

# La valutazione della ricerca scientifica in Italia: una valutazione della valutazione

## 1. Premessa

L'istituzione di un sistema di valutazione della qualità della ricerca è una misura irrinunciabile per un paese che voglia investire nell'innovazione e rimanere competitivo. L'entrata in funzione dell'Agenzia per la valutazione dell'Università e della Ricerca (Anvur) dimostra che anche l'Italia si sta finalmente muovendo nella direzione della valutazione della ricerca e questo fatto non può che essere accolto positivamente. Tuttavia, l'esperienza negativa della valutazione dei progetti di ricerca nazionali (PRIN, FIRB) dimostra che fine ultimo della valutazione deve essere la premialità del merito. *Una valutazione inefficace, che non incida sulla distribuzione delle risorse, è più dannosa di una non-valutazione perché, essendo percepita come ingiusta, mina la fiducia dei ricercatori nelle Istituzioni e nella ricerca stessa.*

Una "valutazione della valutazione" della ricerca in Italia non può non entrare nel merito delle procedure attualmente seguite dal MIUR e considerarne gli aspetti metodologici.

La valutazione della ricerca scientifica riguarda:

- i progetti di ricerca
- le istituzioni (Atenei e Istituti di Ricerca controllati dal MIUR)

## 2. La valutazione dei progetti di ricerca

In Italia la valutazione dei progetti di ricerca finanziati con fondi pubblici (PRIN, FIRB) utilizza da molti anni il sistema dei revisori esterni scelti da una banca dati, quella del Cineca, sulla base dell'area e del settore scientifico disciplinare (Ssd), del panel ERC (es LS1) e del relativo settore (es.LS1\_5), della tipologia (es. docente attivo, ricercatore non universitario etc.), della residenza o meno in Italia, della partecipazione al PRIN o al FIRB corrente, dell'appartenenza alla struttura da cui proviene il progetto da valutare e dei temi di ricerca, mediante parole chiave.

I revisori sono scelti da un Garante per ciascuna area di ricerca in modo che ciascun progetto sia valutato di norma da due revisori. In certi casi, e a suo giudizio, il garante può richiedere la valutazione di un terzo revisore.

Questo sistema è inadeguato per una serie motivi che possono essere schematizzati come segue:

- a) Scarsa trasparenza del processo di revisione e del suo esito.

Il processo di valutazione di ciascun progetto è essenzialmente nelle mani del garante. Il garante non può modificare il giudizio dei revisori ma può influenzare il risultato finale sia attraverso la scelta dei revisori che in itinere, richiedendo l'intervento di un terzo o di un quarto revisore. In alcune tornate del PRIN era prevista la possibilità che i due revisori

confrontassero i giudizi e si accordassero su un punteggio finale. Pertanto, ciascun revisore, protetto dall'anonimato, deve rendere conto, quando questo è previsto, solo all'altro revisore, oltre che al garante. Il garante, a sua volta, per quanto riguarda la valutazione dei progetti individuali, non deve rendere conto a nessuno. A procedura conclusa i revisori non hanno più accesso ai giudizi ed ai punteggi da loro stessi attribuiti nè sono informati sugli esiti della revisione finale.

b) Assenza di una valutazione comparativa.

Il fatto che ciascun progetto sia valutato da due, al massimo da tre revisori e, d'altra parte, il fatto che non esista possibilità di confronto tra i progetti rende impossibile un giudizio comparativo e l'applicazione dello stesso criterio di giudizio a progetti appartenenti alla stessa area.

c) Scarsa presenza o assenza di revisori stranieri.

La banca dati dei revisori del Cineca è costituita prevalentemente o esclusivamente da revisori italiani. Il fatto che la valutazione sia affidata quasi esclusivamente a revisori italiani rende inevitabile il conflitto di interesse dei revisori.

d) Insufficiente peso attribuito alla produttività e qualità della ricerca del proponente il progetto.

A questo aspetto, che indica l'affidabilità del proponente, la sua capacità di effettuare la ricerca e pubblicarne adeguatamente i risultati, viene attribuito solo il 20% del punteggio a disposizione.

e) Dissociazione tra valutazione e budget disponibile.

La parcellizzazione e soggettività del processo valutativo fa sì che non esista alcun rapporto tra criteri di valutazione e budget disponibile per il finanziamento. Paradossalmente in vari PRIN e FIRB il numero dei progetti che hanno riportato un punteggio massimo è stato talmente elevato da far sì che la quota annuale attribuita a ciascuna unità si riducesse a livelli (media PRIN 2008-2009 : 25.000 euro, inclusa quota di cofinanziamento locale) difficilmente compatibili con la realizzazione dei progetti stessi. E' quindi essenziale che per ciascun progetto sia indicato il budget minimo per la sua attuabilità, in misura non inferiore al 30% del budget richiesto.

f) Preselezione

Nel 2012 è stata introdotta una preselezione dei progetti di ricerca a cura degli Atenei. Questa procedura è stata giustificata con la difficoltà di gestire a livello centrale un numero eccessivo di domande. In effetti, organismi nazionali (Fondazione CARIPO) e internazionali (Commissione Europea) utilizzano meccanismi di preselezione. Tuttavia in questi casi la preselezione viene di regola effettuata su una Lettera di Intenti e non sul progetto in extenso, una procedura leggera sia per i ricercatori che per le istituzioni. Una procedura come quella messa in atto impone un carico eccessivo ai partecipanti e agli Atenei, senza vantaggi dal punto di vista della qualità della selezione. Peraltro, il fatto di utilizzare nella preselezione una procedura di valutazione analoga a quella della valutazione finale, e cioè quella dei revisori

esterni scelti attraverso la banca dati del Cineca, ne riproduce, duplicandole, le inadeguatezze. *Per questo motivo riteniamo necessario che la seconda fase della valutazione dei progetti PRIN e FIRB venga effettuata utilizzando il principio della study session.*

### 2.1 La study session

Stante l'attuale congiuntura economico-finanziaria e la necessità di ridurre la spesa pubblica, un uso efficiente dei fondi per la ricerca diventa assolutamente essenziale e prioritario. Da qui l'urgente necessità di riformare radicalmente il sistema di valutazione dei progetti di ricerca. Ma in che modo? Semplicemente imitando l'esempio dei sistemi adottati da altri paesi come il sistema di revisione del National Institute of Health (NIH) americano, dell'European Research Council (ERC), delle agenzie nazionali di Francia, Germania e Regno Unito ed anche quello di fondazioni private italiane, come Telethon, AIRC, che hanno dimostrato di avere un sistema di valutazione efficiente.

Questi metodi sono basati essenzialmente sul fatto che la valutazione finale dei progetti viene effettuata da un panel di esperti dell'area cui appartengono i progetti oggetto della valutazione (*study session*), coordinato e moderato da un Presidente e da uno o più Vicepresidenti che lo affiancano ed eventualmente lo sostituiscono quando siano in discussione progetti per i quali esista un conflitto di interesse.

La *study session* consente una valutazione comparativa e l'applicazione di uniformi e comuni criteri di valutazione, tutela la trasparenza, riduce l'influenza di motivazioni personali e infine consente l'adeguamento dei criteri di selezione al budget disponibile.

Un esempio di applicazione di questo metodo è quello utilizzato dall'ANR, l'agenzia nazionale della ricerca francese e dall'Accademia delle Scienze finlandese, paesi che riservano alla ricerca quote del Pil ben più consistenti dell'Italia.

In questo sistema i ricercatori fanno domanda nelle varie aree, a ciascuna delle quali corrisponde un panel di esperti. Gli esperti, escluso il presidente ed i vicepresidenti, sono per almeno il 50% residenti all'estero. A ciascun membro del panel è affidato un certo numero di progetti (per es. 10). Per circa un terzo di questi progetti l'esperto sarà il referente (referee) nella riunione plenaria del panel, al quale dovrà brevemente descrivere il progetto in discussione, i suoi aspetti di forza e di debolezza, le eventuali revisioni da apportare e la produzione scientifica del proponente negli ultimi 5 anni. Per i restanti due terzi dei progetti a lui affidati l'esperto sarà il lettore (reader) e come tale dovrà affiancare il referente nella discussione plenaria della *study session*, esprimendo il suo giudizio sul progetto. In pratica nella discussione plenaria vengono sentiti di regola prima il referente e il lettore e quindi si apre la discussione cui possono partecipare tutti i membri del panel, ciascuno dei quali ha accesso a tutto il materiale che riguarda i progetti in esame.

La riunione plenaria è preceduta da un lavoro preparatorio di alcuni mesi durante i quali vengono raccolti i giudizi dei revisori esterni, selezionati, in numero di almeno tre da una banca dati internazionale ed, eventualmente, tra i revisori che ciascun proponente dei

progetti ha la facoltà di esprimere. E' da notare che, come i membri del panel, anche i revisori esterni sono scelti tra quelli residenti all'estero.

Ciascun membro del panel dovrà, prima della riunione plenaria, compilare per via telematica una scheda identica a quella ricevuta dai revisori esterni, nella quale dovrà attribuire un punteggio per i vari aspetti di ciascun progetto a lui assegnato e per l'attività del proponente e infine indicarne la categoria di appartenenza per quanto riguarda la finanziabilità (A, sicuramente finanziabile, B, recuperabile al finanziamento in base alla disponibilità residua, C, non finanziabile).

Di fondamentale importanza è il fatto che i revisori sono istruiti a limitare l'inserimento nelle due categorie A e B ad una percentuale prefissata (per es. max 20% in A e 20% in B).

La classifica finale viene stilata sulla base del punteggio assegnato dal panel a ciascun progetto per i due aspetti della qualità del progetto e dell'attività del proponente. Questo sistema è relativamente rapido ed efficiente e la spesa relativa alla ospitalità dei membri del panel è abbondantemente compensata dal guadagno di produttività determinato dal finanziamento dei progetti e dei proponenti più meritevoli.

## 2.2 Organizzazione della valutazione dei progetti di ricerca

Presupposto essenziale del sistema suesposto è la creazione di un organismo ad hoc, con l'esclusivo compito di allestire, implementare e monitorare la revisione e valutazione dei progetti di ricerca. L'Italia, unica tra le nazioni più industrializzate, non possiede un tale organismo e a questo non è estranea l'inefficienza della valutazione dei progetti nazionali di ricerca. Quest'organismo potrebbe essere costituito all'interno dell'Anvur ma dovrebbe essere comunque distinto dall'organismo deputato alla valutazione delle strutture.

## 3. La valutazione delle Strutture (Atenei e Istituti di ricerca)

La creazione e l'entrata in funzione dell'Anvur è da considerare di per sé un fatto positivo dato che allinea l'Italia con gli altri paesi più industrializzati. *Nonostante ciò, l'Anvur è stato fatto segno di critiche chiaramente strumentali, alcune intese non a migliorarne l'efficienza, ma a impedire tout court la valutazione stessa della ricerca.*

La valutazione delle Strutture da parte dell'Anvur è basata prevalentemente sulla valutazione dei prodotti della ricerca e quindi delle pubblicazioni, suddivise per aree di ricerca corrispondenti alle aree CUN. A questo scopo è stato costituito, per ogni area, un gruppo di esperti (Gev). Altri indicatori della qualità delle strutture sono costituiti dalla capacità di attrarre risorse e cioè, dai finanziamenti nazionali e internazionali e dall'internazionalizzazione (turnover di stranieri e presenza di collaboratori stranieri nei lavori eccellenti), dal numero di ricercatori in formazione, etc. etc.

### 3.1 La valutazione dei prodotti della ricerca

Il metodo Anvur, applicato alla valutazione dei prodotti della ricerca, si basa essenzialmente

sulla valutazione di un massimo di 3 prodotti, suddivisi per area di ricerca, per ciascun docente universitario (ricercatori inclusi) e di 6 prodotti per ciascun ricercatore degli Istituti di Ricerca per il periodo di riferimento 2004-2010.

I prodotti sono valutati sulla base del fattore d'impatto della rivista, codificato da ciascun Gev, e del numero di citazioni rilevate da Scopus o da Isi, oppure sulla base dell'esame dei prodotti effettuato da revisori esterni (peer review).

Alla peer review sono destinati tutti i prodotti pubblicati in convegni o riviste senza un codice di impatto (come è spesso il caso del settore umanistico-giuridico) e un certo numero (a campione) dei prodotti per i quali sono applicabili i parametri bibliometrici. Sulla base di questi parametri ciascun prodotto viene classificato in una di 5 categorie a ciascuna delle quali corrisponde un punteggio (eccellente, 1; buono, 0,8; accettabile, 0,5; limitato, 0; non valutabile, -1). La mancanza di un prodotto ha un peso di -0,5. Per ciascuna area, i prodotti eccellenti corrispondono al miglior 20% della produzione internazionale complessiva.

L'Anvur prevede di utilizzare per la valutazione delle Università e degli Istituti, in aggiunta ai prodotti della ricerca anche una serie di altri indicatori, relativi ad attività professionali e imprenditoriali come: prestazioni in conto terzi, creazione di spin-off, brevettazione, incubatori di impresa, Consorzi.

### 3.2. Considerazioni

Sulla base degli aspetti generali del metodo Anvur si possono fare alcune considerazioni:

- a) Il programma di valutazione dell'Anvur è piuttosto complesso e il fatto di ricorrere, nei casi previsti, anche a revisori esterni lo espone ai problemi che affliggono la valutazione dei progetti di ricerca (mancanza di trasparenza, conflitto di interessi etc. etc.). Perciò, al fine di evitare i problemi evidenziati a proposito dell'uso dei revisori esterni nella valutazione dei progetti (vedi punto 2.), sarà essenziale che ciascun Gev istituisca efficaci procedure di controllo della valutazione di ciascun prodotto da parte dei revisori esterni.
- b) Il fatto di aver fissato a 3 in 7 anni per ciascun docente (6 in 7 anni per i ricercatori degli Istituti) il numero di prodotti da sottoporre alla valutazione riducendone fortemente il potere di risoluzione, producendo un effetto livellante delle eccellenze. La possibilità di distribuire le pubblicazioni tra i coautori attenua in parte questo limite, ma penalizza le strutture con un elevato numero di pubblicazioni e/o con pubblicazioni di un singolo o pochi autori. *Si auspica quindi di aumentare il numero di prodotti utilizzabili, fermo restando il minimo di 3 prodotti in 6 anni per i ricercatori universitari. Dato che tale aumento comporterebbe, nel caso di prodotti da analizzare attraverso la peer review, un eccessivo dispendio di risorse e di tempo, bisognerebbe riservarlo alle pubblicazioni dotate di un indice di impatto e da valutare esclusivamente attraverso gli indici bibliometrici.*
- c) La fissazione del fattore d'impatto e delle citazioni al 20% superiore come requisito di eccellenza tende a livellare ulteriormente la valutazione (vedi punto b). *Si propone*

*quindi di restringere al 5% superiore il requisito per l'eccellenza.*

- d) Dato che una valutazione triennale dei prodotti (cadenza prevista quando la valutazione sarà a regime) è essenzialmente prospettica, essendo basata prevalentemente, almeno in ambito scientifico, sul fattore d'impatto della rivista piuttosto che sulle citazioni di ciascun prodotto, *si auspica che ogni 9 anni i prodotti valutabili in base alle citazioni possano venire rianalizzati in modo da fornire una valutazione reale della ricerca delle Strutture.*
- e) Il fine di una valutazione della ricerca è la distribuzione delle risorse secondo il merito. Una valutazione, per quanto accurata e ben condotta, è inutile se non influenza pesantemente il finanziamento dell'Università e degli Istituti di Ricerca. *Di questo l'Anvur dovrebbe farsi promotore e indicare le modalità applicative dei risultati della valutazione, al fine di attivare un efficace meccanismo virtuoso volto a migliorare la qualità della ricerca.*
- f) Il metodo Anvur non valuta la ricerca dei singoli docenti e ricercatori ma ne utilizza i prodotti per valutare le Strutture di ricerca. Tuttavia la valutazione individuale è necessaria ai fini del reclutamento di nuovi ricercatori e docenti e del monitoraggio della produttività individuale in vista dell'istituzione di incentivi salariali, un principio ormai di comune applicazione nella pubblica amministrazione e previsto dalla legge 240. *A questo fine la valutazione potrebbe essere effettuata, nel rispetto dell'autonomia, a cura degli Atenei.* Per la valutazione individuale potrebbero essere considerati i seguenti criteri:
- qualità, rilevanza, eccellenza, impatto e continuità della ricerca scientifica.
  - visibilità e prestigio internazionale (es. letture magistrali/invited lectures a Congressi internazionali),
  - ruolo come opinion leader nel suo campo (membro di commissioni di valutazione internazionali; partecipazione a Consigli scientifici istituzionali)
  - capacità di attrarre risorse finanziarie finalizzate all'autofinanziamento

*Gaetano Di Chiara e Maria Grazia Roncarolo per il Gruppo 2003 \**

20 giugno 2012

\* Il Gruppo 2003 è costituito dagli scienziati italiani, residenti nel nostro Paese, più citati nelle riviste internazionali  
[www.gruppo2003.org](http://www.gruppo2003.org)  
[www.scienzainrete.it](http://www.scienzainrete.it)