di burocratizzazione per un sistema già sottoposto a continui stress procedurali ed informativi. Sarebbe quindi opportuno verificare ex-ante la compatibilità di sistemi in essere con il progetto ANPrePS disegnato dal CUN previa consultazione pubblica e trasmesso al MIUR.

2.2 Approccio bibliometrico vs valutazione dei pari

Un secondo aspetto rilevante da considerare è l'impiego di due metodologie distinte nel processo di valutazione.

Da un lato, un approccio puramente quantitativo che impiega indicatori bibliometrici basati su dati citazionali, che spesso conducono a valutare il prodotto di ricerca sulla base della veste editoriale con la quale è diffuso nella comunità scientifica piuttosto che sul suo diretto valore intrinseco. Dall'altro, un approccio che pone al centro della valutazione il contenuto del prodotto piuttosto che il contenitore, e quindi una valutazione di merito da parte dei pari.

Questi due approcci presentano ognuno vantaggi e svantaggi assai noti, di procedura e di costo. L'utilizzo esclusivo di uno di questi è però indicato da alcuni come inadeguato, preferendo la compresenza di entrambi in modo equilibrato.

Raggiungere questo equilibrio non è tuttavia agevole, in quanto occorre individuare procedure che non comportino distorsioni in un senso o nell'altro, oppure che non diano luogo a risultati contraddittori. Entrambe le situazioni si sono verificate nell'esercizio VQR 2004-2010: in ambito delle aree bibliometriche sono stati verificati risultati contraddittori, mentre nell'area 13 si sono avuti effetti distorsivi a causa delle procedure adottate.

Emblematico è il processo di validazione dei risultati bibliometrici tramite le risultanze della valutazione dei pari per un 10% di prodotti valutati con indicatori bibliometrici, per i quali si dispone di indicatori diretti oppure imputati. L'analisi condotta per la gran parte delle aree disciplinari evidenzia una concordanza di valutazione limitata e spesso molto al di sotto del livello soglia minima. Nel caso dei prodotti dell'area 13, si registrano valori più elevati ma sempre al di sotto della soglia considerata di concordanza moderata. Pesa in particolare per l'area 13 la metodologia adottata per condurre l'analisi che ha condotto ad una sovrastima dell'indice di concordanza, in quanto - ad esempio - le due modalità, bibliometria e valutazione dei pari, non sono indipendenti. I risultati ottenuti non supportano l'uso esclusivo della bibliometria per valutare la qualità dei prodotti di ricerca, impattati o meno, semmai avrebbe suggerito di impiegare in modo congiunto i due strumenti senza porre nessuno dei due in posizione di preminenza nella valutazione.

L'area 13 ha una sua specificità data dalla presenza di una quota di prodotti di ricerca che non possono essere valutati mediante indicatori bibliometrici, e richiedono valutazione dei pari. In tal caso è emerso che le procedure di valutazione hanno penalizzato i prodotti inviati in valutazione dei pari se confrontati con gli altri, penalizzando quindi volumi monografici e saggi in volumi collettanei e quindi anche quelle aree scientifiche che impiegano più di altre tale formato editoriale.

I risultati delle procedure di classificazione delle riviste nelle fasce di merito costruite per l'esercizio VQR sono stati poi trasferiti in modo automatico per definire le riviste di fascia A per le ASN. Nonostante le associazioni scientifiche di riferimento abbiano espresso motivate obiezioni a questa procedura, l'ANVUR con il suo GL per l'area 13 ha sostanzialmente ignorato le proposte ricevute optando per considerare nella classe A solo le riviste classificate nella classe più elevata dell'esercizio VQR, e tale operazione viene fatta: a) assegnando per sub-aree 13 (macro-aree) in modo univoco le riviste di classe A; e quindi b) costruendo la lista delle riviste di classe A indipendentemente dagli SSD. L'effetto è stato quello di costruire liste esclusive per macro-aree ed ammettere poi un numero consistente di riviste come appartenenti a tutte le macro-aree, senza alcuna considerazione degli SSD. Nella successiva fase di aggiornamento per la definizione delle riviste di classe A inizialmente costruite con indici di impatto al 2010, si è poi dovuto affrontare il problema della collocazione delle nuove riviste in cui i candidati avevano pubblicato i loro lavori; in tal caso si è scelto di aggiornare la lista sulla base dei loro nuovi valori bibliometrici, lasciando immutato il posizionamento delle riviste esistenti. Al termine del processo si sono così costruite liste di riviste di classe A ibride, e non internamente coerenti.

Una opzione che è emersa di recente è quella di distinguere nettamente le procedure delle aree bibliometriche da quelle non bibliometriche. Nelle prime si abbandona del tutto ogni controllo o verifica dei pari, nelle seconde ci si affida unicamente a questi.

Questa opzione porrebbe problemi seri nell'area 13 per almeno tre ragioni:

- a) Alcune sub-aree preferirebbero forse optare per la bibliometria assoluta, mentre altre per la valutazione dei pari assoluta. Ciò produrrebbe una frattura nelle procedure di valutazione entro l'area 13, con effetti che non si limitano alla valutazione, ma si estendono anche alle posizioni che difficilmente l'area 13 potrebbe raggiungere in modo unitario in sede istituzionale.

- b) In presenza di prodotti non valutabili con modalità quantitative, si avrebbe la secca alternativa tra non considerare tali prodotti, con effetti ovviamente discriminanti per alcune aree disciplinari e finanche filoni di ricerca, approcci ed anche singoli ricercatori, oppure ritornare alla valutazione dei pari in questi casi, e quindi riproponendo la questione dell'equilibrio tra procedure di valutazione quantitativa e valutazione dei pari.
- c) In presenza di prodotti valutabili con modalità quantitativa, l'esame di validazione condotto per un campione del 10% dei prodotti non supporta la scelta di impiegare in modo esclusivo il metodo bibliometrico, ma suggerisce che bibliometria e valutazione dei pari, proprio perché forniscono informazioni non adeguatamente concordanti, sono modalità di valutazione che dovrebbero essere utilizzati in modo congiunto senza preminenza dell'uno sull'altro.

Il metodo della valutazione tra i pari, d'altronde, soffre di problemi noti - ma diversi da quelli dei metodi quantitativi bibliometrici, e si deve confrontare con vincoli di bilancio in termini di costo.

Un aspetto cruciale è la varietà che dovrebbe essere assicurata circa la popolazione dei valutatori, trasparente e non distorta per approcci disciplinari. A tal fine, un maggior coinvolgimento delle associazioni scientifiche di riferimento è auspicabile, e non può limitarsi alla mera prassi di consultazione formale condotta nel passato dall'agenzia preposta alla valutazione.

Poco appropriata invece appare l'insistenza su procedure *double blind* o *single blind* per prodotti di ricerca che sono pubblicazioni su riviste o volumi.

La valutazione dei pari è ovviamente cruciale con riferimento alla tipologia "prodotto pubblicato su volume" (saggio o monografia). In questo ambito segnaliamo che, in presenza del preminente approccio quantitativo alla valutazione, la mancanza di fasce di riferimento di case editrici e collane può contribuire a valutare tali prodotti di ricerca come marginali. Volumi o saggi anche importanti dal punto di vista contenutistico, in quanto monografia o saggi in volumi collettanei o pubblicazioni organiche e strutturate derivanti da progetti di ricerca, rischiano di essere considerati residuali nella valutazione in ragione di una valutazione non premiante del "contenitore" più che del "contenuto", e nonostante il loro impatto citazionale.

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.

2.3 Procedure di valutazione

Nell'ambito dell'avvio delle attività dell'ANVUR, sin dalla sua costituzione, il confronto pubblico ha portato a far emergere alcune problematiche che non sembrano aver trovato sinora soluzione soddisfacente. Tali problematiche, di natura eterogenea, appaiono aver avuto rilevanza significativa nel produrre esiti della valutazione non ottimali. Si fa osservare che revisioni delle procedure sarebbero necessarie in questi ambiti.

In primo luogo, si segnala che la costituzione dei Gruppi di Esperti della Valutazione (GEV) e dei Gruppi di Lavoro Riviste e Libri Scientifici (GL) non appare avvenuta in modo tale da assicurare una adeguata pluralità di orientamenti scientifici. Una revisione nella composizione dei GEV e GL in area 13 è fortemente auspicabile.

In secondo luogo, dovrebbero essere maggiormente rispettati metodi di trasparenza e pubblicizzazione dei criteri di valutazione, dell'ordinamento delle liste per fasce, con *timing* adeguati (*ex ante*) rispetto alle scadenze a cui devono attenersi i soggetti valutati nella scelta dei prodotti di ricerca per il VQR. Appare infatti essenziale che si realizzi una verificabilità *ex ante* da

parte dei soggetti valutati (strutture od individui) dei criteri utilizzati in sede di valutazione. Ciò purtroppo non è avvenuto in particolare per la collocazione delle riviste in fasce di merito, in generale, ed in particolare in quei casi di non perfetta conformità rispetto alle specifiche aree disciplinari di riferimento. Carenze di trasparenza sono emerse circa la motivazione per eventuali sensibili scostamenti.

In terzo luogo, l'adozione di procedure di normalizzazione per la valutazione dei prodotti di ricerca che fanno riferimento ad aree disciplinare con differenti abitudini citazionali avrebbe potuto evitare che la misurazione dei fattori d'impatto quantitativo (bibliometrici) conducesse ad una non corretta valutazione relativa della produttività e qualità scientifica tra le varie discipline economiche, e tra queste ed altre discipline. Procedure di normalizzazione avrebbero anche dovuto essere applicate per la valutazione di prodotti della ricerca firmati da più ricercatori verso prodotti con un solo autore, dalla assenza delle quali derivano rischi di "scambi opportunistici" di firme.

In quarto luogo, effetti distorsivi a causa della presenza di tali problematiche risultano peraltro accresciuti per il peso prevalente che è stato assegnato alla valutazione bibliometrica, in assenza di adeguate normalizzazioni, rispetto a quello ridotto assegnato alla valutazione dei pari, anche a seguito di come questa è stata condotta in area 13.

2.4 Approccio bibliometrico

Sull'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.

Anzitutto, viene osservato che metodi quantitativi, anche basati sulla bibliometria, non possono essere utilizzati per la valutazione individuale dei singoli ricercatori o dei singoli prodotti di ricerca in ambiti differenti da quelli nei quali sono stati concepiti. Approcci di tipo quantitativo sono più adatti per la valutazione delle strutture, ed in tale ambito possono essere utilmente impiegati. Valutazioni dei prodotti di ricerca in ambito VQR non dovrebbero invece essere impiegate a fini diversi da quelli per cui espressamente sono stati concepite. Lo ricordano sia il CUN che la stessa ANVUR, anche se poi nella pratica dell'accreditamento dei dottorati tale indicazione è stata contraddetta.

Le criticità dell'utilizzo della bibliometria sono peraltro presenti nell'esercizio VQR 2004-2010.

La Commissione Università, Ricerca, Valutazione ha peraltro segnalato alcune di queste criticità delle procedure VQR2004-2010 con ricadute su ASN2012 e ASN2013.

Tra queste, qui segnaliamo con maggiore dettaglio in particolare i seguenti:

- a) la procedura che ha portato alla compilazione delle riviste scientifiche di fascia A per il VQR, quindi impiegata in ambito ASN, con l'imputazione di valori bibliometrici a riviste che non disponevano di fattori d'impatto certificati in ambito WoS mediante procedure non accurate che hanno impiegato indicatori citazionali in ambito Google Scholar;
- b) l'insufficiente considerazione di specificità ed abitudini citazionali che caratterizzano le discipline economiche al loro interno per cui in assenza di opportune procedure di normalizzazione sono risultati penalizzati in termini relativi settori disciplinari significativi;
- c) l'assenza totale di riviste italiane 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 e le

relative aree di ricerca, settori disciplinari, approcci metodologici che hanno abitudini editoriali e citazionali loro proprie.

2.5 Il valore aggiunto della varietà

L'impiego di metodi quantitativi è stato accompagnato dall'idea diffusa che sia non solo auspicabile ma persino irrinunciabile premiare le aree a maggior diffusione della ricerca, che godrebbero del merito di eccellenza, con il rischio di penalizzare la varietà dei campi di ricerca, degli approcci metodologici e linee di pensiero. Infatti, il rischio è quello di premiare comunque aree più *mainstream*, indipendentemente dal loro grado di innovatività, solo perché risultano essere quelle più consolidate, dove i ritorni in termini citazionali sono maggiori e di breve periodo. In tale processo, il rischio di penalizzazione non è tanto per le nicchie di ricerca, quanto per chi si muove sulla frontiera della ricerca, in ambiti poco esplorati e poco consolidati per definizione.

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 con indicatori prevalentemente citazionali 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.

Si può anche aggiungere che la politica più efficace non è quella "elitaria" di far sopravvivere e moltiplicare sempre la specie superiore, quanto indurre le specie considerate - a volte erroneamente - inferiori verso percorsi virtuosi che migliorino il risultato complessivo del sistema. In alti termini, potrebbe essere conveniente per il sistema elevare la media senza penalizzare o discriminare chi sta sotto a favore esclusivo dell'eccellenza, ma incentivando la non eccellenza a perseguire il miglioramento.

Vi sono quindi motivazioni che portano a suggerire che procedure di valutazione che non distinguono tra aree di ricerca e tra metodologie, applicando all'universo della popolazione criteri standardizzati, non risultano le più adatte per preservare, ed ancor meno accrescere, la varietà. Metodi di normalizzazione tra settori nelle valutazioni quantitative, da un lato, e valutazione dei pari, dall'altro, costituiscono due strumenti cardine che invece possono ridurre il rischio della omologazione.

2.6 Produzione scientifica, sua collocazione editoriale nazionale ed effetti di lock-in

Il rischio di distorsione di cui sopra si presenta anche per gli ambiti di ricerca che riguardano campi e temi di indagine con/caratterizzati da una base informativa geograficamente limitata 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. Prodotti di ricerca che hanno queste caratteristiche trovano più difficile collocazione e diffusione internazionale su riviste prestigiose ad alto impatto citazionale, mentre trovano più agevole collocazione su riviste che hanno diffusione in ambito nazionale.

Una criticità specifica - ma significativa - riguarda così 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, pur in presenza di un *upgrading* qualitativo (percorso virtuoso ad esempio verso l'internazionalizzazione) delle stesse, un fenomeno di *lock-in* in fasce inferiori di queste: i ricercatori non sono incentivati a pubblicare su riviste

nazionali perché nessuna di queste è 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).

Peraltro, l'assenza assoluta di riviste italiane in fascia A nelle discipline economiche si è accompagnata con la presenza di numerose riviste italiane in fascia A in altri settori e aree disciplinari, senza che si possano osservare differenze qualitative tali da giustificare questa diversità di collocazione. Ciò determina ovviamente una penalizzazione delle discipline economiche nel processo di valutazione, a vantaggio delle altre aree disciplinari anche affini, quelle umanistiche e sociali, effetto che peraltro appare prodotto dalle scelte compiute nella nostra disciplina.

Sarebbe opportuno invece che la disciplina favorisse un approccio alle "pari opportunità", che non releghi le riviste italiane nelle fasce inferiori in ragione esclusiva dai fattori d'impatto che esse presentano, ma premiasse i percorsi virtuosi che molte di queste hanno avviato, ad esempio nelle procedure di valutazione dei pari, nella diffusione internazionale anche mediante nuovi canali editoriali, nell'utilizzo della lingua preminente di pubblicazione, ecc. Le indagini che la Sie ha condotto nel 2011 e nel 2013 segnalano come una quota significativa di riviste italiane sia impegnata negli anni recenti, anche con successo, in percorsi virtuosi che meritano di essere premiati. L'assenza di riviste italiane in fascia A costituisce invece un cruciale fattore di ostacolo lungo tale percorso.

Una possibile revisione radicale della VQR potrebbe consistere nella creazione di diversi indicatori di VQR ciascuno per obiettivi diversi. Ad esempio, un indicatore di semplice attività-inattività, uno di intensità di pubblicazione, uno di eccellenza, ecc. Tali indicatori non risolverebbero completamente i problemi elencati, ma il fatto di averne diversi renderebbe più agevole l'individuazione dei criteri per ciascun indicatore, e permetterebbe di distinguere diversi profili senza omogeneizzazioni forzate.

3. Fattori di criticità e ipotesi di revisione del processo di valutazione in ambito ASN

In generale, le procedure per il reclutamento universitario devono soddisfare (almeno) due distinte esigenze: assicurare un ambiente di lavoro adeguato al ricercatore che produca innovazione scientifica e garantire la sua integrazione in un organismo complesso come un dipartimento. Ovviamente quando si deve scegliere tra combinazioni diverse, tra le due esigenze si pone un problema di compatibilità che non ha una soluzione ovvia.

Il sistema introdotto con la legge 240 usa il principio, discutibile ma comprensibile e coerente, di separare le due fasi di selezione per i due obiettivi: la abilitazione garantisce sul fronte della qualità scientifica mentre il concorso locale si occupa di valutare la coerenza dei candidati con la posizione richiesta. Per ora abbiamo potuto osservare gli esiti della ASN, mentre i primi bandi per il reclutamento si stanno chiudendo in questi giorni.

Nel seguito vengono proposte quindi alcune considerazioni espresse nella forma di fattori di criticità e possibili loro soluzioni.

Ricordiamo comunque che il contesto nel quale ci muoviamo attualmente è piuttosto mobile, con un tasso di cambiamento atteso che fino a pochi mesi fa non era previsto, in cui, oltre ad essere in presenza di proposte specifiche provenienti da più fonti di revisione delle procedure di abilitazione nazionale, sembra esservi anche un differente posizionamento tra istituzioni, il MIUR in primo luogo, l'organismo di rappresentanza e autogoverno dell'accademia italiana, il CUN, e l'agenzia pubblica di valutazione della ricerca scientifica, l'ANVUR.

Ciononostante, proviamo, assumendocene il rischio, ad indicare possibili aggiustamenti che potrebbero alleviare uno o più delle criticità individuate. Le proposte hanno lo scopo di stimolare la discussione in termini propositivi, non hanno pretesa né di esaustività né di unicità riguardo alla soluzione dei problemi indicati.

3.1 Abilitazione vs Competizione

Probabilmente la singola causa che ha provocato i danni maggiori è che molte commissioni hanno interpretato (e sono state portate a farlo da una gestione amministrativa ondivaga, ad esempio via circolari ministeriali) il loro mandato in termini competitivi, decidendo sulla posizione relativa della qualità dei candidati (X è meglio di Y) e non sull'essere o meno adeguati scientificamente alla posizione richiesta (X ha un profilo da professore, Y no).

Dall'evidenza disponibile non emerge che in generale le commissioni abbiano proceduto a prescindere da valutazioni comparative. In area economica si segnalano difformità in questo senso. La stessa appartenenza ad uno specifico settore disciplinare è stata usata a volte come discriminante 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.

Secondo alcuni, il danno prodotto da questo approccio nella procedura di selezione sembra duplice.

In primo luogo, come già spesso notato, si sono favoriti in alcuni settori disciplinari i profili di ricerca coerenti con il *mainstream* ed è stata fortemente penalizzata la varietà di approcci metodologici ed opinioni, che invece è la principale condizione necessaria di avanzamento scientifico nel lungo periodo.

In secondo luogo, in modo meno evidente ma non meno dannoso, la selezione effettuata surrettiziamente dalle commissioni nazionali di abilitazione rischia di essere sfruttata dalle commissioni di reclutamento locale per evitare qualsiasi selezione, giudicando in modo erroneo tutti gli abilitati come egualmente validi, e quindi selezionando ad esempio di preferenza gli interni rispetto agli esterni, oppure secondo logiche puramente amministrative o di gestione interna.

Gli Atenei dovrebbero invece essere 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.

3.2 Dimensione delle commissioni

Il numero ristretto di commissari (sostanzialmente 4, essendo stato il membro OCSE spesso ininfluente se non per definire al margine maggioranze) porta a due tipi di problemi. La probabilità di avere la maggioranza relativa della commissione omogeneamente schierata su determinate posizioni rende molto facile negare l'abilitazione a chi non raccoglie ampio consenso nella commissione (si ricorda anche la regola dei 4/5 dei voti per la concessione). Da notare che, per ovvi motivi campionari, la composizione della commissione sarà con alta probabilità molto diversa dalla effettiva composizione della comunità che dovrebbe rappresentare. Il problema è esacerbato dalle modalità di scelta dei commissari, non su base elettiva, che porta facilmente a distorsioni a causa della diversa propensione alla partecipazione ad incarichi amministrativi.

3.3 I settori concorsuali

La frammentazione dei settori concorsuali sposta la discussione non solo (e non tanto) sulla qualità scientifica dei candidati ma sulla loro compatibilità con la declaratoria del settore o, peggio, sulla sua interpretazione da parte di fazioni della politica accademica. Questo rischia di penalizzare

fortemente i profili anche solo moderatamente innovativi rispetto alle frontiere settoriali (determinate da risultati scientifici passati), creando un blocco insormontabile ad uno dei processi fondamentali della ricerca, le trasposizioni interdisciplinari di metodi e risultati.

Inoltre, le ovvie dinamiche dialettiche fanno sì che si scoraggino non solo i ricercatori che centrano la loro attività su argomenti di confine, ma anche candidati forti in un dato settore che lavorano *anche* in altri settori.

L'esistenza e la artificiosità normativa di questo problema è testimoniata anche dal semplice fatto che la legge permette agli abilitati di partecipare a tutti i concorsi locali dello stesso macrosettore nel quale hanno preso la abilitazione.

Per indurre le commissioni ad assumere il loro ruolo di "abilitatori" e scoraggiare il comportamento da "valutatori comparativi" dei candidati, una soluzione consiste nell'assicurare che la maggioranza della commissione nella sua interezza *non* sia esperta esclusiva nel settore specifico del candidato, ma esperta per un macro-settore con affinità scientifiche, e quindi in grado di valutare l'impatto della produzione scientifica di ogni candidato sulla comunità allargata. Questo si può ottenere mediante la formazione di commissioni relativamente ampie (rispetto alle attuali) che racchiudano un insieme di settori concorsuali molto vasto. Ad esempio, in area 13 si potrebbe pensare a 3 sole commissioni per 3 sub-aree: Economia, Scienze Aziendali, Metodi Quantitativi per le Scienze Sociali (denominazioni ovviamente da creare). Ogni commissione sarebbe composta da un ampio numero di docenti (da 3 a 4 volte l'attuale numerosità), selezionati con una combinazione di elezione ed estrazione che garantisca la rappresentatività della comunità.

3.4 Commissario OCSE

L'idea che l'obbligo di un commissario estero contribuisse ad allineare gli standard decisionali alle migliori pratiche internazionali si è dimostrata illusoria. Sembra questa essere una valutazione pressoché unanime. Il commissario estero si è spesso poco integrato nei lavori delle commissioni, le barriere linguistiche hanno determinato in alcuni casi effetti di estraniamento, oppure si è verificato l'impiego di metodi di valutazione e criteri propri del commissario ed estranei all'esperienza accademica italiana, se non anche dai criteri stessi previsti dalla normativa o stabiliti dalla commissione.

L'esperienza realizzata suggerisce di valutare l'eliminazione di tale figura, oppure di renderla facoltativa. In quest'ultimo caso però si ha il problema di quale soggetto abbia il potere di esercitare tale facoltà, *ex ante* la costituzione della commissione quindi in via amministrativa, oppure *ex post* la costituzione della commissione quindi in capo alla commissione stessa.

3.5 Criteri e procedure

Le commissioni hanno adottato criteri diversissimi sia nella loro dichiarazione che nella pratica decisionale, andando da algoritmi puri a mere declaratorie molto generiche.

Al contempo, la flessibilità introdotta dai cambiamenti nella normativa via circolari amministrative ha comportato ampia varietà di comportamento. Si va dalle situazioni nelle quali il soddisfacimento del superamento delle mediane è stato il criterio generale adottato, a quello opposto nel quale il criterio delle mediane è stato surclassato da criteri propri stabiliti dalle commissioni, che hanno infatti abilitato una quota significativa di candidati che non superavano le mediane e non abilitato candidati che le superavano.

È comunque generalmente condivisa l'opinione secondo la quale 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.

Qui introduciamo due opzioni differenti per una revisione di criteri e procedure, lasciando al confronto di idee e posizioni l'individuare vantaggi e svantaggi relativi.

Una prima opzione di revisione potrebbe essere quella di una regola comune e vincolante.

Ciò implica che le procedure devono essere tali da individuare ex ante criteri soglia vincolanti, trasparenti e non discrezionali su cui accertare la qualità dei candidati. Al limite, come opzione estrema si potrebbe anche ipotizzare che l'applicazione di tali criteri sia del tutto automatica, 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.

Peraltro, occorre riconoscere che proprio con riferimento alle procedure delle abilitazioni nazionali, le discussioni degli ultimi anni hanno mostrato che la varietà dei metodi di espressione dei ricercatori (i loro prodotti di ricerca) rende problematica la identificazione di criteri universali.

Tanto più che in commissioni molto ampie (come sopra ipotizzate) è facile che un qualsiasi criterio (ad es. numero di articoli su rivista referata) penalizzi gli specialisti che adottano di preferenza altre forme di produzione scientifica (ad esempio monografie). D'altra parte, l'ampiezza delle commissioni potrebbe garantire che vi siano almeno alcuni esperti della specializzazione (o specializzazioni) del candidato, che possono quindi fornire dettagli agli altri ove necessario.

Una seconda opzione, alternativa alla precedente, è quindi quella di non suggerire alcun criterio specifico, ma solo fornire tutti i dettagli necessari alla identificazione del lavoro del candidato (mediante l'accesso al suo record ANPrePS).

Per semplificare le procedure ed, allo stesso tempo, disincentivare comportamenti strategici dei commissari si possono adottare processi che identifichino facilmente i candidati più facili da valutare. Ad esempio, decisioni assunte da maggioranze sopra una soglia alta (ad es. 80-90%), che implicano ampio consenso necessariamente anche da parte di non-specialisti, sono immediatamente adottate, sia in negativo che in positivo, senza obbligo di un dettaglio specifico (ma, ovviamente, con facoltà di motivazione per la minoranza). Per gli altri casi, in cui non si forma un consenso evidente, si richiede una procedura in cui ogni commissario deve motivare esplicitamente il suo giudizio generando un certo numero (auspicabilmente limitato) di "mozioni" pro e contro la concessione della abilitazione.

Formalmente queste mozioni potrebbero essere demandate pubblicamente ad una "commissione di appello", unica e nazionale, che valuta non il candidato specifico, ma i principi alla loro base, cioè la correttezza delle mozioni. Nella sostanza, la necessità di motivare pubblicamente il proprio dissenso dovrebbe scoraggiare l'espressione di motivazioni che esulino da ambiti professionali, mentre faciliterebbe la omogeneizzazione e trasparenza dei criteri rispetto a legittime incertezze.

Inoltre, i vantaggi di un'apertura continua alle domande (opzione "a sportello") suggerita inizialmente dal CUN sono ormai ampiamente condivisi, tanto che questa modifica sarà probabilmente adottata in tempi brevi, forse inizialmente nella forma di scadenze semestrali.

Contrariamente alla abilitazione, il reclutamento locale dovrebbe tenere conto delle esigenze didattiche e di armonizzazione dei potenziali nuovi afferenti con i dipartimenti che emettono il bando. L'esigenza è tanto più forte nel caso di abilitazioni molto "ampie" che non specificano alcun profilo.

Quindi i bandi dovrebbero indicare non solo gli insegnamenti che dovranno essere coperti dal candidato selezionato, ma anche dettagli utili per soddisfare le molteplici funzioni da svolgere entro il dipartimento.

Il rischio di favorire comportamenti cooptativi iper-specificando un bando su un profilo specifico e molto raro potrebbe essere ridotto obbligando il dipartimento a specificare i risultati attesi nel

medio-lungo periodo. Ad esempio, “almeno 3 articoli nei primi 5 anni”, o “almeno 1 monografia nei primi 2 anni”. Quest’obbligo forzerebbe i bandi a contenere obiettivi compatibili con le politiche strategiche dell’ateneo, e fornirebbe il criterio base per la valutazione di conferma del nuovo docente. Quindi, i proponenti di bandi in competizione tra di loro dovrebbero impegnarsi su obiettivi sempre più ambiziosi in linea con gli obiettivi dell’istituzione. L’eventuale fallimento nel raggiungere questi obiettivi porterebbe non solo un danno alla carriera del neo-assunto, ma, cosa più rilevante a livello sistemico, anche alla caduta reputazionale del proponente del bando, che avrebbe meno probabilità di vedere accolte ulteriori richieste.

Documentazione

ANVUR (2014), “Alcune prime valutazioni sulle abilitazioni scientifiche nazionali”, ANVUR, Roma, audizione di Stefano Fantoni, presso la VII Commissione della Camera dei Deputati, 12 giugno: <http://webtv.camera.it/evento/6508>; testo scritto: [http://www.anvur.org/attachments/article/44/Audizione ANVUR 12 06 2014 ASN.pdf](http://www.anvur.org/attachments/article/44/Audizione_ANVUR_12_06_2014_ASN.pdf)

ANVUR (2013), *Rapporto Finale Valutazione della Qualità della Ricerca 2004-2010*, ANVUR, Roma, luglio, <http://www.anvur.org/rapporto/>

Baccini A. (2014), “Lo strano caso delle concordanze della VQR”, *Roars*, 10 febbraio: <http://www.roars.it/online/lo-strano-caso-delle-concordanze-della-vqr/>

Baccini A. (2013a), “Oltre la VQR”, presentazione al convegno *Evoluzione e valutazione della scienza giuridica*, Università di Firenze, 18 ottobre: http://www.slideshare.net/albertobaccini1/unifi-oltre-la-vqr-def?from=ss_embed

Baccini A. (2013b), “Napoleone, la valutazione della ricerca e le sue conseguenze prevedibili in Italia”, presentazione al convegno *La conoscenza organizzata*, Trento, 6 dicembre: http://www.slideshare.net/albertobaccini1/unitn-vqr-e-oltre?from=ss_embed

Bartolucci F. Dardanoni V., Peracchi F. (2013), “Ranking Scientific Journals via Latent Class Models for Polytomous Item Response Data”, EIEF Working Paper 13/13, Roma, maggio: <http://www.eief.it/files/2013/05/wp-13-ranking-scientific-journals-via-latent-class-models-for-polytomous-item-response-data.pdf>

Bertocchi G., Gambardella A., Jappelli T., Nappi C.A., Peracchi F. (2013), “Bibliometric Evaluation vs. Informed Peer Review: Evidence from Italy”, CSEF Working Paper n.344, Napoli, ottobre: <http://ideas.repec.org/p/sef/csefwp/344.html>

Camera dei Deputati (2014), Mozione approvata dalla VII Commissione, Cultura, Scienza, Istruzione, sui lavori delle commissioni per l’abilitazione scientifica nazionale dei professori universitari, Atto conclusivo 8-00064 del 18 giugno 2014, Roma.

Cicero T., Malgarini M., Nappi C.A., Peracchi F., (2013), “Bibliometric and Peer Review Methods for Research Evaluation: A Methodological Appraisal”, MPRA Munich Personal RePEc Archive, testo in italiano dal titolo “Metodi bibliometrici e revisione dei pari per la valutazione della ricerca: un confronto metodologico”, ottobre: http://mpra.ub.uni-muenchen.de/50470/1/MPRA_paper_50470.pdf

CUN (2014), “Proposte per la revisione delle procedure finalizzate all’attribuzione dell’abilitazione scientifica nazionale”, Adunanza CUN, 9 aprile: https://www.cun.it/uploads/4971/do_2014_04_09_proposte_procedure_asn.pdf?v

- CUN (2013a), Consultazione Pubblica per l'istituzione dell'Anagrafe Nazionale Nominativa dei Professori e dei Ricercatori e delle Pubblicazioni Scientifiche, ANPrePS, Rapporto sugli Esiti, CUN, Roma, ottobre:
https://www.cun.it/uploads/4533/rapporto_esiti_consultazione_pubblica_cun.pdf?v
- CUN (2013b), “Proposta «Criteri identificanti il carattere scientifico delle pubblicazioni e degli altri prodotti della ricerca» ai sensi art. 3-ter, comma 2, l. 9 gennaio 2009, n. 1 e successive modificazioni, Adunanza CUN, 22 ottobre:
https://www.cun.it/uploads/4532/proposta_cun_criteri_scientifici%20C3%A0.pdf?v=
- CUN (2012), “Criteri, modalità e procedure per il riconoscimento di scientificità delle riviste, Proposta CUN”, Adunanza CUN, 5 dicembre:
www.cun.it/media/118035/proposta_scientificita_riviste.pdf
- Pini P., Valente M. (2014), “Alla ricerca del Santo Graal (e come vivere senza)”, *Roars*, 16 giugno:
<http://www.roars.it/online/alla-ricerca-del-santo-graal-e-come-vivere-senza/>
- Torelli N. (2014), “Bibliometria: fin dove arrivare? Alcune riflessioni su VQR e ASN per l'area 13”, presentazione per l'area 13 all'incontro promosso dalle aree CUN 10-14 “Umaniste e sociali” sulla revisione delle procedure di valutazione della ricerca, 6 giugno, La Sapienza Università di Roma, Roma.