

Andrea Bonaccorsi

*Questioni teoriche della valutazione della ricerca nelle discipline
umanistiche e sociali*

1. *Introduzione*

La valutazione della ricerca è un processo istituzionale ma ha dimensioni epistemiche e sociali che richiedono continue elaborazioni, sia teoriche che operative, da parte di comunità scientifiche attive, critiche e partecipi. Le istituzioni che sono deputate alla valutazione o che la utilizzano in modo più intenso (nel nostro caso, il MIUR e l'ANVUR), ciascuna nell'ambito delle proprie prerogative, devono mantenere un canale continuo di dialogo con le comunità scientifiche, anche esponendosi alle critiche – e, aggiungo, in una misura molto maggiore di quanto accada in altri ambiti della pubblica amministrazione.

Per quanto mi riguarda non ho cessato, dopo la fine del mandato (2011-2015) in ANVUR, di studiare il problema della valutazione della ricerca nell'ambito delle discipline umanistiche e sociali (SSH nel linguaggio internazionale, *Social Sciences and Humanities*). In questo contributo cerco di fare il punto su alcuni temi aperti, ampliando le analisi recentemente pubblicate presso Il Mulino (Bonaccorsi, 2015) e presso Springer (Bonaccorsi 2018).

Le domande preliminari che pongo sono, nell'ordine:

- (i) È possibile giungere ad un consenso nella comunità scientifica in riferimento ad un insieme di criteri di qualità della ricerca?
- (ii) Se si risponde positivamente alla prima domanda, è possibile formulare tali criteri una forma riflessiva e dichiarativa, in modo che possano essere sottoposti a verifica intersoggettiva e possano essere interpretati correttamente da un sottoinsieme della comunità (ad esempio, da esperti)?
- (iii) Se si risponde positivamente anche alla seconda domanda, è possibile e utile sintetizzare in una misura quantitativa il giudizio qualitativo sulla qualità della ricerca?

Una volta argomentate, cerco di affrontare alcune delle critiche più recenti alla valutazione nelle SSH, in particolare l'argomento della minaccia al pluralismo

epistemico, della distorsione nei prodotti della ricerca umanistica e della integrità della ricerca.

2. Le domande preliminari

2.1 Consenso su criteri di qualità della ricerca

Chiunque lavora in ricerca possiede un insieme di criteri in base ai quali riconosce la qualità di un contributo scientifico. Ciò accade per i propri lavori (ciascuno di noi sa bene quali sono i suoi lavori migliori) ma anche per quelli degli altri. Su questo punto sembra difficile non convergere.

La domanda si fa però più ardua se ci chiediamo se sia possibile giungere ad un consenso intersoggettivo su tali criteri, oppure se essi siano irriducibilmente soggettivi, personali, non comparabili, idiosincratici. In fondo si potrebbe qui riadattare la formula di Agostino sulla percezione del tempo: Se non mi chiedi cosa è il tempo lo so, se me lo chiedi non lo so più.

A sostegno di una risposta negativa alla domanda sul consenso sui criteri si possono avanzare diversi argomenti.

In primo luogo le scienze umane e sociali sono caratterizzate da conflitti paradigmatici, spesso acuti. Tra i più noti basti pensare al conflitto tra psicologia clinica e psicologia sperimentale, tra filosofia analitica e tradizione continentale, tra economia neoclassica ed economia *non mainstream* (Viola 2017). Si tratta di casi nei quali non vi può essere consenso sui criteri di qualità, perché essi sono esattamente una parte essenziale del paradigma, inteso nel senso di Kuhn. Seguire paradigmi diversi significa quindi, secondo questo argomento, aderire a criteri di qualità in conflitto tra loro. Per fare degli esempi concreti, se si ritiene che un argomento filosofico non può essere sostenuto senza una rigorosa ricostruzione storica dei suoi antecedenti, si riterrà che un articolo analitico che utilizza la logica ma non cita gli autori classici non sia di buona qualità. Se si ritiene che l'economia politica debba procedere per modelli matematici, un articolo che utilizza modelli di simulazione verrà ritenuto di livello inferiore. Riprenderò più tardi le implicazioni del pluralismo epistemico ai fini delle scelte operative della valutazione.

Su un piano meno elevato si può osservare che, in parte come conseguenza di pluralismo epistemico, in parte per dinamiche accademiche nazionali e locali, si formano scuole di pensiero in conflitto tra loro¹. Poiché la valutazione ha delle

¹ Ho appreso da alcuni colleghi giuristi che una componente della formazione dei giovani ricercatori nelle discipline giuridiche ha a che fare con quali autori devono essere citati e quali autori devono essere accuratamente evitati, a pena di conseguenze serie per la carriera scientifica. Spero sia una informazione parziale.

conseguenze, dirette o indirette, sulla allocazione delle risorse, sarà impossibile far convergere scuole di pensiero diverse sullo stesso insieme di criteri, perché ciascuna delle parti in campo temerà che se ne avvantaggi l'avversario accademico.

Un argomento un poco più nobile ha a che fare con una posizione epistemologica radicale, che afferma la incommensurabilità della conoscenza, che sarebbe sempre irriducibile a schemi di confronto con altri elementi di conoscenza. La conoscenza sarebbe solo di casi singoli. Si tratta della radicalizzazione del metodo idiografico, già praticato nella filosofia greca in contrapposizione ai grandi sistemi aristotelico e platonico.

Infine si potrebbe temere che, attraverso il consenso sui criteri di qualità si facciano passare surrettiziamente operazioni di restrizione della libertà di ricerca. In fondo, si sostiene, chi può dire quale sarà la buona ricerca in futuro? Vincolarsi a dei criteri oggi potrebbe significare scoraggiare la esplorazione di nuove direzioni di ricerca, limitando l'avventura intellettuale che è parte costitutiva del mestiere di studiosi. Si potrebbe in altre parole imporre il conformismo intellettuale.

Trovo questi argomenti seri e credo che si debba ingaggiare un confronto intellettuale alto e motivato, supportato dalla conoscenza diretta e di prima mano della letteratura specializzata, ma anche dei dibattiti filosofici, politici e in senso lato culturali che stanno dietro a questi argomenti.

Un primo controargomento muove dalla osservazione che il conflitto paradigmatico e, in misura diversa, anche il conflitto tra scuole accademiche non precludono di per sé la convergenza su un insieme di criteri condivisi. Si può convergere su alcune regole minime del gioco e divergere su tutto il resto. In *La valutazione possibile* esamino il caso della storia, una delle discipline umanistiche nella quale il conflitto politico, culturale, valoriale è, per natura, più acuto. Grandi controversie storiche hanno attraversato tutto il Novecento, e ancora animano in vario modo il dibattito nelle società avanzate, su temi come l'identità nazionale, la memoria della guerra, i totalitarismi. Eppure uno studio attento del dibattito interno alla comunità degli storici² mostra un fatto sorprendente: esiste un robusto nucleo di criteri di qualità della ricerca che viene condiviso da storici di ogni estrazione. Questi si riferiscono principalmente alla qualità dell'esame delle prove o degli indizi storici, ai metodi di consultazione degli archivi, a tecniche di datazione e a criteri di completezza delle fonti.

La convergenza su criteri di qualità nei metodi di raccolta dei dati è presente anche in altre discipline umanistiche e sociali. Kagan (2009) sostiene addirittura che la convergenza su tali criteri nella psicologia sia all'origine dello status che questa

² In *La valutazione possibile* esamino tre episodi cruciali del dibattito interno alla comunità degli storici: la reazione al caso Paul de Man, il revisionismo sulla Shoah, la reazione alla tesi post-strutturalista di Hyden White. Nel dialogo con Carlo Olmo discutiamo il ruolo della microstoria di Ginzburg e Levi, del paradigma indiziario e della discussione storiografica intorno al regime della prova.

disciplina ha acquisito, a fronte della difficoltà di altre scienze sociali (come la sociologia) nella c.d. crisi delle scienze sociali. De Ketele e Paquay (2011) esaminano un caso simile, le discipline pedagogiche.

Non si tratta di imporre un insieme di regole valide per tutte le scienze umane e sociali o di rincorrere una impossibile unità della scienza. Si tratta, al contrario, di osservare il grado di maturazione del dibattito intra-accademico che verte sui criteri di validità delle asserzioni. Tale dibattito non coincide necessariamente con la riflessione epistemologica formalizzata³, ma ha più spesso a che fare con contingenze specifiche delle discipline, dibattiti e contenziosi aperti per la necessità di difendere lo statuto della disciplina, o anche eventi esterni. Si pensi all'imponente dibattito storiografico aperto dalla creazione di grandi musei di storia del Novecento in Germania e in Austria, fino ad epoca recente. Oppure, nelle scienze sociali, alla riapertura del dibattito metodologico sulla economia politica dopo la Grande crisi iniziata nel 2008. Fatto sta che le diverse discipline umanistiche e sociali hanno affrontato questi temi con diversi gradi di intensità e profondità.

In un importante studio sui criteri con cui i ricercatori valutano i progetti di ricerca, Michèle Lamont ha mostrato che le discipline più dinamiche, che hanno subito meno la crisi delle immatricolazioni e dei finanziamenti, sono quelle caratterizzate da un attivo dibattito teorico e metodologico sulla validità delle asserzioni della disciplina e sui criteri di qualità della ricerca.

Avanzo quindi la tesi secondo la quale le comunità scientifiche che rispondono negativamente alla domanda posta all'inizio possono certamente evitare spinose discussioni, ma finiscono per pagare nel medio termine un prezzo molto elevato in termini di chiusura.

2.2 Natura dichiarativa dei criteri di qualità della ricerca

Anche ammettendo la possibilità di un consenso sui criteri, si potrebbe argomentare che esso si eserciti solo all'interno delle comunità scientifiche in senso stretto. In altri termini, non ci sarebbe bisogno di esplicitare e formalizzare i criteri, ma solo di consentire a gruppi di volta identificati di studiosi (comitati, commissioni, *panel*, consigli vari) di prendere decisioni.

³ Occorre tuttavia riflettere sulla circostanza che l'epistemologia del '900 è quasi interamente basata sulle scienze dure, inizialmente fisica, più recentemente biologia, con aree specializzate in filosofia della matematica o dell'informatica e della medicina. Una riflessione epistemologica di insieme sulle discipline umanistiche è appena iniziata (si veda BOD 2013 per una analisi introduttiva). Nelle scienze sociali la riflessione è più avanzata ma notevolmente frammentata (WALLISER 2009; 2015).

Si potrebbe sostenere infatti che i giudizi di qualità sono inevitabilmente idiosincratici e formano una base di conoscenza personale, inseparabile dalle persone che la detengono. Quando una persona esperta vede un caso singolo, allora formulerà un giudizio sul quale potrà trovare il consenso di altri esperti, senza alcun bisogno di esplicitarne le ragioni. Una interessante conseguenza pratica di questa tesi è che l'unica procedura legittima per la valutazione della ricerca sarebbe la delega alle società scientifiche (nella ipotesi che esse siano rappresentative della comunità scientifica) delle decisioni di nomina degli esperti.

Questo problema è stato esaminato magistralmente da Polanyi con la nozione di *conoscenza personale*. Esiste un residuo di conoscenza non codificata, impossibile da esprimere in un linguaggio formale, e quindi da comunicare attraverso mezzi generali. Si tratterebbe di una particolare forma di conoscenza tacita, che gli studiosi guadagnano solo dopo lungo tirocinio, come effetto della lunga accumulazione di esperienza e di frequentazione personale dei rispettivi maestri.

Ora è possibile accettare la tesi di Polanyi in generale (ad esempio essa spiega molti rilevanti fenomeni in economia), ma traendone conseguenze del tutto diverse. La codificazione della conoscenza è infatti un processo dinamico, storico e sociale. Il confine tra ciò che può essere codificato e ciò che resta tacito, implicito, non verbalizzato non è fissato una volta per tutte. Il confine si sposta per effetto di azioni deliberate e istituzionalizzate.

Si può infatti considerare la scienza moderna come il principale processo di avanzamento del dominio della conoscenza codificata. Secondo il *Novum Organum* di Bacone, «l'unico esperimento che vale è quello scritto». La scienza moderna nasce non solo con la invenzione degli esperimenti, cioè con Galileo, ma anche con la invenzione delle riviste scientifiche, cioè delle sedi ufficiali nelle quali si mette per iscritto la descrizione dell'esperimento, in modo che il lettore possa, almeno mentalmente, ripeterlo, anche se non era testimone oculare. La scienza di corte, come è noto, procedeva per circoli ristretti di testimoni diretti, invitati ad assistere alle prove sperimentali al fine di poter certificare, con la propria autorevolezza e credibilità, ciò che era effettivamente accaduto. La scienza moderna si basa, al contrario, su estesi processi di codificazione e formalizzazione, che consentono in linea di principio la validazione intersoggettiva da parte della comunità scientifica (Ziman 2005).

Il che non significa negare l'importanza per la scienza della componente pre-scientifica, estetica, valoriale o emotiva, ma distinguerla da quella dichiarativa. Nel magnifico *The book of why* Judea Pearl distingue tra assunzioni (che devono essere dichiarative) e conoscenza (che può avere anche componenti implicite). È solo quando le assunzioni “vedono la luce del giorno” che la scienza inizia a operare (Pearl 2018).

Alla tesi della natura non codificata dei criteri di qualità si può anche obiettare che essa tenda a perpetuare (magari con nobili motivazioni) la presa del potere accademico

più tradizionale basato sull'opacità. La assenza di criteri formalizzati di qualità finisce infatti per mantenere come unica fonte di giudizio coloro che hanno l'esperienza prolungata per maturare criteri taciti. A tale giudizio non si potrebbe opporre nulla.

2.3 Comparazione e commensurazione

La terza questione riguarda la possibilità di passare dal consenso su formulazioni dichiarative di tipo qualitativo a misure quantitative. A questa pretesa molti oppongono una serie di argomenti molto seri.

Si può sostenere, ad esempio, che esista uno scarto tra giudizio qualitativo e misurazione. Si può dire che un certo articolo A dell'autore X è migliore dell'articolo B dell'autore Y (oppure anche dell'articolo C dello stesso autore X), ma è impossibile dire *di quanto* sia migliore. La pretesa di trasformare un giudizio qualitativo, che distingue al massimo ordinando in grandi categorie, in un giudizio quantitativo, sarebbe una forzatura.

Dopo tutto, nelle scienze sociali si distingue in modo netto tra variabili categoriali, ordinali o cardinali.

Inoltre ci si potrebbe chiedere *perché* si cerca di trasformare giudizi qualitativi in quantitativi. Qui la critica si fa filosofica e politica: la quantificazione sarebbe un dispositivo, nel senso di Foucault, ovvero uno strumento di governo della società che non passa attraverso ordini espliciti, ma viene veicolato con la interiorizzazione di regole che appaiono naturali, evidenti, inoppugnabili. Cosa c'è nella società di più oggettivo, quindi apparentemente neutrale, del numero? Assegnare dei numeri al merito delle ricerche significa fornire una certificazione che legittima e consolida le gerarchie esistenti.

Alla critica nel senso di Foucault si potrebbe aggiungere un argomento nella tradizione sociologica di Pierre Bourdieu: gli scienziati sono attori sociali che lottano tra di loro per la conquista di un campo, o sfera di influenza. La quantificazione è funzionale alla costituzione di credenziali da spendere nella competizione all'interno del campo dei titoli di istruzione superiore (quindi dell'accesso agli strati elevati della società) e della autorità scientifica.

Infine, in parte collegato con le critiche ora indicate, in parte indipendente, vi è l'argomento secondo il quale la quantificazione sarebbe una forma di mercificazione, di riconduzione del valore ad un unico metro, esattamente come avviene nel sistema di mercato. Quindi la quantificazione sarebbe uno strumento del neoliberismo, inteso come il movimento politico e di interessi economici che predica la estensione del dominio del mercato ad ambiti via via più vasti della società, da sottrarre all'intervento pubblico. Il capitalismo neoliberale utilizza la quantificazione come premessa della

mercificazione (del feticcio delle merci, per riprendere la celebre immagine marxiana) e dell'assoggettamento delle persone.

Va detto che i controargomenti su questo punto devono essere articolati. Certamente la trasformazione delle qualità in quantità, ovvero la commensurazione (*commensuration*) è un processo cognitivo che ha condizioni di validità restrittive. Ma anche in questo caso non esiste una differenza ontologica tra qualità e quantità. In particolare se si cambiano i livelli di osservazione, un processo molto comune nella scienza. Possiamo trovare un accordo qualitativo su cosa intendiamo per “avere gli occhi azzurri”. Per passare da questo giudizio ad una misura possiamo compiere due mosse. Una è acquisire una teoria scientifica del colore, che ci dia una misura, condivisa dagli scienziati, sullo spettro di luce che si può convenzionalmente definire azzurro, sulla natura della materia di cui sono fatti gli occhi etc. Nel campo della valutazione della ricerca, tuttavia, questa teoria potrebbe non essere disponibile o avere condizioni restrittive di applicabilità. Ad esempio potremmo concordare con l'idea che una buona ricerca è originale o creativa, ma mancare di una teoria della creatività sufficientemente precisa da consentire una misura. Allora potremmo seguire un'altra strada, che è spostare il livello di osservazione alla popolazione. Potrei non avere la certezza che la persona A ha senza alcun dubbio gli occhi azzurri, ma potrei accontentarmi del fatto che, osservando un certo numero di persone a caso, due o più osservatori concordino sul numero di coloro che hanno gli occhi azzurri a meno di una differenza finita. L'attenzione si sposterebbe allora su quanto grande sia questa differenza ed eventualmente sui criteri per ridurla: se per un osservatore ritiene che in un gruppo vi sia il 20% di persone con gli occhi azzurri, mentre per un altro nello stesso gruppo ve ne sono l'80% dobbiamo concludere che il criterio “occhi azzurri” non è sufficientemente definito. Ma se invece osserviamo una convergenza non su singoli casi ma su insiemi sufficientemente grandi di casi, allora il passaggio al livello della popolazione è legittimo. Ciò è esattamente ciò che accade con la valutazione basata sul giudizio di esperti.

Qui è utile soffermarsi sulle ragioni per cui riteniamo che la bibliometria non si applichi, se non marginalmente, nelle discipline umanistiche e sociali, mentre invece è accettabile nelle scienze dure. La bibliometria misura la qualità della ricerca utilizzando un indicatore (citazioni) opportunamente standardizzato e normalizzato. Perché le citazioni sono un indicatore accettabile in un caso, poco appropriato nell'altro? Non perché le discipline umanistiche facciano a meno delle citazioni (tutt'altro), ma perché partecipano in misura diversa di una caratteristica che è invece centrale nelle scienze dure, ovvero la *cumulatività*. Nelle scienze dure è accettabile che la citazione sia un indicatore di qualità, perché citando un lavoro recente l'autore “riassume” tutta la scienza precedente che è rilevante per il suo argomento. Non ha bisogno di citare tutti gli scienziati che hanno condotto allo stato dell'arte, perché questi sono impliciti, in

virtù della cumulatività, nella citazione più recente. Per la stessa ragione il formato tipico delle scienze dure è l'articolo, cioè un testo che impone severi limiti al numero delle citazioni. Quindi gli autori dovranno accuratamente scegliere chi e cosa citare. A fronte di cumulatività e di selettività delle citazioni è possibile procedere a regole di standardizzazione e di normalizzazione – esattamente le premesse metodologiche per una accettabile quantificazione. Nelle discipline umanistiche la cumulatività ha natura diversa, non solo perché gli autori citati sono per definizione molto spesso defunti (il che crea un problema difficilissimo di normalizzazione), ma anche perché una parte rilevante del lavoro ha per oggetto la ridefinizione del passato, la riapertura di temi, la ricapitolazione, la messa in prospettiva storica. Serve uno spazio più ampio (la monografia) e le citazioni non hanno un significato univoco.

Quindi occorre passare dall'idea che il mondo si divide nettamente tra qualità e quantità alla questione di quali procedure siano corrette per trasformare l'una nell'altra senza perdere informazione preziosa.

In una prospettiva storica, insisto sul fatto che la quantificazione è stata, ed è tuttora, un potente strumento di riduzione dei poteri tradizionali basati su conoscenze idiosincratiche e di affermazione dei diritti di libertà moderni. La statistica sociale moderna non nasce, come sostiene una lettura parziale di Foucault, per irreggimentare la devianza, ma per estendere i diritti sociali e il sistema del *welfare* nelle società democratiche occidentali.

Su un piano filosofico, occorre notare che gli orientamenti più recenti della teoria delle decisioni e della filosofia pratica suggeriscono che gli spazi per la commensurazione siano decisamente più ampi di quanto si ritenesse fino a qualche decennio fa (ad esempio in seguito al celebre teorema di impossibilità di Arrow).

La condizione perché la commensurazione sia fruttuosa è che le comunità condividano un linguaggio sufficientemente ricco e articolato per formulare giudizi qualitativi sui quali esercitare il consenso. Se ciò avviene e le categorie semantiche veicolate dal linguaggio sono articolate, allora una procedura di commensurazione può essere iniziata. Essa deve sottostare a regole formali di contenuto e di metodo.

3. Le discipline umanistiche e sociali di fronte alla valutazione. Problemi epistemici

La tesi che sviluppo nei due volumi citati nasce dalla osservazione che le varie discipline umanistiche e sociali formulano risposte molto differenziate alle tre domande poste sopra, in uno spettro che va da tre sì convinti a tre no altrettanto convinti, con una ulteriore modulazione sulla intensità della reazione (in alcuni casi, diciamo, piuttosto indignata).

Naturalmente qui assumo come disciplina una stilizzazione delle posizioni prevalenti nelle comunità scientifiche, come emergono da documenti, convegni, *position paper*, e come le ho potute apprezzare anche sul campo durante l'esperienza istituzionale. Non mi sfugge la presenza di differenze interne, sulle quali tornerò. Infine la tesi che propongo non deriva esclusivamente dalla esperienza italiana, ma prende in carico il forte dibattito dell'ultimo ventennio, soprattutto in area francese, inglese e americana, sulla valutazione della ricerca, la crisi delle discipline umanistiche, le politiche pubbliche per l'università e la ricerca. Le reazioni sperimentate in Italia non sono un caso isolato.

Ecco in sintesi la mia tesi. Le discipline nelle scienze umane e sociali percepiscono e recepiscono la valutazione della ricerca in modo positivo o negativo, e con differente intensità, in funzione dei seguenti fattori:

- (a) Storia della istituzionalizzazione come disciplina accademica;
- (b) Orientamento epistemico generale;
- (c) Accettazione o rigetto di due tradizioni dominanti nella epistemologia del XX secolo.

Di conseguenza occorre ingaggiare con le comunità scientifiche un dibattito epistemico elevato, per comprendere le specificità e i problemi aperti. Vediamo le dimensioni del modello.

3.1 *Istituzionalizzazione come disciplina accademica*

Premessa. Le discipline scientifiche emergono per una duplice dinamica: una dinamica epistemica, che ha a che fare con la emergenza di oggetti di indagine, metodi, tecniche, e una politico-istituzionale, che dipende invece da complesse vicende nazionali o internazionali, finanziamenti, azioni di governi o di leader scientifici.

Le nuove discipline seguono dinamiche epistemiche e istituzionali differenziate. Nella fase iniziale di istituzionalizzazione accademica si osserva un dibattito, spesso acceso, sui confini disciplinari, ovvero sulle differenze rispetto alle discipline confinanti. Suggesto che esista una differenza importante tra i casi nei quali la istituzionalizzazione deriva dalla emergenza della disciplina da differenziazione interna, oppure da conflitto con una disciplina dominante. Nel primo caso (gemmazione) la nuova disciplina nasce con uno statuto definito, spesso su iniziativa, o con il consenso, di studiosi attivi nella comunità di origine. O i confini vengono definiti una volta per

tutte all'inizio e raramente messi in discussione. La nascita della storiografia accademica, o della filologia, o della economia politica seguono questa dinamica.

Nel secondo caso, al contrario, la nuova disciplina nasce da un conflitto epistemico ed è costretta a mantenere a lungo una coda polemica nei confronti della disciplina dalla quale prende le distanze. I confini sono sempre problematici e devono essere affermati e riaffermati contro tentativi di riassorbimento, marginalizzazione o esclusione. La nascita della antropologia in opposizione alla sociologia positivista, la nascita della critica letteraria in opposizione alla filologia, la separazione degli *English Studies* dalla critica letteraria basata sul canone occidentale possono essere lette in questa chiave.

Le discipline, attraverso meccanismi istituzionali (es. società scientifiche, riviste, comitati editoriali di collane di libri) preservano i confini che sono stati stabiliti e «tengono memoria» dei conflitti iniziali. Le definizioni disciplinari sono periodicamente (ad intervalli lunghi) soggette a controversie a causa dell'emergenza di nuove prospettive filosofiche o metodologiche, che obbligano ad un ripensamento radicale e conducono ad una accettazione/rigetto.

La mia tesi è che le discipline nate “per gemmazione” hanno avuto un dibattito interno con acquisizioni teoriche e metodologiche stabili, che consentono di reagire positivamente alle pretese della valutazione della ricerca. Al contrario le discipline nate “per conflitto” mantengono una duratura attitudine a sospettare, nella valutazione della ricerca, la fonte di operazioni di riduzione della loro autonomia e specificità.

3.2 Orientamento epistemico generale

Qui si tratta della distinzione consolidata tra orientamento idiografico e nomotetico. Come affermano due celebri scienziati politici: «Questo tipo di tensione disciplinare è familiare: tra campi idiografici e nomotetici, tra studiosi del particolare che si lamentano della attitudine dei teorici a trascurare complessità cruciali al fine di assimilare la storia all'interno dei loro schemi esplicativi preferiti, e studiosi del generale che si lamentano del fatto che i colleghi del particolare sono così immersi nello studio dei dettagli da perdere di vista il quadro teorico generale» (Tetlock e Lebow 2001).

Non vi è dubbio che le discipline a orientamento nomotetico (ad esempio la scienza politica empirica di tradizione americana, o la sociologia quantitativa) recepiscono la valutazione quasi come un prolungamento, con altri fini, dei propri metodi. Maggiore difficoltà viene percepita dagli studiosi che si occupano sistematicamente di casi singoli, in quanto la comparabilità e la generalizzazione del giudizio di qualità della ricerca appaiono, in prima istanza, poco credibili. In queste

discipline si accetta la valutazione solo a fronte di un dibattito interno prolungato, come nel caso discusso sopra della storiografia del Novecento.

3.3 Reazione alle sfide epistemologiche del Novecento

Con una semplificazione estrema, che naturalmente non rende ragione dei temi in discussione, suggerisco che siano due le sfide poste dalla filosofia del Novecento alla organizzazione delle scienze sociali e umane: il positivismo logico e il post-strutturalismo.

Secondo il positivismo logico proposto dal Circolo di Vienna di Schlick, Neurath e Carnap, e in seguito da autori quali Reichnbach, Hempel e Ayer, la verità deve essere definita come corrispondenza di proposizioni a fatti (dati) e i dati di senso sono la fonte ultima delle esperienze (empirismo logico). In questo quadro non vi è una distinzione di principio tra scienze naturali e scienze umane e sociali, in quanto ogni conoscenza scientifica deve essere intesa come scoperta di regolarità esprimibili attraverso proposizioni in forma di legge (*lawlike regularities*). Inizialmente formulato in riferimento alle scienze dure, esso inizia a influenzare le scienze sociali negli anni '30 e soprattutto nel secondo dopoguerra. Alcune scienze sociali (in particolare, la psicologia sperimentale e cognitiva, la sociologia americana, la scienza politica empirica, la politologia che si ispira alla teoria dei giochi o alla Public Choice) si mostrano pronte ad accogliere la sfida di dimostrare uno statuto scientifico in nulla inferiore rispetto alle scienze dure.

Di segno opposto il movimento post-strutturalista legato a filosofi, sociologi e intellettuali francesi (Derrida, Foucault, Lacan, Barthes, Lyotard, Baudrillard) e al loro impressionante successo di opinione negli ambienti accademici americani a partire dagli anni '60-'70. Assumendo una versione estrema di Nietzsche e di Wittgenstein questi autori insistono che non esistono fatti, esistono solo narrazioni; non esistono verità, esistono solo convenzioni linguistiche. Il sistema sociale utilizza forme di dominio (*dispositivi*) che normalizzano i comportamenti dei soggetti attraverso la interiorizzazione di categorie assunte come naturali, mentre le gerarchie di saperi non hanno alcun fondamento epistemico. Di conseguenza ogni prodotto scientifico deve essere interpretato come un testo e non vi sono differenze tra testi di varia natura: ad esempio, secondo la celebre formulazione di Hayden White, il libro di uno storico non contiene alcuna verità storica, ma solo una particolare narrazione, indistinguibile, in linea di principio, da una narrazione letteraria.

Aderiscono a questa linea alcune discipline umane e sociali moderne: una importante componente della antropologia basata sull'ermeneutica, le critiche letterarie

anglosassoni (*English Studies*) e le numerose discipline emerse in seguito all'ondata post-strutturalista come *Cultural Studies*.

Devo dire che, dopo aver lavorato molto nel tentativo di studiare in dettaglio la critica devastante alla valutazione che ha utilizzato autori post-strutturalisti o post-moderni (in particolare, come è noto, Foucault) e di identificarne principi epistemologici generali, sono stato spiazzato dalla recentissima analisi che Elio Franzini ha dedicato al problema. Franzini depotenzia completamente il contenuto epistemologico del post-moderno e lo riporta in un ambito estetico. Scrive infatti: «Pensare a uno stile epistemologico comune del post-moderno è un errore interpretativo... il post-moderno è una strategia di pensiero che sottolinea [...] la dimensione legata alla persuasione, cioè alla retorica, all'evento, all'estetica o, meglio, al mondo delle arti (p. 166). Sul piano epistemologico le affermazioni postmoderne hanno in effetti un valore spesso risibile, che riveste solo un vago senso etico, retorico, estetico (p. 171). Non va dimenticato che questa galassia [...] non solo non ha intenti epistemologici o ontologici, ma è nata per decostruire le "immagini del mondo" che hanno accompagnato e perseguitato il pensiero occidentale "moderno" (p. 177)» (Franzini 2018). Confesso che mi era venuto il sospetto ma leggerlo in uno dei più autorevoli studiosi fa un certo effetto.

Conforta anche trovare un giudizio critico in una recente analisi sistematica delle teorie della letteratura:

«.. Si tratta di un tipico pregiudizio post-modernista, secondo cui qualunque pretesa che esistano valori universali sia in sé reazionaria perché lesiva dei diritti delle minoranze e delle particolarità. Va da sé che queste posizioni prestano all'universalismo solo intenzioni arroganti e omologanti e trascurano quanto di destabilizzante e al limite rivoluzionario ci sia nella volontà di valorizzare quel che ci unisce come specie umana (come appunto fa l'arte di qualunque epoca e provenienza) rispetto a quanto ci divide» (Brugnolo *et al.* 2016, p. 26). E ancora: «... Siccome la distinzione tra vero e falso è spesso problematica e aperta a revisioni e correzioni allora essa viene ricusata come insussistente: nessun discorso è vero, nessun discorso è falso. Il che significa ancora una volta forzare una intuizione ragionevole: una cosa, infatti, è dire che i fatti ci vengono sempre mediati dalle parole (vero) e un'altra che essi non esistano al di fuori delle parole (falso)» (*ibid.* p. 61).

3.4 Dibattito epistemico e attitudine verso la valutazione

Quali implicazioni derivano da queste differenze epistemiche?

Innanzitutto direi che la valutazione fa esplodere le contraddizioni interne e/o i ritardi delle discipline nazionali rispetto al dibattito metodologico internazionale. In un

recente lavoro ho esaminato lo strano caso di sociologia e delle discipline aziendali, accomunate dal fatto di aver avuto una valutazione molto severa già nella VTR del 2001-2003, e di aver attraversato un intero decennio senza apprezzabili miglioramenti, al punto di aver ricevuto valutazioni deludenti nella VQR 2004-2010 (e, va aggiunto, solo di poco migliori nella VQR 2011-2015). In queste discipline si è assistito ad un dibattito che ha accusato la valutazione di attentare al pluralismo paradigmatico o alla rilevanza pratica della ricerca e di attuare una (inesistente) discriminazione valutativa a favore dei metodi di ricerca quantitativi rispetto ai metodi qualitativi. Si è trattato di una reazione difensiva e tutta politicista, mentre nell'insieme è mancato un intervento (serio) di miglioramento degli aspetti metodologici e di raccolta dei dati delle discipline, almeno rispetto agli standard internazionali (Bonaccorsi 2016).

Nelle aree umanistiche il dibattito ha messo in campo argomenti più fondati. Nel dibattito internazionale le reazioni delle comunità scientifiche alla valutazione sono state più intense e di segno negativo in presenza di modelli epistemici condivisi di tipo idiografico, una significativa influenza della critica post-strutturalista e conflitti epistemici irrisolti. Non è un caso, mi pare, che tra gli autori più attivi contro la valutazione compaiano a livello internazionale gli studiosi di letteratura inglese e americana (*English studies*), che combattono contro la gerarchizzazione delle opere letterarie imposta dal canone occidentale (nel senso di Bloom), gli antropologi e, in area francese, studiosi di letteratura comparata e psicologi di orientamento clinico e psicoanalitico.

Al contrario la ricezione della valutazione è stata perlopiù positiva in presenza di basi metodologiche condivise, rigetto della critica post-strutturalista, conflitti paradigmatici gestiti sul piano delle metodologie di raccolta dei dati e di un continuo confronto inter-teorico.

4. *Il problema del pluralismo epistemico*

In un recente lavoro Viola (2017) ha sostenuto, contro le mie tesi espresse nel testo de *Il Mulino*, che la valutazione finisce inevitabilmente per limitare il pluralismo epistemico, definito come la compresenza di due o più scuole rivali di pensiero all'interno nello stesso dominio di ricerca. Per scuole rivali Viola intende distinti gruppi di ricerca che sostengono assunzioni e impegni concettuali e/o metodologici in conflitto tra loro, ma sono in competizione per spiegare lo stesso insieme di fenomeni.

A questa tesi opporrei argomenti teorici e empirici, da riprendere in altra sede. Qui mi limito ad alcuni cenni. Sul piano epistemologico direi che la competizione tra scuole di pensiero è costitutiva della ricerca nelle SSH e segue dinamiche di *topic generation* e di aggregazione epistemica e sociale del tutto indipendenti dalla valutazione. Fa parte di

tali dinamiche la possibilità che si formino *mainstream* dominanti e posizioni minoritarie. La valutazione in questi casi deve assumere la difesa del pluralismo epistemico come obiettivo principale, ma non è responsabile delle dinamiche interne alle discipline.

Come altri autori, Viola è scettico su questo punto. Sostengo che non vi è alcuna evidenza sistematica che la valutazione abbia prodotto la eliminazione di scuole di pensiero. I casi citati non sono convincenti. Si prenda l'economia politica: l'economia neoclassica ha generato un *mainstream* per poderosi motivi epistemici (modellizzazione, matematizzazione, raffinamento delle tecniche econometriche) e politico-economici (formazione delle classi dirigenti, istituzioni internazionali, premi Nobel), su cui si possono avere opinioni molto diverse, ma che in ogni caso sono indipendenti da (e anche enormemente più potenti di) qualunque processo di valutazione della ricerca. Allo stesso modo la discussione su filosofia analitica vs continentale, o la antica querelle sulla natura della psicoanalisi sono fenomeni di scala talmente elevata da dubitare che i processi istituzionalizzati di valutazione possano avere avuto, o avere in futuro, un effetto significativo.

Sul piano pratico si prenda l'esempio di economia in Italia. È noto che nella VTR 2001-2003 un esponente di spicco della teoria *non mainstream* (Luigi Pasinetti) propose una mozione di minoranza, sostenendo che la valutazione sarebbe stata distorta a favore della maggioranza neoclassica. Nella VQR 2004-2010 tutte le valutazioni sono state approvate all'unanimità dal GEV Economia. Cosa è successo? Forse che le posizioni minoritarie sono state eliminate a causa della valutazione e il GEV era composto solo dalla maggioranza vincente? Niente affatto. Il GEV includeva due tra i leader delle posizioni non ortodosse, rispettivamente la c.d. scuola italiana di Cambridge di taglio strutturalista e istituzionalista (Neri Salvadori) e l'economia evolutiva (Giovanni Dosi). Sono forse stati vittime di una sindrome di Stoccolma, essendo prigionieri di una valutazione che li avrebbe infine eliminati? In realtà si tratta di due minoranze estremamente attive, dotate di sedi di pubblicazione proprie e di grande impatto, con una estesa rete internazionale di allievi e colleghi. Pur all'interno di un profondo conflitto epistemico, è stata possibile la convergenza su criteri comuni di qualità della ricerca.

Va riconosciuto che esistono rischi di riduzione della diversità (in particolare nei confronti della ricerca interdisciplinare), ma questi sono un esempio di un più generale problema della scienza contemporanea, alla quale chi è interessato seriamente al pluralismo epistemico dovrebbe contribuire. Tra questi occorre citare i sistemi di *peer review* sotto pressione per sovraccarico di lavoro (c.d. *evaluator fatigue*), le distorsioni soggettive (mitigabili ma difficilmente eliminabili) degli esperti enfatizzate dal sovraccarico dei sistemi di selezione *ex ante* dei progetti e delle sottomissioni a rivista. Rispetto a queste tematiche, i critici della valutazione hanno talora opposto l'idea di

abbandonare non solo la valutazione della ricerca ma anche la selezione *ex ante* delle riviste e dei progetti di ricerca, sostituendola con la valutazione *ex post* basata sulla pubblicazione ad accesso aperto su piattaforme libere (*repositories*). Invece di avere la *peer review* prima della pubblicazione (con tutto il suo seguito di *impact factor* della rivista e di citazioni), si tratterebbe di consentire ad ogni ricercatore di pubblicare su piattaforme libere, ad esempio di ateneo, e di raccogliere successivamente le critiche dei colleghi.

Va detto con chiarezza che la valutazione *ex post* è un mito ideologico, frutto di una visione idealizzata della repubblica della scienza, che non fa i conti con la organizzazione della ricerca. Semplicemente non funziona. Vi sono ragioni profonde per cui non può funzionare. Queste hanno a che fare con il sistema di incentivi dei ricercatori.

Quando un ricercatore riceve un invito dall'*editor* di una rivista per fare un *referee report*, questo invito porta con sé tre messaggi: (i) il *referee* viene invitato personalmente, quindi riceve una legittimazione personale da parte della comunità scientifica per mezzo di una sua istituzione, come la rivista scientifica; (ii) il *referee* è membro di un piccolo gruppo di esperti, in genere 2 o 3 persone, quindi la sua valutazione conta molto; (iii) esiste una decisione finale sulla quale l'opinione del *referee* ha un impatto formalizzato e visibile. Vi è inoltre un ulteriore incentivo, che è la possibilità di accedere a informazioni preziose (non solo sui risultati, ma anche sui metodi, sui dati, sulle referenze) prima che esse divengano di dominio pubblico. A fronte di queste condizioni i ricercatori allocano il loro tempo (scarso) ad attività di referaggio che di per sé non sono visibili e non contribuiscono direttamente alla produzione scientifica.

Al contrario, quando un lavoro è già pubblicato su una piattaforma libera: (i) il *referee* si autolegittima, come uno tra i tanti; (ii) il *referee* ha un impatto indeterminato, che dipende dal numero (sconosciuto) di altri contributori; (iii) non vi è alcuna garanzia di impatto sulla decisione finale (revisione dal lavoro, ritiro). A mia conoscenza a queste obiezioni non è mai stata data risposta convincente.

In conclusione, come ogni sistema sociale complesso, la valutazione della ricerca produce effetti nei comportamenti sociali, inclusi effetti non desiderati prodotti da comportamenti opportunistici (*gaming*). Tuttavia i principali effetti non desiderati possono essere mitigati efficacemente da sistemi di valutazione flessibili e adattabili.

5. Evidenze sull'impatto della valutazione sulle discipline umanistiche e sociali

A fronte dei numerosi argomenti teorici avanzati contro la valutazione, che evidenze empiriche abbiamo? Possiamo sostenere che essa abbia prodotto i danni che le

vengono attribuiti? Una analisi spassionata suggerisce che vi sia una enfattizzazione degli effetti negativi. Negli ultimi cinque anni diversi studi hanno esaminato l'impatto della valutazione in vari paesi europei⁴. Tra gli effetti osservati vi sono i seguenti:

- aumento delle pubblicazioni scientifiche prodotte dai ricercatori nei periodi successivi alla introduzione della valutazione;
- aumento della quota di pubblicazioni prodotte in inglese;
- adozione generalizzata della *peer review* da parte delle riviste scientifiche;
- ricerca della pubblicazione su riviste con IF elevato e/o più prestigiose.

Al contrario:

- non si osservano modifiche strutturali nel mix di prodotti bibliografici (libri, articoli su rivista, capitoli di libro);
- non sono scomparse le monografie;
- non si osserva una riduzione nella dimensione media degli articoli su rivista (*salami slicing*);
- nelle indagini con questionario non emerge una rinuncia a perseguire temi di ricerca stabiliti autonomamente a vantaggio di temi di più facile pubblicabilità o conformità con le posizioni dominanti nelle discipline scientifiche.

Certamente sono necessari molti studi ripetuti nel tempo e svolti in vari paesi. Allo stato delle conoscenze, tuttavia, molte delle previsioni pessimistiche sull'impatto distruttivo della valutazione sulle discipline umanistiche e sociali sembrano non suffragate dall'evidenza.

6. L'impatto della valutazione sulla integrità della ricerca

Più recentemente viene avanzata la tesi che la valutazione induca comportamenti opportunistici nei soggetti valutati, fino al punto di compromettere la integrità della ricerca. Il tema ha assunto importanza a fronte dell'aumento, sia pure entro dimensioni non allarmanti, di fenomeni diversi tra loro per natura e gravità come le ritrattazioni, la

⁴ Rinvio ai volumi presso Il Mulino e Springer per le referenze su questi punti. Su questi temi è utile consultare il sito della iniziativa europea COST sulla valutazione della ricerca umanistica e sociale: <https://enressh.eu>

produzione di dati falsi, il plagio e il furto di idee. Il tema della integrità della ricerca è molto ampio e richiederebbe una analisi dedicata⁵.

Qui vorrei sottolineare che non esistono dimostrazioni del fatto che l'aumento di casi di violazione della integrità della ricerca sia *causato* dalla valutazione. Questi casi aumentano anche in paesi che *non* hanno la valutazione. Nell'unico caso in cui si può impostare un controllo è vero il contrario: la sospensione delle procedure di valutazione in Australia non ha comportato alcuna riduzione del fenomeno.

Credo al contrario che la valutazione possa contribuire a ridurre le violazioni della integrità. Elenco di seguito alcuni degli interventi introdotti dall'ANVUR che, a mio avviso, aumentano la integrità della ricerca.

Pubblicità e trasparenza

- Definizione di rivista scientifica;
- Introduzione della *peer review* nelle discipline umanistiche e delle scienze sociali in Italia;
- Pubblicazione integrale dei criteri di valutazione (VQR, ASN);
- Revisione frequente delle classificazioni (rivista scientifica, rivista classe A);
- Rotazione degli esperti.

Adattamento nella definizione degli indicatori

- *Full counting / fractional counting*;
- Autocitazioni.

Qualificazione della *peer review*

- Ridefinizione dei criteri nel testo delle domande valutative nella seconda VQR;
- Coinvolgimento società scientifiche;
- Codici etici sviluppati autonomamente dalle società scientifiche.

In riferimento al tema della nomina di esperti, che è strumento irrinunciabile della valutazione, vi è stata assunzione di responsabilità piena, apertura di *Call* per esperti e

⁵ Ho discusso il rapporto tra valutazione e integrità della ricerca al convegno organizzato dal CNR su *Research integrity nelle scienze umane e sociali*, il 4 giugno 2018. La relazione non è pubblicata, ma disponibile su richiesta.

pubblicazione della composizione dei *panel* e dei CV dei membri. In aggiunta, per le tematiche che più hanno interessato le comunità umanistiche, ovvero la classificazione delle riviste, è stato praticato sistematicamente un insieme di criteri:

- durata limitata dell'incarico;
- principio di rotazione;
- estrema attenzione a profili di pluralismo metodologico o culturale;
- procedure trasparenti di revisione periodica del *rating* riviste, inizialmente con cadenza annuale, poi rivista in seguito alle pronunce del Consiglio di Stato;
- accurata gestione dei conflitti di interesse, con una definizione estesa (che include anche conflitti di scuola) e pratiche di astensione dal giudizio e utilizzo di *referee* esterni.

In conclusione, confermo l'idea che la valutazione della ricerca "sta o cade" sul terreno della valutazione nelle discipline umanistiche e sociali. Qui la discussione è ancora accesa. Ritengo che essa debba essere collocata innanzitutto sul piano alto delle motivazioni teoriche e metodologiche, poi su quello delle evidenze empiriche. Su questo piano si deve essere intellettualmente aperti e pronti anche al conflitto delle idee. Al contrario, la generalizzazione dei casi singoli o la polemica precostituita hanno durata limitata.

Riferimenti bibliografici:

BOD 2013

R. Bod, *A New History of Humanities. The Search for Principles and Patterns from Antiquity to the Present*, Oxford.

BONACCORSI 2015

A. Bonaccorsi, *La valutazione possibile. Teoria e pratica nel mondo della ricerca*, Bologna.

BONACCORSI 2016

A. Bonaccorsi, *L'impatto della valutazione sulle scienze sociali in Italia. Lo strano caso delle discipline aziendali e della sociologia*, «Notizie di Politeia» CXXIII, 36-45.

BONACCORSI 2018

A. Bonaccorsi (a cura di), *The Evaluation of Research in Social Sciences and Humanities. Lessons from the Italian Experience*, New York.

BRUGNOLO *et al.* 2017

S. Brugnolo, D. Colussi, S. Zatti, E. Zinato, *La scrittura e il mondo. Teorie letterarie del Novecento*, Roma.

DE KETELE – PAQUAY 2011

J.M. De Ketelee, L. Paquay, *Quels critères de qualité dans les recherches empiriques? Le cas des recherches en éducation*, in P. Servais (éd.), *L'évaluation de la recherche en sciences humaines et sociales*, Louvain-la-Neuve.

FRANZINI 2018

E. Franzini, *Moderno e postmoderno. Un bilancio*, Milano.

KAGAN 2009

J. Kagan, *The Three Cultures. Natural Sciences, Social Sciences, and the Humanities in the 21st Century*, Cambridge.

LAMONT 2009

M. Lamont, *How Professors Think. Inside the Curious World of Academic Judgment*, Cambridge (Mass.).

PEARL – MACKENZIE 2018

J. Pearl, D. Mackenzie, *The Book of Why. The New Science of Cause and Effect*, London.

TETLOCK – LEBOW 2001

P.E. Tetlock, R.N. Lebow, *Poking Counterfactual Holes in Covering Laws. Cognitive Styles and Historical Reasoning*, «American Political Science Review» XCV.4, 829-43.

VIOLA 2017

M. Viola, *Evaluation of Research(ers) and its Threat to Epistemic Pluralisms*, «European Journal of Analytical Philosophy» XIII.2, 55-78.

WALLISER 2009

B. Walliser, *La cumulativité du savoir en sciences sociales*, Paris.

WALLISER 2015

B. Walliser (éd.), *La distinction des savoirs*, Paris.

ZIMAN 2000

J. Ziman, *Real Science. What it Is, and What it Means*, Cambridge.