Giornale italiano di psicologia

Estratti riservati ai soci AIP

il Mulino

GIORNALE ITALIANO DI PSICOLOGIA
Estratti riservati ai soci AIP

GIORNALE ITALIANO DI PSICOLOGIA, 1-2020, pp. 13-48

La valutazione della psicologia: considerazioni su un decennio turbolento, *di Luca Tommasi*

GIORNALE ITALIANO DI PSICOLOGIA, 3/4-2020, pp. 1047-94

Lo sfondo scontato della valutazione, *di Chiara Berti* (pp. 1047-54)

Valutare la qualità di persone, strutture, didattica, ricerca... alla ricerca del kit di sopravvivenza, *di Santo Di Nuovo* (pp. 1055-60)

Per una valutazione della didattica oltre i processi, *di Raffaella Ida Rumiati* (pp. 1061-66)

Un sistema di valutazione della psicologia attento al generale e al particolare, *di Remo Job* (pp. 1067-70)

Valutare processi e valutare prodotti: la complessità nella valutazione della qualità in università, *di Lorenzo Avanzi, Amelia Manuti, Monica Molino, Guendalina Graffigna e Dina Guglielmi* (pp. 1071-76)

Autovalutazione, valutazione periodica e accreditamento dei processi o degli esiti?, *di Fabio Lucidi* (pp. 1077-82)

La valutazione della qualità: tra formalismo, aziendalismo, *compliance* e «merito», *di Terri Mannarini* (pp. 1083-88)

La psicologia del futuro dipende da noi?, *di Tomaso Vecchi* (pp. 1089-94)

GIORNALE ITALIANO DI PSICOLOGIA
1-2020, pp. 13-48

Estratti riservati ai soci AIP, © Società editrice il Mulino

LA VALUTAZIONE DELLA PSICOLOGIA: CONSIDERAZIONI SU UN DECENNIO TURBOLENTO

LUCA TOMMASI

Università di Chieti-Pescara «Gabriele d'Annunzio»

Riassunto. Sono trascorsi circa dieci anni dall'avvio delle attività di valutazione del mondo accademico in Italia. Alcune di queste (ASN, VQR) hanno già maturato l'esperienza di molteplici esercizi successivi, spesso con echi e riverberi di non poco conto per le carriere individuali, per l'offerta formativa e per l'assegnazione di risorse alle strutture nelle quali hanno luogo formazione e ricerca. In questo lasso di tempo la psicologia accademica è stata interprete attiva, talvolta esprimendo coesione e condivisione al suo interno, altre volte segnalando l'esigenza di una maggiore articolazione e differenziazione, più in generale non facendosi trovare impreparata rispetto all'impegno comportato dalla valutazione. L'intento dell'articolo è quello di comprendere, guardando a questo pur non lungo passato, se vi siano aspetti dei vari livelli della valutazione (con particolare riferimento ad AVA, VQR, ASN, PRIN) che possano essere considerati irrinunciabili o che al contrario dovrebbero essere modificati per migliorare prassi e metodologie operative future e, più in particolare, se vi siano delle peculiarità della valutazione della psicologia che debbano essere tenute in particolare considerazione nell'interlocuzione con le istituzioni di governo e politiche.

1. INTRODUZIONE

Da ormai un decennio l'università italiana ha adottato, con un certo ritardo rispetto ad altri paesi appartenenti allo Spazio Europeo dell'Istruzione Superiore (EHEA), un insieme di sistemi di valutazione che hanno l'obiettivo di verificare periodicamente il livello qualitativo della ricerca e della didattica nell'ottica di raggiungere e possibilmente mantenere degli standard elevati (Baldissera, 2004; Bonaccorsi, 2018; Decataldo, 2018). In modo crescente, però, le attività di valutazione hanno interessato in modo ben più pervasivo la vita degli accademici, perché i sistemi valutativi non hanno riguardato soltanto la didattica e la ricerca, che restano probabilmente le due colonne portanti dell'attività universitaria, ma hanno investito i meccanismi di reclutamento delle persone, la distribuzione dei fondi di ricerca, il giudizio sulla *performance* degli atenei e dei dipartimenti, compresa la loro capacità di coinvolgimento della società e di penetrazione nel mondo dei servizi e produttivo (la cosiddetta terza missione). È indubitabile che negli ultimi dieci anni tale innovazione della modalità di

Estratti riservati ai soci AIP, © Società editrice il Mulino

concepire il funzionamento dell'università, preceduta da una fase preparatoria caratterizzata da una pressione normativa e gestionale tutto sommato moderata (a partire dall'adesione al cosiddetto Processo di Bologna avviato nel 1999), abbia assunto un peso e un'accelerazione sempre maggiori. Ciò è accaduto sia per il progressivo aggiungersi degli oggetti suscettibili di valutazione menzionati più sopra, sia per la crescente articolazione dei meccanismi gestionali che tali innovazioni hanno comportato. A partire dall'entrata in vigore della Legge 30 Dicembre 2010, n. 240, la cosiddetta Riforma Gelmini (proposta nel 2008 e modificata nel 2010), il tempo dedicato allo svolgimento di didattica e ricerca di un accademico italiano ha risentito in modo sempre più pressante della parallela necessità di rendere conto in modo diretto o indiretto delle proprie attività, al punto che è oramai condivisa tra gli accademici la sarcastica constatazione che non sia più possibile mantenere elevata la qualità del proprio lavoro dovendo dedicare così tanto tempo ed energie alla rendicontazione di quanto sia elevata la qualità del proprio lavoro. È tuttavia altrettanto indubbio, al di là della comprensibile insofferenza che tutto ciò ha comportato per la vita quotidiana degli accademici, che la valutazione sia utile e probabilmente indispensabile, poiché il suo fine ultimo, sicuramente nobile e virtuoso, non è certo quello di tediare gli attori della ricerca e della formazione con sempre più insopportabili richieste amministrative, bensì quello di mantenere elevati gli standard dei processi e la qualità dei risultati del lavoro accademico, a maggior ragione se tale lavoro è finanziato da risorse pubbliche.

Il sistema della valutazione così come lo conosciamo oggi, adottato in Italia a cavallo tra il vecchio e il nuovo millennio con interventi normativi in alcuni casi strutturali (per esempio, attraverso l'istituzione dei Nuclei di Valutazione in tutti gli atenei [1994] e l'istituzione dell'Osservatorio per la Valutazione del Sistema Universitario [1996]) ed in altri un po' più sotto traccia (l'istituzione del Comitato di Indirizzo per la Valutazione della Ricerca [1999], la trasformazione dell'Osservatorio in Comitato Nazionale per la Valutazione del Sistema Universitario [1999]), ha posto sin dagli esordi l'accento sulla valutazione della *qualità*, uniformandosi ad una rappresentazione intuitiva di ciò che tale concetto possa significare in ambito accademico (Rizzi e Silvestri, 2002; Bonaccorsi, 2018; Borrelli e Stazio, 2018). Sarebbe interessante ricorrere ad un dizionario della lingua italiana per cercare di delimitare con pretesa di oggettività la semantica del termine «qualità», provando a comprenderne le declinazioni rilevanti per il sistema universitario. Non volendo tuttavia inoltrarci in questa sede in un'analisi ontologico-lessicale, ma affidandoci al significato condiviso del concetto di qualità per quel che riguarda la nostra comunità, vale la pena sottolineare che in tutte le aree dell'educazione

Estratti riservati ai soci AIP, © Società editrice il Mulino

superiore tale concetto rimanda ad alcuni prerequisiti di base che dovrebbero caratterizzare la qualificazione del lavoro accademico. Che si parli di ricerca o di didattica, la qualità dovrebbe avere a che fare con la profondità e l'ampiezza delle conoscenze, con il rigore e la padronanza dei metodi, con l'autorevolezza riconosciuta all'interno della comunità di riferimento, con l'essere al passo coi tempi inteso come originalità e innovatività (nel caso della ricerca) oppure come aggiornamento (nel caso della didattica), con la capacità di stare al servizio della società, intesa come bacino studentesco ma anche come popolazione e territorio. Se queste condizioni e prerequisiti possono essere condivisi sul piano intuitivo – e ciascuno potrebbe aggiungerne altri rifacendosi ai propri più alti ideali di verità, giustizia e bellezza – si resta un po' delusi nel verificare che la declinazione operativa della qualità, così come è stata intesa a livello europeo dal Processo di Bologna, rimanda più ad una rete semantica cresciuta per agglutinamento progressivo di altri concetti, che non ad una definizione fondativa in grado di orientare gli attori dell'alta formazione verso obiettivi chiari e precisi. È utile notare che il progressivo arricchimento di tale rete semantica ha sistematicamente incorporato costrutti provenienti da un approccio aziendalista all'idea di qualità, mutuando concetti tipici dell'ambito industriale e manufatturiero, intorno al quale il campo della «gestione della qualità» si sviluppò dopo la seconda guerra mondiale, per trasformarsi negli anni '80 in direzione della «qualità dei servizi» e della «qualità totale» (Deming, 1982; Kano, Nobuhiku, Fumio e Shinichi, 1984). A partire dalla stessa Dichiarazione di Bologna del 1999, in cui si faceva riferimento in una sola occorrenza al costrutto, raccomandando «la promozione della cooperazione europea nell'assicurazione della qualità», nessuna definizione di qualità o, eventualmente, dei requisiti ai quali essa potrebbe corrispondere, è mai stata proposta. L'interpretazione è stata lasciata per lo più implicita o nel migliore dei casi è consistita in un aggancio incidentale ad altri costrutti. Ad esempio, nei comunicati rilasciati a seguito delle riunioni biennali dei Ministri dell'Università dei Paesi dell'Unione Europea, che hanno avuto (e continuano ad avere) la funzione di vigilare sull'andamento del Processo di Bologna, si è fatto riferimento a (in sottolineato i contenuti nuovi):

– *sistemi nazionali* di assicurazione della qualità; *standard* di qualità; *reti* di assicurazione della qualità (Praga 2002);

– *efficacia* dei sistemi di assicurazione della qualità; *criteri e metodologie mutuamente condivise* dell'assicurazione della qualità; *prerequisiti* dei sistemi di assicurazione della qualità nazionali: 1) *definizione delle responsabilità* degli organismi e delle istituzioni; 2) *valutazione* dei programmi includente verifiche interne ed esterne, partecipazione

Estratti riservati ai soci AIP, © Società editrice il Mulino

degli studenti e pubblicazione dei risultati; 3) sistemi di *accreditamento*; *internazionalizzazione e networking* (Berlino 2003);

– *peer review delle agenzie* di assicurazione della qualità; *registro europeo delle agenzie* di assicurazione della qualità; qualità *dell'insegnamento* (Bergen 2005);

– *forum* dell'assicurazione della qualità europea (Londra 2007);

– *educazione* di qualità; qualità dei programmi ed *eccellenza della ricerca* (Lovanio 2009);

– qualità e *rilevanza dell'apprendimento e dell'insegnamento*; benefici delle *tecnologie digitali*; rafforzamento del *legame tra apprendimento, insegnamento e ricerca*; intensificazione delle attività di sviluppo della *creatività, innovazione e imprenditorialità*; sviluppo delle competenze che più possono soddisfare le *aspirazioni personali* e i *bisogni sociali; descrizione trasparente dei risultati di apprendimento e del carico di lavoro; percorsi di apprendimento flessibili; metodi di insegnamento e verifica appropriati* (Yerevan 2015);

– costruzione della carriera accademica basata su *attività di ricerca di successo* ed insegnamento di qualità; riconoscimento dell'alta qualità e dell'*insegnamento innovativo* nelle carriere (Parigi 2018).

Appare curioso che il concetto di qualità non sia quindi stato definito nemmeno ad uso dei Ministri, per semplificare la comprensione dei progressi fatti in itinere dal Processo di Bologna, se non come rimando a costrutti che a loro volta facevano riferimento alla qualità dandola per scontata, o prevedendo azioni e interventi che si riproponevano di perseguirla senza tuttavia definirla. Nel migliore dei casi (comunicato di Yerevan del 2015) va apprezzato lo sforzo di elencare aspetti e requisiti specifici che potrebbero caratterizzare tale concetto, anche se nulla di tutto ciò risulta esplicitato chiaramente, e resta il dubbio che si tratti più che altro di un comodo nominalismo. Risulta quindi evidente che anche tra i legislatori e i *policymaker* che hanno avviato e gestito il Processo di Bologna, il concetto di qualità sia rimasto più agganciato all'intuizione personale e al buon senso che non ad una precisa statuizione di cosa essa debba veramente essere.

Adeguandosi a quanto era stato concepito e concordato in sede europea, il nostro Paese è oramai entrato in pieno nell'era valutativa, in particolar modo a partire dalla comparsa sulla scena dell'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (Anvur), istituita nel 2006 ed avviata operativamente a partire dal 2010, in quasi perfetta sincronia con la già citata Riforma Gelmini. In questo decennio, Anvur, che è un'agenzia di natura tecnico-operativa indipendente dal Ministero dell'Università e della Ricerca (ma da quest'ultimo vigilata) e che nel 2019 è stata accreditata ufficialmente dalla *European Association for Quality Assurance in Higher*

Estratti riservati ai soci AIP, © Società editrice il Mulino

Education (ENQA), ha avuto il compito di mettere a punto e gestire le iniziative di valutazione che, assieme alle innovazioni normative comportate dalla Riforma Gelmini per quel che attiene il funzionamento di atenei, corsi di studio e reclutamento universitario, hanno oggettivamente rivoluzionato la mentalità e le prassi quotidiane degli accademici italiani. Gli acronimi di tali iniziative, oramai scolpiti nell'immaginario (e in alcuni casi negli incubi) della maggior parte di noi, sono presto ricordati: ASN (Abilitazione Scientifica Nazionale), AVA (Autovalutazione, Valutazione periodica e Accreditamento), VQR (Valutazione della Qualità della Ricerca). Oltre a queste iniziative di valutazione va ricordato il ruolo centrale di Anvur in altre non meno rilevanti missioni valutative, come per esempio l'accreditamento dei dottorati, la premialità riservata ai «dipartimenti di eccellenza», la messa a punto di indicatori relativi alle conoscenze disciplinari e trasversali degli studenti (TECO: TEst sulle COmpetenze), la premialità destinata ai ricercatori e docenti più produttivi (FFABR: Finanziamento delle Attività Base di Ricerca). Per continuare il ragionamento avviato sulla definizione di qualità, che vuole qui essere un *leitmotiv* volutamente provocatorio utile a promuovere la riflessione e la discussione, va detto che anche a livello di sistema nazionale della valutazione, così come a livello europeo, è invalsa una concezione sostanzialmente aziendale di qualità. Perdura, cioè, l'affidamento ad una non-definizione, o a una definizione intuitiva di tale concetto, a cui si aggiunge il rimando all'armamentario dei metodi e degli strumenti per valutarla e assicurarla. Ciò riecheggia la ben nota definizione di intelligenza data da Edwin Boring che, eludendo volutamente l'*impasse* che si creerebbe nel tentativo di definire questo complesso costrutto, la fece equivalere a «ciò che è misurato dai test di intelligenza» (Boring, 1923). Anche nel nostro Paese, quindi, pare non si sia ritenuto utile stabilire in modo condiviso una definizione di qualità della formazione e della ricerca, o interrogarsi preliminarmente su cosa essa dovrebbe o potrebbe essere; di fatto è stata accettata l'equivalenza della qualità con i risultati dei processi necessari a misurarla e/o con la mera esistenza di tali processi.

Questo articolo trae spunto da una tavola rotonda organizzata in seno al XXV Congresso Nazionale della Sezione di Psicologia Sperimentale dell'AIP tenutosi all'Università Vita-Salute San Raffaele di Milano. A partire da alcune considerazioni sull'impatto delle prassi di valutazione per la psicologia accademica, intendo qui mettere in luce alcuni punti di forza e meriti, così come problemi e criticità di tali prassi, nella consapevolezza che la discussione sulla valutazione e sulla qualità all'interno di una comunità scientifica, non essendo ancora stata affrontata in modo coordinato in Italia né per la psicologia, né per altre discipline, a questo punto del cammino sia un'operazione

Estratti riservati ai soci AIP, © Società editrice il Mulino

possibile, opportuna e probabilmente necessaria (si veda anche Cubelli, 2015, per una precedente disamina di ASN e VQR). La trattazione riguarderà quattro ambiti, tre dei quali corrispondono proprio ai tre «trigrammi diabolici» di cui sopra, e cioè ASN (carriere), AVA (strutture e corsi di studio) e VQR (ricerca), ma si estenderà anche ad un altro ambito fortemente (e giustamente) dipendente dalla valutazione, vale a dire l'attribuzione di finanziamenti nazionali per la ricerca, con particolare riferimento ai cosiddetti fondi per Progetti di Rilevante Interesse Nazionale (PRIN). Sono convinto che su questi temi tutti gli psicologi accademici italiani abbiano un'esperienza diretta e posizioni molto ben definite, così come sono certo che tali esperienze e posizioni costituiscano oggetto di comunicazione e scambio tra colleghi per buona parte del loro tempo, dentro e fuori le università. Quel che spero possa sortire da questo contributo di riflessione è una rinvigorita consapevolezza che l'opinione della comunità degli psicologi accademici dovrebbe avere peso come valore condiviso e, operativamente parlando, come strumento di lavoro in questo strano campo di forze che si chiama università italiana. Mi rendo conto che tale speranza possa apparire scoraggiata a priori, per via delle diffuse spinte individualistiche, delle territorialità di settore e del poco tempo a disposizione che caratterizza il nostro sforzo quotidiano di fare della buona scienza e della buona formazione. Tuttavia credo che valga la pena sacrificare un po' di tempo come comunità scientifica per metterci d'accordo su che cosa funzioni e che cosa non funzioni della valutazione, come possa essere definita e misurata la qualità, che cosa dovrebbe caratterizzare in modo irrinunciabile l'identità scientifica della psicologia agli occhi dei sistemi di valutazione. Credo inoltre che sia importante fare tutto ciò *nonostante* la pesante interferenza comportata dallo sforzo valutativo nella vita quotidiana di ciascuno di noi. Per necessità di sintesi, nel prosieguo darò per scontata la conoscenza della natura e delle caratteristiche dei meccanismi di valutazione, pur delineandone alcuni aspetti di base.

2. AVA E PSICOLOGIA

Il sistema di Autovalutazione, Valutazione periodica e Accreditamento (AVA), giunto ormai alla terza revisione (i decreti ministeriali sono stati emanati nel 2013, nel 2016 e nel 2019), rappresenta la fattispecie più emblematica di iniziativa di valutazione dell'ultimo decennio, perché la sua finalità è quella di valutare – rispetto a standard predeterminati – sia il sistema, sia alcune delle sue parti, e quindi ambisce ad ampiezza e ad articolazione. Inoltre i suoi potenziali esiti hanno effetti importanti ed eventualmente fatali, vi-

Estratti riservati ai soci AIP, © Società editrice il Mulino

sto che qualificano i valutati come pienamente degni, oppure come soddisfacenti (pur con gradazioni di giudizio e raccomandazioni di migliorie più o meno rilevanti), o infine come suscettibili di soppressione. Il ciclo di valutazione prevede che gli attori coinvolti (in particolare i coordinatori dei corsi di studio e i direttori dei dipartimenti, ma sostanzialmente tutti i docenti di un ateneo) provvedano di anno in anno e autonomamente alla verifica del soddisfacimento di alcuni standard (autovalutazione) e solo più sporadicamente, ma ciclicamente, siano soggetti ad una valutazione esterna alla quale farsi trovare pienamente attrezzati, con l'obiettivo di ottenere l'agognato accreditamento, possibilmente di alto rango (valutazione periodica). Va ricordato che gli oggetti della valutazione sono gli atenei, i dipartimenti/scuole e i corsi di studio, e che di queste ultime due entità avviene una selezione di un campione rappresentativo del loro numero totale in funzione della dimensione (misurata come numerosità dei corsi di studio) dell'intero ateneo. Restringerò l'analisi ai soli dipartimenti e corsi di studio di predominante interesse psicologico che hanno ricevuto le visite di accreditamento periodico programmate delle Commissioni di Esperti di Valutazione (CEV), come risultanti dai rapporti finali pubblicati sino ad ora da Anvur. Sebbene il campione delle sedi che hanno sostenuto la visita delle CEV per un dipartimento di prevalente rilevanza psicologica, o per uno o due corsi di studio di ambito psicologico (L-24, LM/51), o ancora per una combinazione dipartimento/CdS non sia numerosissimo (una dozzina, delle quali tre università telematiche), dall'analisi degli esiti si nota una sostanziale equivalenza tra la valutazione finale della sede e la specifica valutazione conseguita da dipartimenti e/o corsi di studio. Casomai, può essere rimarchevole segnalare che in un paio di casi la valutazione di un singolo dipartimento psicologico o con rilevante presenza della psicologia è stata superiore alla valutazione finale della sede. Ciò testimonia il fatto che il contributo della psicologia alla performance valutativa degli atenei non è stato mai più deludente del risultato d'insieme, e che in alcuni casi esso è stato superiore a quello di altri ambiti non prevalentemente psicologici; questo suggerisce, ove possibile, di promuovere e far pesare maggiormente in futuro il modello organizzativo che si è data la psicologia nelle *governance* degli atenei.

Tra tutte le attività di valutazione messe in atto nel sistema universitario italiano e descritte in questo articolo, AVA è l'unica a misurare direttamente il livello di qualità degli aspetti didattici. Questo evoca una prima riflessione, e cioè che l'attenzione dedicata dai meccanismi di valutazione ai corsi di studio potrebbe apparire sottodimensionata rispetto a quella dedicata alle strutture (AVA di sede e di dipartimento, piano della *performance*), alla ricerca e alla terza

Estratti riservati ai soci AIP, © Società editrice il Mulino

missione (VQR, PRIN) e alle persone (ASN, FFABR), a maggior ragione se si considera che il servizio offerto dalle università dovrebbe attribuire alla formazione un peso almeno analogo a quello attribuito alla ricerca. Si può però rovesciare il ragionamento fatto poc'anzi, assumendo che la valutazione dei corsi di studio svolta attraverso il sistema AVA sia di fatto appropriatamente dimensionata, e deducendo che le attività di valutazione di altri aspetti (in particolare dei dipartimenti) siano per contro gravate da una sostanziale ipertrofia e ridondanza. Considerando infatti *i*) che i dipartimenti sono già di per sé soggetti ad una verifica periodica relativamente alla qualità dei risultati della ricerca dei propri componenti attraverso la VQR, *ii*) che essi devono già rendere conto della messa a punto e del monitoraggio di obiettivi rispondenti al piano della *performance* dell'ateneo di cui fanno parte, e *iii*) che esistono già sistemi di incentivazione che dovrebbero spingere al perseguimento di standard elevati (l'iniziativa dei «dipartimenti di eccellenza»), si può concludere che la valutazione della qualità dei dipartimenti realizzata attraverso il sistema AVA sia sovrabbondante e probabilmente un po' sterile. Ciò è testimoniato da due aspetti: in primis, una disamina delle narrative con le quali sono formulati i giudizi stilati da Anvur fa emergere che la valutazione dei dipartimenti si riduce spesso alla conferma della messa in atto di processi di monitoraggio e di controllo con retroazione, gestiti da direttori, delegati o commissioni preposte a dimostrare l'esistenza e la tracciabilità di quei processi, con uno spazio insignificante o nullo per la messa in luce di aspetti di merito (o di difficoltà) specificamente attinenti alle discipline che caratterizzano questo o quel dipartimento. Se è vero infatti che ciascun ambito disciplinare è caratterizzato da una storia e da modalità operative che potrebbero trovare nella formulazione delle attività di un dipartimento uno spazio di rivendicazione del valore intrinseco di quelle tipicità disciplinari, è sconfortante che ciò non venga in qualche modo riconosciuto, se necessario anche caso per caso, contesto per contesto, dalla visita delle CEV. In altre parole, le CEV, essendo costituite da esperti di sistema, impongono un modello di processo uniforme e universale a cui attenersi, ma sembrano evitare deliberatamente di far tesoro delle peculiarità disciplinari, il cui insieme costituisce quell'articolata varietà di usi e costumi che qualifica la ricchezza del mondo dell'alta formazione, e che potrebbe permettere di creare un catalogo, differenziato per disciplina o per area, di prassi, di distinguo ed eventualmente anche di legittime deviazioni dallo standard qualitativo predefinito. Si noti bene, non viene qui negata l'utilità o l'indispensabilità di criteri che permettano di valutare la «buona condotta» dei contenitori dipartimentali attraverso gli standard di AVA, né si vuole dimenticare che i contenuti sono già monitorati in altro modo sotto forma di valutazione dei

Estratti riservati ai soci AIP, © Società editrice il Mulino

risultati della ricerca (VQR). Si vuole piuttosto sottolineare l'esistenza di una dimensione intermedia tra il contenitore e il contenuto, fatta di particolarità di una disciplina, che AVA potrebbe tenere in considerazione per modificare a sua volta in modo retroattivo i propri criteri, arricchendoli e differenziandoli. Per fare un esempio concreto applicato al caso della psicologia, è ben nota la situazione caotica e fuori controllo relativa alla verifica dei progetti di ricerca da parte dei Comitati Etici. È una circostanza, quella del vaglio etico, che riguarda moltissime delle attività quotidiane degli psicologi attivi nella ricerca (per esempio, ottenimento di fondi e finanziamenti, pubblicazione sulle riviste) e che, al di là del ruolo strumentale o utilitaristico (poter partecipare a bandi, poter pubblicare), risponde ad un principio fondamentale per lo psicologo accademico: il vaglio etico è *a priori* un passaggio obbligato per moltissime fattispecie di ricerca psicologica. Si dà il caso che nelle varie sedi e dipartimenti tale obbligo sia assolto nei modi più diversi (Comitati Etici d'Ateneo, *Institutional Review Board*, Comitati Etici Dipartimentali o Interdipartimentali, ricorso a pareri esterni, ecc.) con svariate criticità. Tra queste vanno in particolare annoverate l'inappropriatezza dell'ambito di valutazione di molti Comitati Etici d'Ateneo (spesso istituiti per valutare esclusivamente la ricerca clinica in ambito medico) rispetto alle concrete necessità di valutazione etica degli psicologi, così come la non terzietà dei Comitati Dipartimentali «fatti in casa». Non è raro che tali problematiche di sede o di dipartimento possano indurre docenti e ricercatori a «gettare la spugna», eludendo il preliminare vaglio etico anche lì dove esso dovrebbe essere presente. Non è inoltre un caso se la Commissione Etica dell'Associazione Italiana di Psicologia ha recentemente realizzato e diffuso un questionario con lo scopo di mappare questa frastagliata geografia di prassi e di problemi. Può darsi che mi sbagli, ma mi pare che nemmeno un accenno in merito all'annoso problema dei Comitati Etici abbia avuto modo di emergere nelle visite dipartimentali AVA, non sia stata cioè né segnalata dai valutati, né scrutinata dai valutatori, alcuna disfunzionalità in tal senso, laddove è noto a tutti che un problema c'è, e che esso vede proprio nei dipartimenti (che devono deliberare sui progetti) i portatori del problema.

Un altro aspetto che testimonia di una certa ridondanza della valutazione dei dipartimenti è rappresentato dallo spazio in fin dei conti piuttosto scarno che tali valutazioni finiscono per occupare nell'insieme dei rapporti stilati da Anvur a valle delle visite delle CEV. Anche a giudicare dalla lunghezza delle narrative impiegate per descrivere tali valutazioni, ci si accorge infatti che il peso attribuito alla valutazione dipartimentale è tutto sommato ritenuto più marginale da parte dello stesso valutatore. Tutto ciò sembra suggerire che potrebbe essere abbandonata la dimensione della valutazione dipartimentale e che, se

Estratti riservati ai soci AIP, © Società editrice il Mulino

proprio essa sarà mantenuta in futuro, potrebbe essere valorizzata e tenuta in maggior considerazione la geografia delle differenze e delle specificità. Questo potrebbe valere a maggior ragione per la psicologia, che per vocazione e storia recente vede la propria collocazione in dipartimenti di natura fortemente interdisciplinare (con le scienze della vita, la filosofia, le scienze sociali, economiche, dell'educazione, ecc.), soprattutto perché la comprensione delle «regole del gioco» dell'interdisciplinarietà vincente avrebbe senza dubbio un enorme interesse generale, dato che l'interdisciplinarietà appare da non poco tempo sia come un correlato qualificante, sia come un obiettivo virtuoso in ambito accademico (Graff, 2015). In altre parole, se la missione valutativa non si limitasse alla verifica del soddisfacimento di standard predefiniti, ma si riproponesse contemporaneamente di comprendere i meccanismi del successo in presenza di singolarità o di deviazioni da quegli standard, tutti ne trarremmo giovamento.

Se la valutazione AVA delle strutture intermedie (dipartimenti) può apparire quindi, anche per la psicologia, necessariamente perfettibile oppure *tout court* rinunciabile, quello dei corsi di studio sembra invece essere il livello di valutazione AVA più delicato e forse irrinunciabile, per via dell'aspetto menzionato più sopra, e cioè che la valutazione AVA dei corsi di studio è l'unica modalità di valutazione formale di tali entità, al di là del successo di fatto che i corsi possono avere in termini di numero di strumenti immatricolati. Credo sia particolarmente importante distinguere tra il successo di un corso di studio misurato in termini di popolosità studentesca (che può dipendere da fattori di mercato, di posizionamento geografico, di mode passeggere, ecc.) o addirittura in termini di occupazione post lauream, e la qualità AVA di quel corso: un corso di studio potrebbe essere popolatissimo e portare ad elevati tassi di occupazione pur non essendo affatto di qualità. L'annoso problema della sovrappopolazione dei laureati in psicologia in Italia dovrebbe tuttavia far tornare a riflettere sull'interazione tra il fattore della numerosità degli studenti e quello dei livelli di occupazione, perché la popolosità studentesca nel caso della psicologia è sia un indicatore di successo, sia un vulnus della qualità. Se da un lato è infatti vero che i corsi di studio psicologici continuano ad intercettare una domanda superiore all'offerta, è altrettanto vero che l'offerta lavorativa in ambito pubblico e privato rimane insufficiente a fronte del numero dei laureati in psicologia (si veda a questo proposito l'articolo bersaglio di Legrenzi, Malaguti e Umiltà apparso sul Giornale Italiano di Psicologia nel 2017), pur riconoscendo gli sforzi messi in atto da diverse istituzioni per sensibilizzare il mondo dei servizi, della scuola, della sanità e aumentare così lo spazio occupazionale dei laureati in psicologia. Una domanda che sorge spontanea analizzando i rapporti di valutazione delle CEV sui corsi

Estratti riservati ai soci AIP, © Società editrice il Mulino

di studio di natura psicologica riguarda proprio questo aspetto, ed è la seguente: come è possibile che la valutazione dello stato di salute di un corso di studio psicologico si limiti alla constatazione che alcuni parametri vitali rientrano nella norma, quando di fatto quello che si sta auscultando è un organismo afflitto da un problema di salute non rilevabile in base a quei parametri? In altre parole, anche per quel che riguarda la valutazione della didattica, ben venga la pletora di criteri da soddisfare in termini di numerosità dei docenti, chiarezza degli obiettivi formativi, verifica dei risultati di apprendimento, presenza di servizi agli studenti, forme innovative di didattica, monitoraggio delle valutazioni degli studenti, ecc. Sono cose sacrosante, che anche per la psicologia hanno un grande merito, cioè rendere sempre più chiari, fruibili e *smart* i corsi di studio, tuttavia anche in questo caso le CEV sembrano perdere un'occasione, a maggior ragione se si considera che si tratta di commissioni «disciplinari» (ad esempio di area psicologica): misurare il rispetto delle prescrizioni normative e la messa in atto di politiche attive di assicurazione della qualità è un aspetto necessario ma di per sé non sufficiente ad assicurare la «qualità totale» di un corso di studio, perché non si tiene conto degli eventuali segnali di allarme che il corso di studio può fornire al valutatore per aggiornare il sistema relativamente a quello specifico ambito disciplinare. Se la qualità deve essere totale, e cioè se deve guardare al sistema in tutte le sue parti e interazioni, perché non tenere in ampia considerazione tutte le osservazioni di ritorno che un corso di studio può fornire al valutatore per aggiornare quella qualità totale, possibilmente ponendo attenzione sulle particolarità disciplinari? Come già detto per quanto riguarda la mancata raccolta di *feedback* dalle strutture dipartimentali psicologiche, anche il ritorno informativo proveniente dalla didattica psicologica avrebbe un valore potenziale che al momento appare sostanzialmente ignorato. Con un paragone forse un po' bizzarro ma a mio parere sensato nella sua semplicità, sarebbe opportuno che le CEV di Anvur si prendessero la briga di definire e classificare progressivamente le particolarità delle aree scientifiche in un modo molto simile a quella dei disciplinari dei prodotti alimentari: è evidente che una mozzarella di bufala deve rispondere a criteri di qualità tanto quanto un vino (e si potrà collocare l'uno e l'altro prodotto su diversi livelli di una medesima scala scadente-eccellente a seguito della loro valutazione), ma gli specifici criteri e descrittori della qualità della mozzarella sono ben diversi da quelli del vino e, per la nostra fortuna di consumatori e *gourmet*, non potrebbe che essere così. Proviamo però a essere concreti e rinunciamo alla chimera di perfezionare le prassi valutative con degli ideali di miglioramento: purtroppo rispunta il problema evidenziato all'inizio. Appare chiaro cioè che coloro che devono gestire i corsi di studio (in primis il coor-

Estratti riservati ai soci AIP, © Società editrice il Mulino

dinatore, poi i docenti impegnati nei vari gruppi, ma per estensione l'intero corpo docente) profondono così tante energie ed attenzione a curare i dettagli di interesse per la valutazione di sistema (cioè i parametri vitali secondo Anvur) da non avere più il tempo per riflettere come comunità scientifica sulle variabili più importanti per lo specifico della disciplina o, peggio, finiscono per perderle di vista. La prassi valutativa, schiacciando tutti sotto il peso dell'adeguamento a prerequisiti universali, fiacca la riflessione sulle esigenze, sulle tipicità e sui prerequisiti irrinunciabili di area. Credo che questo si sia visto molto bene per la psicologia: ad esempio, il problema della numerosità di classe resta un problema irrisolto *nonostante* diverse siano state le occasioni per trovare nei valutatori degli «alleati disciplinari» in grado di retroagire sul sistema circa la gravità di questo problema. Il momento valutativo viene vissuto dai valutatori come momento di verifica di prerequisiti, dai valutati come momento di esposizione della propria adeguatezza circa quei prerequisiti, e si dimenticano, o si fa finta di non vedere, i veri problemi.

3. VQR E PSICOLOGIA

Commentare la VQR (Valutazione della Qualità della Ricerca) e ciò che essa ha rappresentato – e probabilmente rappresenterà – per la psicologia nel suo insieme non è cosa semplice, ma tutto sommato è un'operazione che dovrebbe risultare facilitata da una discreta semplicità e comparabilità degli elementi a disposizione. Gli esercizi completati nel decennio valutativo sono stati infatti «soltanto» due (VQR1: 2004-2010; VQR2: 2011-2014), mentre il terzo è ai blocchi di partenza (VQR3: 2015-2019); la valutazione ha portata nazionale e non locale; sono disponibili rapporti comprensivi di tabelle che ben dettagliano metodi e risultati ottenuti per ciascuna tornata. Questo dovrebbe teoricamente permettere di trarre delle utili conclusioni sul percorso della valutazione della ricerca psicologica, ma già la comparazione tra le due tornate svoltesi nel decennio pone di fronte ad un primo importante problema generale, vale a dire la presenza di elementi di cambiamento nella pur non lunghissima storia di questo sistema di valutazione. A tal proposito, è importante chiarire subito che il cambiamento è stato causato sia da scelte interne alla psicologia, sia da scelte esterne o non conseguenti ad alcun tipo di interlocuzione con la psicologia. Le scelte interne, come ad esempio la non scontata adesione all'approccio di tipo bibliometrico sin dal primo esercizio di valutazione, sono state il frutto di una vivace discussione e di dibattito all'interno della disciplina (talvolta con posizioni di non facile conciliazione) e testimoniano della capacità della nostra disciplina di

Estratti riservati ai soci AIP, © Società editrice il Mulino

collocare sé stessa nel panorama scientifico contemporaneo in modo autonomo e maturo (Cubelli, 2015; Salvatore, Pazzagli, Lucarelli, Gelo e Giannone, 2015). Queste scelte interne hanno dato l'idea di una partecipazione attiva della psicologia e di una sua autonomia, che ha beneficiato di un vivo dialogo bidirezionale tra gli attori della ricerca, i loro rappresentati nelle società scientifiche e le istituzioni di riferimento del sistema di valutazione, in particolare i componenti dei Gruppi di Esperti della Valutazione (GEV). Tale quadro positivo è anche dovuto ad una sorta di disposizione autoctona verso la valutazione, preesistente all'avvento dei sistemi che conosciamo oggi (Laudanna, 1995, 1996; Laudanna, Miceli e Welin, 2001). Le scelte esterne, per contro, sono state spesso subìte, o quanto meno recepite con difficoltà, come è normale che sia laddove l'impegno profuso dai singoli aveva tenuto conto di meccanismi noti, per vedere quei meccanismi ridimensionati, rivisitati o addirittura accantonati a favore di meccanismi nuovi nell'esercizio successivo. È vero che se un sistema di valutazione della ricerca ripresentasse sistematicamente ad ogni esercizio i medesimi criteri e i medesimi parametri, ciò potrebbe consentire ai soggetti valutati di programmare in modo ottimale la propria produttività, permettendo di giocare d'anticipo in maniera opportunistica e facendo così risultare massimamente performante la produzione scientifica. Tuttavia è anche vero che se una certa programmazione della produttività scientifica risponde precisamente a criteri che sono stati considerati degli indicatori appropriati della qualità in un primo momento, non si capisce perché cambiarli in un secondo momento, finendo per far apparire gli attori della ricerca accademica come dei volponi dediti esclusivamente a mettere in atto comportamenti opportunistici, piuttosto che degli operosi ricercatori dediti a migliorare la qualità del proprio lavoro.

È in ogni caso un dato di fatto che i cambiamenti nei meccanismi di valutazione contenuti nel Decreto Ministeriale che ha recentemente annunciato il nuovo esercizio (VQR3), quasi tutti confermati nel relativo Bando emanato da Anvur all'inizio del 2020, hanno sortito un'unanime levata di scudi proprio per queste ragioni. Anche nella comunità psicologica, così come in tutta la comunità accademica nei suoi risvolti istituzionali (Consiglio Universitario Nazionale – CUN) e in quelli più informali, è stato infatti rilevato che la presenza di differenze tra le caratteristiche della valutazione di VQR3 rispetto a quelle di VQR2 ha introdotto fattori di distorsione che appaiono indebolire incomprensibilmente o addirittura rendere inservibile il nuovo impianto valutativo rispetto a quello precedente. Ciò è vero per numerosi aspetti, che vanno dalla definizione del metodo di valutazione dei singoli prodotti, che torna capricciosamente a dare priorità alla revisione tra pari rispetto alla collocazione in ranghi di impatto e ci-

Estratti riservati ai soci AIP, © Società editrice il Mulino

tazioni (anche nelle aree bibliometriche come è l'Area 11b, a cui afferisce tutta la psicologia), alla definizione delle diverse classi di qualità dei prodotti, che si discosta significativamente da quella precedente (e che diviene comunque poco confrontabile con essa, visto il cambio di approccio), per arrivare infine alla questione della prescrittiva ripartizione percentuale dei prodotti nelle diverse classi di qualità, che del nuovo esercizio VQR rivela la più pericolosa delle mutazioni, vale a dire la pretesa che di un insieme di prodotti, che potrebbero essere tutti oggettivamente eccellenti, venga forzata dall'alto una ripartizione che prevede in ogni caso l'assegnazione di una quota di essi alla classe dei prodotti scadenti (o viceversa). Tra gli altri problemi di VQR3 degni di suscitare una certa preoccupazione per la psicologia, sembra essere rilevante anche la scomparsa dell'articolazione della valutazione nei diversi Settori Scientifico-Disciplinari, poiché tale articolazione aveva costituito un indicatore utile per ogni settore psicologico per comprendere come migliorare la qualità della propria produzione scientifica. Per fare un esempio di questo effetto si prenda a riferimento quanto accaduto tra il primo ed il secondo esercizio VQR, vale a dire la constatazione di una significativa crescita del numero di articoli in lingua inglese in taluni SSD psicologici. Ora, che il pubblicare in riviste scientifiche in lingua inglese possa essere fatto equivalere *tout court* ad un indicatore di qualità è cosa che potrebbe essere discussa a lungo. Tuttavia, che piaccia oppure no, pubblicare in inglese è sicuramente giovevole ad una maggior diffusione di ipotesi, risultati e teorie, e quindi ad una maggior probabilità che quelle ipotesi possano essere verificate, quei dati replicati, quelle teorie confermate o confutate. La sparizione dei Settori Scientifico-Disciplinari dal radar della VQR impedirà di continuare a tenere sotto controllo quella che era risultata essere una trasformazione virtuosa per alcuni ambiti, rimescolando in un unico contenitore tutta la produzione scientifica generata dalla psicologia. Questo è solo un piccolo campionario delle storture annoverate nel ricco *cahier des doléances* del nuovo Bando VQR, sul quale tuttavia si nutrono speranze di modifica che, dato l'inevitabile ritardo tra la produzione e la pubblicazione di questo articolo, non possono essere qui confermate o smentite con tempismo.

Ciò che è stato da più parti segnalato come inaccettabile (Consiglio Universitario Nazionale, Conferenza della Psicologia Accademica, Associazione Italiana di Psicologia) è soprattutto il fatto che la VQR3 «cambi le regole in corsa», rivedendo in modo addirittura peggiorativo un dispositivo che era probabilmente perfettibile, ma sul cui impianto generale individui e strutture hanno investito tempo, energie e fondi, e che (salvo ragionevoli correttivi) sarebbe stato opportuno confermare. Al di là dell'effetto di comprensibile irritazione suscitato, quel che è certo è che dei cambiamenti marcati nelle modalità

Estratti riservati ai soci AIP, © Società editrice il Mulino

adottate dai sistemi di valutazione in esercizi successivi sono particolarmente inopportuni perché impediscono di raffrontare in modo affidabile l'evoluzione nel tempo della qualità della ricerca dei soggetti della valutazione, o addirittura impediscono di valutare gli stessi soggetti**,** come si è già detto del caso della scomparsa degli SSD. Una domanda a questo punto sorge spontanea: perché cambiare? La prima risposta è quella delineata più sopra, che vede nel cambiamento uno strumento utile a spiazzare gli attori della ricerca, impedendo loro di farsi trovare opportunisticamente attrezzati al momento della verifica. Oltre a non essere condivisibile su un piano di principio e probabilmente nemmeno su quello giuridico, poiché non si capisce come si possa operazionalizzare la qualità della ricerca in modo condiviso in un dato momento per modificare poi in modo unilaterale quelle scelte, la spiegazione dello «spiazzamento degli opportunisti» non è nemmeno vera sul piano fattuale, dal momento che la risposta reale alla domanda sembra essere molto più semplice, e giace in un breve documento. Tale documento è la «Nota di accompagnamento alla pubblicazione del Rapporto degli Alti Esperti VQR», pubblicata nel marzo del 2019 a seguito di una revisione delle caratteristiche di VQR avviata volontariamente da Anvur, affidata ad un «un gruppo di esperti indipendenti, con rilevanti esperienze di valutazione a livello comunitario o in altri paesi europei», e fortemente ispirata al sistema di valutazione britannico (*Research Excellence Framework* – REF). Secondo il *panel* di esperti, l'esperienza delle prime due VQR non era da considerarsi eccessivamente criticabile, ma poteva essere affinata. Oltre al suggerimento di alcuni correttivi di natura organizzativa (per esempio, attraverso un aumento delle quote rosa e dei componenti stranieri nei GEV), nella nota si raccomandava di mantenere l'approccio della revisione informata tra pari tramite l'uso di indicatori bibliometrici e di aumentare il numero dei GEV, «al fine di poter meglio rappresentare la granularità dell'attività di ricerca» (si noti che il documento parla dell'aumento della numerosità dei GEV, non dei loro componenti). Si può quindi concludere che quei cambiamenti di metodo che tanto malumore hanno creato nel lancio della VQR3 non siano stati affatto introdotti per scoraggiare comportamenti di natura opportunistica, bensì per far aderire l'apparato valutativo a prescrizioni autorevoli (leggasi mantenimento della *peer review* informata), ma che per quanto autorevoli potevano essere ignorate (leggasi sparizione degli SSD, che avrebbero potuto conservare l'analisi granulare invocata dagli esperti). Appare dunque evidente che si vuol far maturare il sistema VQR nei decenni attraverso quelli che sembrano essere aggiustamenti dettati volta per volta da pareri di alti esperti di valutazione oppure da scelte autonome, finendo tuttavia col rinunciare agli aspetti positivi di quanto era stato già realizzato in nome di un non

Estratti riservati ai soci AIP, © Società editrice il Mulino

ben definito processo di miglioramento. Buttando cioè il bambino con l'acqua sporca.

È comunque plausibile – quali che saranno i dettagli – che il nuovo esercizio VQR contribuirà come i precedenti esercizi a gettare luce su importanti aspetti comparativi della qualità della ricerca svolta nelle università e negli enti di ricerca del nostro paese. Sicuramente sarà possibile confrontare il livello d'insieme della qualità della ricerca psicologica rispetto alla VQR2, così come le varie sedi le une con le altre (cosa che era già possibile nella seconda tornata consultando la valutazione dell'intera Area 11b), verificandone così la tenuta e, si spera, confermando un continuo miglioramento. Due dati d'insieme che continueranno ad essere valutati, costituendo utili *proxy* qualitativi dei cambiamenti avvenuti nella psicologia al di là delle possibili differenze tra specifici settori, saranno *i*) la proporzione dei contributi in riviste scientifiche sul totale dei prodotti (rispetto a monografie, capitoli in monografie e altri prodotti) e *ii*) la proporzione dei contributi in lingua inglese. Focalizzandosi su questi due indicatori, il rapporto finale del GEV di Area 11b della VQR2 aveva potuto riscontrare, rispetto ai medesimi indicatori forniti ad esito della VQR1, uno «straordinario progresso della psicologia». È abbastanza prevedibile che tale progresso sarà ulteriormente confermato dall'esito della VQR3, dal momento che la produzione scientifica psicologica, così come quella di tutte le altre discipline bibliometriche, nel periodo di riferimento 2015-2019 è stata drammaticamente influenzata in quella direzione (cioè verso la totalità dei contributi su riviste in inglese) dal contemporaneo e sempre più concorrenziale svolgimento della valutazione delle persone ai fini di reclutamento e progressioni, come si vedrà nella sezione dedicata all'Abilitazione Scientifica Nazionale. Tuttavia, non si può non rimarcare che la diversa modalità di classificazione dei prodotti ed il diverso modello valutativo presenti nei due successivi esercizi comporteranno una sostanziale limitazione ai possibili raffronti VQR2/VQR3 circa la collocazione nelle diverse classi di qualità dei prodotti, impedendo così di realizzare un obiettivo importante per un sistema di valutazione, cioè consentire di monitorare l'evoluzione nel tempo del livello qualitativo di un dato campo scientifico. Anche concedendo momentaneamente che quella di eliminare il livello di granularità dei SSD sia stata una scelta sensata, la non comparabilità della qualità scientifica tra i due esercizi rappresenta a tutti gli effetti un'occasione perduta per la psicologia. Si badi bene, una simile situazione si era già data tra la VQR1 e la VQR2, tuttavia in quel caso si poteva ragionevolmente invocare l'imperfetta definizione dei parametri in un sistema che, essendo in fase di avvio, necessitava di rodaggio (VQR1) ma che, ciononostante, nel passaggio alla VQR2 aveva visto addirit-

Estratti riservati ai soci AIP, © Società editrice il Mulino

tura l'affrancamento ufficiale della psicologia dal resto del contenitore multiforme rappresentato dall'Area 11. Il valore di quella differenziazione disciplinare per la psicologia, che fu possibile allora e che oggi appare come una genuina scelta di campo, segnala per contrasto il peggioramento che sembra oggi interessare il dialogo tra accademia, società scientifiche, decisori politici e istituzione valutativa, un dialogo che ha visto sicuramente momenti migliori. Di nuovo, Anvur sembra aver rinunciato a far tesoro delle caratteristiche interne alle discipline, che tanto avrebbero da dire per ottimizzare il sistema di valutazione interagendo costruttivamente con esso. Più sostanzialmente, con il nuovo esercizio di valutazione Anvur sembra proprio far capire che non c'è più tempo per ascoltare chicchessia e appare procedere a testa bassa, dando priorità assoluta al proprio mandato in un momento sicuramente complicato di cambiamenti interni (rinnovo delle cariche) ed esterni (mutazioni ministeriali e parallela nascita dell'Agenzia Nazionale della Ricerca).

Una previsione che mi pare possibile fare riguarda la distribuzione della qualità della ricerca psicologica nelle sedi e nell'ulteriore articolazione dei dipartimenti. Pur con la limitazione esposta sopra (difficile comparabilità tra VQR3 e VQR2), la nuova valutazione dovrebbe continuare a permettere di raffrontare sedi e dipartimenti per quel che riguarda lo specifico della qualità della ricerca psicologica nel suo insieme. In questo caso sembra possibile intuire che l'eliminazione dei SSD possa essere addirittura utile, perché le informazioni di intersezione SSD x Sede e SSD x Dipartimento riportate nelle tabelle della prima e della seconda VQR risultavano spesso frammentarie (soprattutto nei casi di dipartimenti ibridi), mentre nel caso dei dipartimenti più grandi il dato era suscettibile di innescare antagonismi tra settori psicologici (o tra dipartimenti). L'effetto della sparizione degli SSD dovrebbe essere quello di rafforzare la psicologia nel suo insieme, restituendo al lavoro corale degli psicologi accademici di una determinata sede o di un determinato dipartimento, quale che sia il loro settore di appartenenza, una natura più coesa. Quello che si può concludere, tuttavia, è che probabilmente sarebbe stata una scelta più efficace nel nuovo esercizio VQR prevedere la valutazione dei singoli SSD, ma solo a livello di risultato aggregato nazionale, ferma restando la valutazione delle Aree nella doppia e tripla interazione con le sedi ed i dipartimenti. Ciò avrebbe consentito di sfruttare i risultati della valutazione ad un doppio livello: a livello di sede, rafforzando la coesione disciplinare unitaria della psicologia di quella sede, e a livello di settore, rafforzando la coesione nazionale tra gli attori della ricerca in quel settore. Un'altra occasione mancata di migliorare il sistema e, nel caso della psicologia, un'occasione mancata di sfruttare la valutazione per migliorarci.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Nemmeno nel caso dell'Abilitazione Scientifica Nazionale si può obiettare in radice che il principio ispiratore non sia cosa buona e giusta: tale meccanismo valutativo si prefigge infatti di pre-selezionare in modo qualificato gli aspiranti candidati a posizioni accademiche, misurandoli su un insieme di parametri oggettivi, e consentendo così di formare delle liste di nominativi che garantiscano la presenza di alcuni prerequisiti di qualità, a beneficio delle operazioni di reclutamento e progressione messe in atto da parte delle Università. Nonostante il suo principio di fondo e la sua finalità abbiano riscosso un consenso pressoché unanime, essendo stata posta la meritocrazia al centro del sistema di reclutamento, di tutte le operazioni di valutazione l'ASN è certamente quella che ha maggiormente contribuito ad agitare le coscienze e a cambiare la quotidianità degli accademici italiani, vuoi perché in ballo ci sono le vite delle persone, vuoi perché tra i valutatori e i valutati si è generata una vera e propria situazione di tensione generale, determinata dalla posta in gioco e terribilmente concentrata sulle regole del gioco. Anche chi (come il sottoscritto) non ha mai partecipato direttamente (né come candidato, né come valutatore) al meccanismo ASN, può comunque testimoniare come nell'ultimo decennio sia avvenuta una trasformazione radicale delle abitudini, dei tempi e modi di lavoro, dei contenuti delle conversazioni con i colleghi, sia senior che (soprattutto) junior. Inevitabilmente, anche se non coinvolti in prima persona, siamo stati tutti trascinati dal cambiamento, se non altro perché con i colleghi coinvolti direttamente ci interagiamo spesso e (più o meno) volentieri, e quello che il cambiamento ha comportato per loro non poteva non ripercuotersi anche su tutti gli altri. Questo è vero sia per la vita quotidiana nella sua dimensione più informale (dalle chiacchiere di corridoio ai suggerimenti da dare ai giovani circa la loro carriera), sia per aspetti di natura più istituzionale o formale. A questo proposito ricordo come se fosse ieri una conversazione che ebbi con un autorevole collega genetista. Doveva essere il 2006, chiacchieravamo di pubblicazioni, di *impact factor* delle riviste e di citazioni, e io gli descrissi l'algoritmo di calcolo dell'indice di Hirsch, o «indice H» (Hirsch, 2005), del quale lui non era ancora a conoscenza. Pochi anni più tardi quell'indice sarebbe divenuto, tanto per la genetica quanto per la psicologia (Nucci, Boscolo e Sartori, 2012), uno dei parametri di riferimento cruciali per valutare il livello qualitativo di docenti, ricercatori e aspiranti tali. Lungi da me l'idea di esprimere qui un parere favorevole o contrario all'esistenza di un indice bibliometrico, che di per sé assolve ad una mera funzione di operatore quantitativo, e la cui bontà dipende esclusivamente dall'intelligenza di chi ne fa uso. Rilevo

Estratti riservati ai soci AIP, © Società editrice il Mulino

semplicemente che se quattordici anni fa diversi di noi, come il mio collega genetista, potevano magari credere che l'indice H fosse una malattia affine all'alluce valgo, ormai il concetto è talmente radicato nell'immaginario comune da rendere quella chiacchierata con il collega genetista il mio personale marcatore autobiografico del cambio di scenario occorso, il confine tra il prima e il dopo dello tsunami valutativo. Al di là delle reminiscenze personali, non si può comunque non essere d'accordo che l'ASN, giunta ormai verso la conclusione del suo terzo esercizio, sia diventata lo strumento di valutazione che più di altri ha rivoluzionato, volenti o nolenti, favorevoli o contrari, interessati o meno, la vita di tutti. All'apprezzamento unanime circa i sacrosanti obiettivi che l'ASN si prefigge (cioè selezionare i migliori) e al rumore di fondo che fisiologicamente accompagna l'adeguamento a nuove regole, si è tuttavia affiancato sin dal suo esordio un ampio coro di critiche, perché l'obiettivo virtuoso del meccanismo è stato più volte messo in crisi dal metodo, e questo è stato certamente vero e tangibile nell'ambito della psicologia, rivelando in momenti diversi e per le più varie ragioni importanti turbolenze, alcune condivise con tutte le aree bibliometriche, altre interne alla disciplina.

Per avere un'idea più chiara di come si sia evoluto il reclutamento della psicologia in funzione di quest'ultimo decennio di ASN, si consideri la figura 1 (in tab. 1 sono riportati i valori numerici). In primo luogo, in essa è riportato l'andamento della numerosità dei professori ordinari, dei professori associati, dei ricercatori a tempo indeterminato e di quelli a tempo determinato di tipo B di ambito psicologico in servizio dal 2010 al 2019, per la precisione al 31 dicembre di ognuno di questi anni (dati ricavati dal database Cineca). Ogni punto di ciascuna di queste spezzate rappresenta la somma degli appartenenti alla corrispondente categoria in tutti gli otto settori scientifico-disciplinari psicologici da M-PSI/01 a M-PSI/08 (oppure, se si preferisce, di tutti e quattro i settori concorsuali psicologici da 11/E1 a 11/E4). Si può osservare come nel decennio si sia assistito quasi al raddoppio del numero totale dei professori associati (precisamente da 336 a 563); ad una diminuzione del numero dei professori ordinari (da 318 a 283) seppur con una prima lunga fase di trend decrescente, solo più recentemente in via di inversione; ad una diminuzione di più della metà del numero dei ricercatori a tempo indeterminato (da 548 a 217); alla comparsa a partire dal 2012, seguita da un progressivo aumento, dei ricercatori a tempo determinato di tipo B (da 1 a 113).

Due considerazioni immediate sull'aumento dei professori associati: *i)* la riduzione progressiva del numero dei ricercatori a tempo indeterminato a partire dal 2014 è quasi perfettamente simmetrica all'aumento del numero dei professori associati, ed è ben osservabile come ad una più ripida riduzione degli RU nel periodo 2013-2015

Estratti riservati ai soci AIP, © Società editrice il Mulino

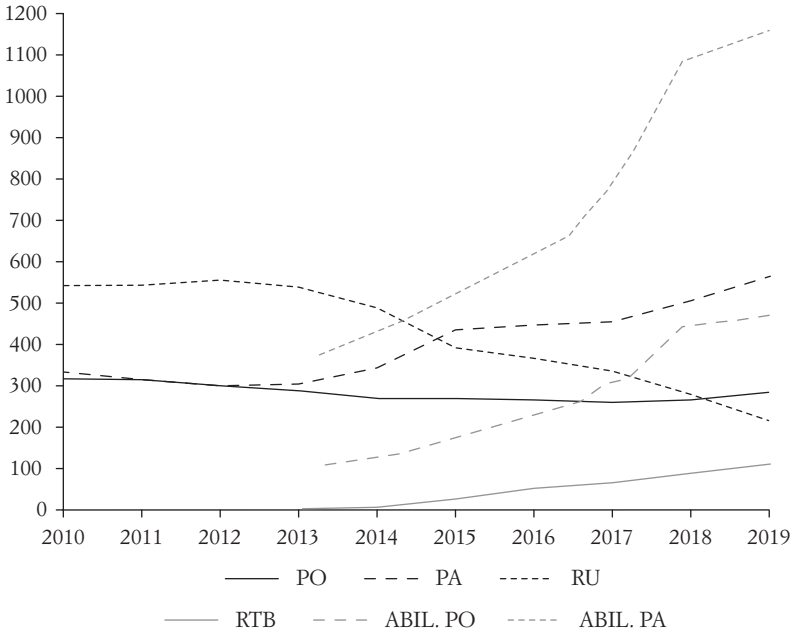


FIG. 1. Evoluzione del numero dei professori, dei ricercatori e degli abilitati di area psicologica dal 2010 al 2019. Ciascuna spezzata continua rappresenta l'andamento della somma degli strutturati M – PSI/01 ÷ M – PSI/08 (PO Professori ordinari; PA Professori associati; RU Ricercatori a tempo indeterminato; RTB Ricercatori a tempo determinato di tipo B) in servizio alla data del 31 dicembre di ciascuno degli anni riportati in ascissa. Le due spezzate tratteggiate rappresentano le funzioni cumulative del numero totale dei soggetti entrati in possesso di Abilitazione Scientifica Nazionale nei quattro settori concorsuali 11/E1 ÷ 11/E4 nelle dieci distinte tornate di abilitazione (per ciascuna tornata è stata considerata la data di pubblicazione del più tardivo tra i risultati delle quattro commissioni ASN). NB: le numerosità degli abilitati non sono depurate della quota di coloro che hanno conseguito una seconda volta l'abilitazione.

corrisponda la prima «ondata» di crescita del numero degli associati nel medesimo periodo, mentre la seconda ondata di aumento dei PA (2017-2019) trova solo parziale corrispondenza nella riduzione degli RU; infatti, *ii*) la *tenure track*, vale a dire il percorso di transizione da RTB a professore associato, subordinata ad un concorso valutativo verso il completamento del triennio del ricercatore, rende probabilmente conto della seconda «ondata» di crescita del numero dei professori associati, avvenuta in particolare dopo il 2016 (cioè dopo il completamento del triennio dei primi contingenti di RTB). Tutto ciò indica che la Riforma Gelmini è stata presa alla lettera nell'ambito della psicologia accademica portando ad un aumento davvero rilevante della fascia dei professori associati, sia in funzione di avan-

Estratti riservati ai soci AIP, © Società editrice il Mulino

TAB. 1. *Numerosità dei professori, dei ricercatori e degli abilitati di area psicologica dal 2010 al 2019 (nella riga relativa al 2019 sono inclusi anche i soggetti che hanno ricevuto l'abilitazione a gennaio 2020). NB: le numerosità degli abilitati non sono depurate della quota di coloro che hanno conseguito una seconda volta l'abilitazione*

	PO	PA	RU	RTB	TOT	ABIL. PO	ABIL. PA
2010	318	336	548	0	1202	0	0
2011	315	322	544	0	1181	0	0
2012	297	308	555	1	1161	0	0
2013	288	302	541	4	1135	0	0
2014	271	344	489	6	1110	109	385
2015	269	434	393	24	1120	139	468
2016	263	445	370	50	1128	139	468
2017	260	453	338	64	1115	303	776
2018	265	505	283	86	1139	451	1081
2019	283	563	217	113	1176	488	1159

zamenti di carriera di coloro che già erano ricercatori a tempo indeterminato, sia in funzione del meccanismo della *tenure track* che la stessa Riforma Gelmini ha introdotto. Questo ha creato tuttavia una asimmetrica e infelice competizione tra le due categorie nell'accesso alla sempre più inflazionata categoria dei professori associati**.** Se si considera che la diminuzione del numero dei professori ordinari nel decennio risponde anche al dato del progressivo pensionamento dei colleghi di questa fascia, l'aumento dei professori associati appare ancora più impressionante, perché anche molti professori associati – così come anche molti ricercatori a tempo indeterminato – sono stati posti in quiescenza nel decennio, e tuttavia il trend è stato decisamente crescente. È ancora più interessante andare a considerare questi andamenti alla luce delle abilitazioni di area psicologica, pure riportate in figura 1 (vedere la tab. 1 per i valori numerici). Nella figura (e nella tabella) sono infatti presentati anche i valori cumulativi del numero degli abilitati a professore associato e a professore ordinario, risultanti cioè dalla somma degli abilitati (per ciascuna tornata) dei quattro settori concorsuali di area 11/E. Appare evidente come il numero degli abilitati di entrambe le fasce risulti crescere in modo molto più accelerato del numero dei docenti della fascia corrispondente, con la conseguenza che ad oggi ci sono in Italia poco meno di 1200 abilitati a professore di seconda fascia e poco meno di 500 a professore di prima fascia in area psicologica, sebbene in questa ricognizione le numerosità degli abilitati non sono state depurate né della quota di coloro che hanno conseguito una seconda volta l'abilitazione, né del numero totale degli abilitati divenuti nel frattempo professori associati. Risulta in ogni caso palese che il potenziale di ulteriore crescita del numero dei professori associati non potrà cer-

Estratti riservati ai soci AIP, © Società editrice il Mulino

tamente reggere il passo a fronte del numero assoluto raggiunto dagli abilitati (e destinato ad aumentare), con buona pace di coloro che avendo conseguito l'abilitazione dovranno ripiegare su altre carriere, e di quei ricercatori a tempo indeterminato che dovranno «farsene una ragione». Ancora più preoccupante appare il fatto che, se dal punto di vista del sistema generale immaginato dalla Riforma Gelmini, l'aspettativa è che si continui a mantenere la cosiddetta «piramide» (pochi professori ordinari, più professori associati, ancor più ricercatori), il numero dei professori associati in capo ad un altro decennio dovrebbe crescere così tanto da far diventare un problema la previsione di copertura finanziaria che potrà continuare a mantenere in piedi il sistema stesso. È vero che il tasso di ricambio è ancora alquanto insoddisfacente e che la numerosità relativa degli addetti all'università italiana è molto inferiore a quella di moltissimi altri paesi (come conseguenza della bassa quota di investimento in ricerca del PIL nazionale); tuttavia, se lo Stato non metterà mano a pesanti interventi di rilancio finanziario dell'università, la prospettiva è paradossalmente quella di assistere proprio ad una inevitabile limitazione del numero di RTB reclutabili, e quindi al venir meno dei presupposti stessi del sistema. Dal momento che in questa ricognizione non sono stati affatto considerati i ricercatori a tempo determinato di tipo A e nemmeno gli assegnisti di ricerca, che costituiscono certamente un ampio bacino di interessati al reclutamento di ulteriori professori associati, sembra ragionevole spostare il discorso anche sulle condizioni di accesso a quel precariato, costituite essenzialmente dal conseguimento del titolo di dottore di ricerca. Poiché sarebbe assurdo tentare di risolvere alla radice il problema del reclutamento riducendo il numero dei dottori di ricerca (e il successivo bacino di «precariato abilitato»), l'ovvia soluzione prospettabile per realizzare concretamente il sistema piramidale e contemporaneamente aumentare la numerosità complessiva dell'università italiana sembra essere quella del ripristino della posizione di ricercatore a tempo indeterminato, ferma restando l'utilità dell'abilitazione per il passaggio ai ruoli di professore. Un discorso analogo a quello fatto per i professori associati vale anche per la numerosità dei professori ordinari, rispetto alla quale è altrettanto evidente che il numero crescente degli abilitati PO è decisamente troppo alto per poter soddisfare una qualsivoglia speranza di progressione in tutti i sempre più numerosi aspiranti.

Queste considerazioni sulla psicologia nel suo insieme riflettono quanto sta diffusamente accadendo come conseguenza dell'Abilitazione Scientifica Nazionale nel sistema universitario italiano, tuttavia è importante sottolineare almeno un aspetto che ha sicuramente valore per la psicologia. Se il numero totale di docenti e ricercatori è rimasto praticamente costante nel decennio (colonna TOT in tab. 1), ciò

Estratti riservati ai soci AIP, © Società editrice il Mulino

è particolarmente suggestivo di un aumento di quella ambita qualità che il sistema ASN mirava ad incentivare: la natura dei meccanismi di abilitazione aveva posto infatti nella definizione di parametri quantitativi e nel loro aggiornamento delle asticelle sempre più elevate. Che si consideri o meno il superamento di soglie bibliometriche quale valido criterio per definire la tanto decantata qualità, si deve prendere atto che nonostante la progressiva elevazione degli standard di accesso, il numero totale dei docenti di area psicologica è rimasto pressoché invariato e non è diminuito. Ciò significa che i nuovi professori di psicologia sono figli di una risposta molto *compliant* alla sfida valutativa. Questo aspetto, come ben si sa, cela tuttavia delle insidie di non poco conto, perché rimanda alla questione delle *modalità* con cui l'abilitazione è conseguita, spesso dipendenti dalla moltiplicazione opportunistica del numero delle pubblicazioni e dal valzer delle citazioni (Baccini, De Nicolao e Petrovich, 2019; Bruton, Medlin, Brown e Sacco, 2020), e dall'inarrestabile aggiornamento al rialzo delle soglie bibliometriche, di cui tanto si è parlato nei mesi recenti anche da parte dei non addetti ai lavori (ad esempio Magnani, 2019). Tuttavia, lasciando ad altri l'approfondimento di questi aspetti spinosi, preferisco qui spendere qualche parola di maggior dettaglio sul rapporto tra numerosità del personale in servizio e degli abilitati ASN. Considerando separatamente i quattro settori concorsuali della psicologia, è utile ripercorrere gli stessi ragionamenti fatti più sopra per ciascun settore, verificando se ci siano delle differenze tra settori. Come si può vedere in Figura 2, ove sono riportati gli andamenti delle numerosità dei docenti, dei ricercatori e degli abilitati nel decennio per ciascuno dei quattro settori concorsuali (in tab. 2 i dati numerici), l'andamento generale risulta essere abbastanza simile (aumento dei professori associati, diminuzione dei ricercatori a tempo indeterminato, eccedenza di abilitati), tuttavia si possono notare delle differenze rilevanti tra i settori. La prima e più rimarchevole peculiarità è la diversa situazione di eccedenza degli abilitati nei quattro settori concorsuali, che vede nel settore 11/E1 un rapporto molto più problematico tra il numero degli abilitati ed il numero dei candidati «naturali» (RU e RTB), rispetto agli altri tre settori di area psicologica. A tale proposito è utile calcolare un indice che esprima tale rapporto ricorrendo a delle assunzioni ragionevoli. Restringendo il ragionamento alla sola fascia dei professori associati che, come si è visto, rappresenta la fascia portatrice della maggior «criticità di sistema», sarà calcolato un indicatore dato dal rapporto che pone al numeratore la somma degli attuali RU e RTB (i candidati naturali alla seconda fascia), e al denominatore il numero totale degli abilitati di seconda fascia nel decennio, depurato dal numero di coloro che sono già diventati professori di seconda fascia a partire dalle prime abilitazioni ASN (cioè dal 2013). Un valore

Estratti riservati ai soci AIP, © Società editrice il Mulino

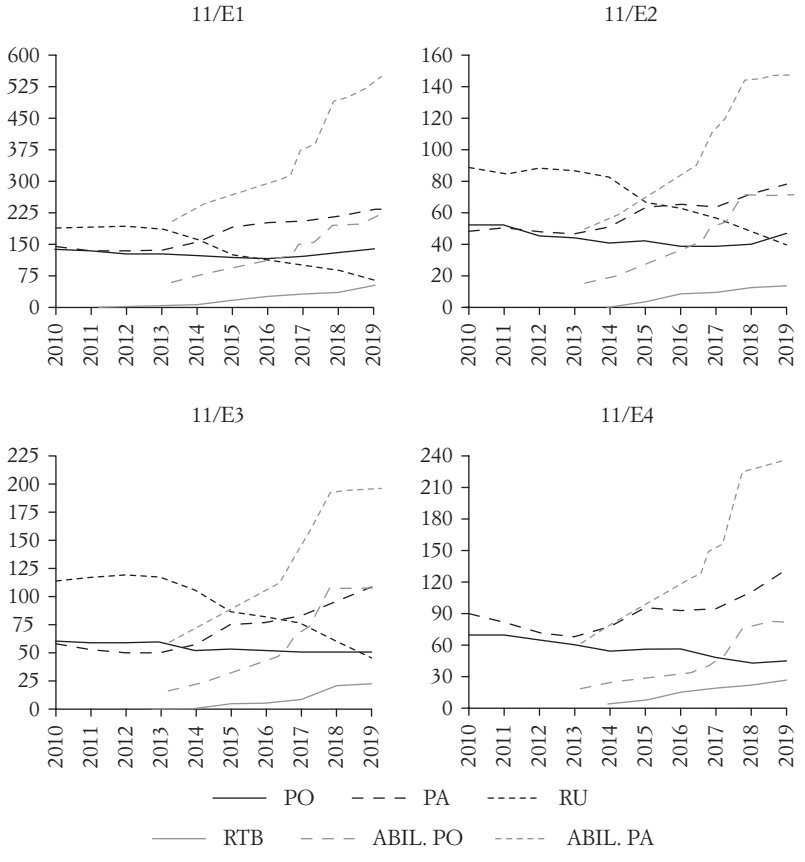


FIG. 2. Evoluzione del numero dei professori, dei ricercatori e degli abilitati dal 2010 al 2019 per i quattro settori concorsuali della psicologia (si faccia attenzione alle diverse scale utilizzate in ordinata).

dell'indicatore pari a 1 rappresenta la situazione più auspicabile (parità tra abilitati e candidati naturali), un valore dell'indicatore inferiore a 1 rappresenta una situazione problematica di «troppa disponibilità qualificata» (eccedenza di abilitati rispetto ai candidati naturali), un valore dell'indicatore superiore a 1 rappresenta una situazione problematica nel verso opposto di «troppo poca disponibilità qualificata» (carenza di abilitati tra i candidati naturali). Si può calcolare che per il SC 11/E1 (Psicologia generale, psicobiologia e psicometria) tale rapporto è pari a 0.24, per il SC 11/E2 (Psicologia dello sviluppo e psicologia dell'educazione) il rapporto è 0.44, per il SC 11/E3 (Psicologia sociale, del lavoro e delle organizzazioni) il rapporto è 0.5,

Estratti riservati ai soci AIP, © Società editrice il Mulino

TAB. 2. *Numerosità dei professori, dei ricercatori e degli abilitati dal 2010 al 2019 nei quattro settori concorsuali della psicologia (incluso il terzo quadrimestre ASN 2018, i cui risultati sono stati pubblicati a gennaio 2020)*

11/E1	PO	PA	RU	RTB	TOT	ABIL. PO	ABIL. PA
2010	138	140	193	0	471	0	0
2011	136	137	193	0	466	0	0
2012	129	138	196	1	464	0	0
2013	126	138	191	3	458	0	0
2014	124	159	166	4	453	57	217
2015	119	198	126	13	456	75	256
2016	117	208	116	25	466	75	256
2017	123	210	102	31	466	150	380
2018	132	224	87	34	477	202	512
2019	141	241	62	52	496	225	564
11/E2							
2010	52	49	89	0	190	0	0
2011	52	51	85	0	188	0	0
2012	45	48	89	0	182	0	0
2013	44	47	87	0	178	0	0
2014	41	51	83	0	175	17	47
2015	42	64	67	3	176	20	59
2016	39	66	63	8	176	20	59
2017	39	64	57	9	169	50	110
2018	40	73	48	12	173	68	149
2019	47	79	40	14	180	69	154
11/E3							
2010	62	61	119	0	242	0	0
2011	61	56	123	0	240	0	0
2012	61	53	125	0	239	0	0
2013	61	52	123	1	237	0	0
2014	55	60	111	1	227	19	62
2015	55	79	91	4	229	24	77
2016	54	81	86	5	226	24	77
2017	53	87	79	8	227	65	144
2018	53	101	63	21	238	112	202
2019	53	115	48	23	239	114	205
11/E4							
2010	66	86	147	0	299	0	0
2011	66	78	143	0	287	0	0
2012	62	69	145	0	276	0	0
2013	57	65	140	0	262	0	0
2014	51	74	129	1	255	16	59
2015	53	93	109	4	259	20	76
2016	53	90	105	12	260	20	76
2017	45	92	100	16	253	38	142
2018	40	107	85	19	251	69	218
2019	42	128	67	24	261	80	236

e per il SC 11/E4 (Psicologia clinica e dinamica) il rapporto è 0.52. Appare di tutta evidenza che tale indicatore, che per tutti i settori

Estratti riservati ai soci AIP, © Società editrice il Mulino

concorsuali ben esprime un'eccedenza di disponibilità qualificata rispetto al bacino dei candidati naturali, mette in luce un rapporto di circa 2:1 per tutti i settori concorsuali psicologici, eccezion fatta per il settore 11/E1, per il quale il rapporto è di circa 4:1. Inutile appellarsi al fatto che il numero dei settori scientifico-disciplinari «costitutivi» varia da un settore concorsuale all'altro (11/E1 è costituito da 3 SSD, 11/E2 da 1 SSD, 11/E3 ed 11/E4 da 2 SSD), o che le numerosità assolute di riferimento dei tre settori concorsuali sono diverse, perché il rapporto rappresenta una misura relativa e indipendente dalla numerosità di riferimento. In buona sostanza, nel SC 11/E1 esiste una eccedenza di disponibilità qualificata addirittura doppia rispetto agli altri SC, probabilmente perché 1) esso è il SC che meno degli altri ha da offrire in termini di percorsi professionali non accademici (attraendo così più giovani con un'aspettativa nell'abilitazione), e forse anche perché 2) il SC è per sua tradizione in qualche modo «preadattato» ad una competizione bibliometrica (le soglie sono infatti le più alte di tutti e quattro i settori, per quel che può contare tale confronto). Tale spiegazione può essere invocata anche ponendo attenzione alla seconda differenza, più sottile, tra i grafici relativi ai quattro settori riportati in figura 2, consistente nell'intersezione tra la curva discendente dei RU e la curva ascendente dei PA. Se il momento storico corrispondente a tale intersezione rappresenta un indicatore semplificato ma ragionevole della «prontezza di risposta» del settore al sistema valutativo, si può notare come questa risposta sia stata la più rapida per 11/E1 (2014), un po' più lenta per 11/E2 (2015-2016), ancor più lenta per 11/E3 (2016-2017) e la più lenta per 11/E4 (2017-2018). La diversa progressione temporale nell'adeguamento al sistema di reclutamento con abilitazione potrebbe essere spiegata anche con altre interpretazioni, ad esempio con differenze dell'anzianità media dei ricercatori a tempo indeterminato dei diversi SC, oppure tenendo presente che le discipline psicologiche godono di due diversi statuti sul piano didattico (base e caratterizzante) nella definizione quanto meno degli ordinamenti triennali, e quindi la necessità di docenti di SSD di base *versus* SSD caratterizzanti potrebbe aver imposto una priorità di reclutamento scaglionata nei vari settori. Sono tutte interpretazioni plausibili (che per altro non escludono quella della diversa «velocità di adeguamento»), e altre ancora se ne potrebbero trovare, tuttavia non mi avventuro oltre, sperando di aver sollecitato in altri la voglia di approfondire ulteriormente queste valutazioni differenziali. Infatti, ritornando dal particolare al generale e prendendo in esame l'insieme della psicologia accademica, l'indicatore risulta essere pari a 0,36 sui dati aggregati delle abilitazioni a professore associato, cioè un rapporto di 3:1 tra abilitati e candidati naturali. L'anno di svolta nella transizione bibliometrica del reclutamento dei professori associati

Estratti riservati ai soci AIP, © Società editrice il Mulino

dell'intera psicologia, consultando al volo la figura 1, appare infine essere stato il 2015.

Un'ultima analisi non può non focalizzarsi sulla fascia dei professori ordinari. Calcolando per l'intero aggregato della psicologia l'indice già utilizzato per i professori associati, ma assumendo in questo caso il numero degli associati (candidati naturali alla prima fascia) al numeratore, ed al denominatore il numero degli attuali abilitati alla prima fascia depurato dal numero dei professori di prima fascia che hanno preso servizio nel periodo 2013-2019, si ottiene un valore pari a 1,14. Per il SC 11/E1 si ottiene un valore identico a quello aggregato (1,14), per il SC 11/E2 il valore è pari a 1,19, per il SC 11/E3 il valore è 0,94, e per il SC 11/E4 il valore è 1,34. In base alla definizione dell'indicatore, sembra possibile affermare che nel caso delle abilitazioni a professore ordinario la situazione numerica della psicologia sia decisamente più in equilibrio che nel caso delle abilitazioni a professore associato, probabilmente anche a causa delle condizioni meno permissive rappresentate dai requisiti bibliometrici più elevati da soddisfare per la prima fascia. Va però tenuto presente che questi indicatori di abilitazione e reclutamento dei professori ordinari sono in un certo qual modo sensibili agli effetti a cascata dell'inflazione di associati descritta più sopra, dal momento che il numeratore dell'indice include la grande quota di nuovi professori associati immessi in servizio nell'ultimo sessennio, che sappiamo tendere all'aumento e abbiamo visto essere in grado di adeguarsi a condizioni sempre più sfidanti. Si può quindi prevedere che la dinamica dell'evoluzione dell'indicatore nei prossimi anni sia destinata ad assomigliare sempre più a quella già vista per gli associati. Ne consegue una scontata ma inevitabile raccomandazione per i direttori dei dipartimenti ed i rettori che hanno a cuore la psicologia: dal momento che nei bacini di interessati le abilitazioni crescono (e probabilmente continueranno a farlo) come i funghi, le strategie del reclutamento negli atenei, nelle scuole e nei dipartimenti dovranno costituire il vero locus di selezione, ragionando d'anticipo (e con il bilancino) sulla quantità di bandi delle varie fasce, conciliando il reclutamento di nuovi RTB (necessari per mantenere la piramide gerarchica, ma destinati a raddrizzarne automaticamente la forma dopo tre anni) con le legittime aspirazioni di progressione di chi è già in servizio, contemperando inoltre le diverse quote di disponibilità di abilitati nei diversi settori concorsuali psicologici (che si traducono in valori di «pressione umana» spesso difficili da governare), e infine cercando di soddisfare appieno e al contempo sia i requisiti di docenza AVA dei corsi di studio, sia i parametri di qualità VQR della produzione scientifica e della terza missione della struttura (ateneo e dipartimento). Non una *mission impossible*, ma nemmeno una passeggiata.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Un altro argomento che vale la pena discutere, riprendendo quanto già evidenziato nelle precedenti considerazioni su AVA, VQR e psicologia, riguarda le peculiarità della disciplina e il peso delle sue articolazioni all'interno del meccanismo valutativo di ASN. Ciò solleva ad esempio il problema dello scorporo delle soglie per SSD diversi all'interno di alcuni settori concorsuali, che non poca rilevanza ha avuto nell'accendere il dibattito e nel creare malumori diffusi, in particolare nel settore 11/E1 con la divisione territoriale (successiva ad unificazione) tra la psicobiologia ed «il resto del settore concorsuale». Tale vicenda resta a mio parere una ferita aperta, dal momento che ha riguardato (al di là delle ovvie implicazioni corporativistiche relative all'ASN) importanti riflessioni sullo statuto della disciplina e sulla sua natura, soprattutto alla luce del cambiamento storico che la psicologia, nelle sue componenti teoretiche, metodologiche e applicative, sta vivendo sia al suo interno, sia nell'interazione sempre più intensa con altri ambiti del sapere, che a loro volta stanno subendo importanti mutazioni. Senza voler focalizzare la questione sul rapporto tra psicologia e neuroscienze invocato nella vicenda del SC 11/E1, ma cercando di trarre da quella diatriba dei principi di utilità generale per la psicologia, è importante sottolineare che se alla base di certe operazioni funzionali ai meccanismi ASN vi sono ragioni di parte (mascherate ad esempio da questioni di numerosità dei settori scientifico-disciplinari), esse mal si sposano con una valutazione ponderata e condivisa dello statuto della disciplina, anzi finiscono in qualche modo per stabilirne surrettiziamente la fisionomia senza il necessario coinvolgimento della comunità scientifica di riferimento. Da questo punto di vista è chiaro che il ruolo svolto dalle società scientifiche, ma anche quello offerto dalle occasioni di discussione pubblica, come il presente articolo vuole sollecitare, dovrebbe essere proprio quello di mantenere vigile e aggiornata la consapevolezza dell'identità della disciplina e delle sue articolazioni in un mondo ed in una società che cambiano, e non solo in caso di emergenza (a questo proposito si vedano Cubelli, 2018; Di Nuovo e Cassibba, 2018).

5. PRIN E PSICOLOGIA

In quest'ultima parte, seppur molto più brevemente che nelle precedenti, il fuoco dell'attenzione si sposta da quei meccanismi valutativi costruiti per saggiare la qualità di strutture, corsi di studio e persone, ai finanziamenti per la ricerca, in particolare ai Progetti di Rilevante Interesse Nazionale (PRIN), l'ultimo bando dei quali ha rappresentato una fonte di importanti scontenti e malumori. Si dirà subito che è tipico degli accademici, un po' come nel tifo sportivo,

Estratti riservati ai soci AIP, © Società editrice il Mulino

lamentarsi in caso di pubblicazioni e domande di finanziamenti rigettati, attribuendone la ragione agli arbitri che di quelle pubblicazioni o progetti non hanno saputo valutare oggettivamente la bontà, o ad altri fattori estrinseci. Credo infatti che sia opportuno mantenere una posizione di maggior cautela sul versante del rapporto tra valutazione e finanziamenti, partendo dall'assunto che la valutazione messa in campo in un meccanismo di redistribuzione di fondi pubblici per la ricerca deve tenere conto di un obiettivo che, per quanto imparentato, non è affatto identico a quello che caratterizza la valutazione delle persone, delle strutture e dei corsi di studio. Sebbene anche questi soggetti possano derivare benefici economici dai sistemi AVA e VQR (leggasi FFO, Dipartimenti di Eccellenza, FFARB, ecc.), in questi sistemi è evidente che devono essere soddisfatti dei criteri di minimo relativamente a cose e persone in funzione del loro presente e della loro storia pregressa, cioè deve essere dato riscontro del raggiungimento di determinati standard qualitativi. Diversamente, nel caso dei finanziamenti per la ricerca, è evidente che devono essere adottati criteri diversi, perché le risorse in gioco riguardano essenzialmente «cosa futura», vale a dire progetti sulla cui bontà (e possibilità di successo) la valutazione assolve un ruolo di giudizio più simile a quello richiesto di fronte a un investimento o, per certi versi, a una scommessa. È anche vero che pure nel caso dell'ASN si tratta di valutare il presente e il passato delle persone in vista della loro potenziale capacità futura di ricoprire ruoli didattici, di ricerca e gestionali, tuttavia si deve fare attenzione tra abilitazione (possesso dei requisiti minimi da parte di una specifica persona *oggi*) e procedura di reclutamento, in cui si dovrebbe valutare quale candidata/o, tra molti con quei requisiti, è la/il più meritevole per ricoprire una posizione *domani*. Tuttavia, anche se il reclutamento ha un obiettivo prospettico, esso non può che avvenire in base a caratteristiche che possono essere anche predittive di potenzialità future, ma che devono inevitabilmente attenersi a ciò che un curriculum di titoli e conseguimenti pregressi racconta (Moher, Naudet, Cristea, Miedema, Ioannidis e Goodman, 2018).

Della valutazione dell'ultima tornata dei PRIN è stata contestata la sostanziale irrilevanza del valore dei progetti in sé, che dalle commissioni valutatrici sono stati considerati soltanto a valle della valutazione dei profili dei proponenti. Alcune indicazioni di massima che erano state rese esplicite dal bando PRIN rimandavano fin da subito al fatto che l'ispirazione proveniva dai bandi dell'*European Research Council* (ERC), che solitamente premiano in modo generoso pochi progetti collocati al livello *top-notch* (il vertice della piramide qualitativa). L'aspettativa sull'ultimo bando PRIN, i cui predecessori erano stati criticati in passato per altre ragioni certamente non estranee al tema valutativo anche nella psicologia (consorterie, scambio di favori,

Estratti riservati ai soci AIP, © Società editrice il Mulino

ecc.), è stata quindi molto forte da parte dei proponenti dei progetti, perché il nuovo bando lasciava presagire la rottura con schemi passati di cordate, favoritismi e mafiette, cedendo finalmente il posto a modalità di selezione più oggettive e meritocratiche. Date le premesse, l'aspettativa non poteva certo essere anche quella di veder premiata una quota molto più ampia che una piccola percentuale dei progetti presentati – evidentemente i migliori progetti. Che ad essere premiati siano stati in pochi è un dato di fatto che ha confermato le previsioni, ma si è tuttavia segnalato da più parti – e va sottolineato in questa sede – che i progetti finanziati di ambito psicologico sono stati comunque più numerosi nell'ultimo bando PRIN che in quelli passati. Questo è certamente vero, e a mio parere è un risultato importante che – preso nel suo complesso – merita ampia considerazione: l'aver riconfigurato l'impianto valutativo dei PRIN adottando il formato europeo è stato un successo per la psicologia sul piano del risultato. Quel che è piaciuto meno, e da questo punto in poi avverto la presenza di sabbie mobili, è che ad essere premiati siano stati soprattutto coloro i quali avevano già avuto in passato importanti finanziamenti, o i cui curriculum vantavano liste di pubblicazioni imbattibili, finendo per sacrificare alcuni ambiti della psicologia a discapito di altri, e probabilmente – a parità di valore progettuale – anche alcuni progetti presentati da proponenti meno competitivi a discapito di altri progetti il cui valore è stato rafforzato dai *pedigree* dei proponenti.

Va rilevato che nella valutazione dei progetti ERC ai quali l'ultimo bando PRIN si sarebbe ispirato, il giudizio – che procede in due fasi consecutive – non pone, come pare sia accaduto in Italia, un'attenzione disgiunta e seriale prima sul solo valore del proponente e poi sul solo valore del progetto, ma guarda ad entrambi gli aspetti in entrambe le fasi della valutazione. C'è tuttavia da tenere presente che nella prima fase della valutazione dei PRIN i profili dei proponenti esaminati dalle commissioni sono stati normalizzati per l'età accademica e per il Settore Scientifico-Disciplinare (rispetto alle mediane della fascia immediatamente superiore a quella del proponente), il che dovrebbe aver avuto l'effetto di rendere comparabili le differenze tra i candidati, quanto meno in termini di pubblicazioni. È inutile comunque dire che anche secondo il buon senso la valutazione determinante non potrebbe che essere – verificata preliminarmente la consistenza scientifica del proponente – quella centrata sul progetto. È cioè ragionevole che, se mai si dovesse disgiungere il giudizio separando il valore di un progetto in termini di innovatività e originalità dalla collocazione scientifica del suo proponente, l'ordine dei fattori valutati da parte di un giudice interessato a scommettere assumendo il livello minimo di rischio non potrebbe che essere quello di fatto adottato nell'ultimo bando PRIN, e cioè profili prima e progetti poi. La garan-

Estratti riservati ai soci AIP, © Società editrice il Mulino

zia costituita dalla storia pregressa dei proponenti dei progetti sembra quindi spiegare abbastanza bene le scelte fatte, che probabilmente non sono state molto coraggiose, ma che si sono ancorate agli aspetti più diagnostici in termini di probabilità di successo futuro, vale a dire il curriculum dei proponenti. Non condivido fino in fondo una critica a tali criteri, perché nell'essere volutamente prudenziali essi ci informano che il giudizio dei valutatori ha fortemente risentito di un mandato pubblico, e quindi del peso della responsabilità etica nell'attribuzione di cospicue risorse finanziarie, il che è per certi versi rassicurante. Casomai appare ragionevole immaginare una differenziazione, in futuro, tra progetti meritevoli di attenzione indipendentemente dal curriculum del proponente, e progetti maggiormente ancorati ad esso (d'altra parte il modello ERC prevede diverse articolazioni di bandi proprio in tal senso), perché la psicologia avrebbe davvero un enorme potenziale di progetti promettenti. Non c'è molto da sperare, tuttavia, poiché le iniziative di finanziamento alla ricerca nel nostro Paese soffrono di una annosa problematica di frammentarietà e sporadicità, tanto da non poter essere nemmeno sicuri che di altri bandi PRIN da criticare ce ne saranno ancora in futuro.

6. CONCLUSIONI

Scrivendo questo articolo sulla valutazione mi sono purtroppo reso conto di non aver potuto affrontare – per evidenti ragioni di spazio – tantissimi dei punti dolenti e delle complesse e articolate implicazioni per la psicologia che il dipanarsi del discorso mi suggeriva mano a mano. Sono tuttavia persuaso che l'aver passato in rassegna in modo almeno parzialmente organizzato tanti aspetti diversi collegati alla valutazione non sarà stato del tutto inutile, e in un certo senso sono soddisfatto di non aver nemmeno voluto affrontare molti di quegli aspetti, perché al di là delle mie considerazioni e di qualche dato portato a beneficio della comunità della psicologia accademica, confido che questo contributo potrà generare discussione in molteplici direzioni, quanti possono essere i risvolti ed i nessi rilevanti per la psicologia. Sono certo che chi avrà avuto la pazienza di leggere questo articolo fino in fondo sarà stato sufficientemente sollecitato a riflettere su uno o più specifici punti di interesse che qui non sono stati toccati o che non sono stati sviluppati a dovere. Si pensi ad esempio al rapporto tra i diversi ambiti di valutazione in chiave di prospettiva futura per la psicologia, ad esempio nell'evoluzione dimensionale delle sedi: se alla valutazione della qualità delle strutture conseguiranno attribuzioni differenziali di risorse finanziarie, le variazioni geografiche nel reclutamento alle quali probabilmente assisteremo cambieranno il

Estratti riservati ai soci AIP, © Società editrice il Mulino

destino della psicologia italiana, anche in rapporto alla sua storia? La qualità dei corsi di studio psicologici può essere garantita dalla qualità della ricerca dei docenti che in essi insegnano? Una delle conseguenze indotte dalla valutazione è che di questioni come queste, essendosi moltiplicati i costrutti e i loro effetti, se ne possono generare innumerevoli. Ovviamente non saranno tutte questioni cruciali, il che dovrebbe condurre a scremare i meccanismi davvero importanti dalle sovrastrutture inutili, tuttavia tra le tante suggestioni future sono convinto che ve ne saranno di interessanti, che un domani ci permetteranno di vedere la valutazione come qualcosa che sarà stato utile non solo a misurare lo stato di salute della psicologia, ma – c'è da augurarsi – a meglio definirne e consolidarne l'unità e l'identità.

BIBLIOGRAFIA

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2011). Valutazione della Qualità della Ricerca 2004-2010 (VQR 2004-2010). Bando di partecipazione, 7 novembre 2011 (disponibile al link: https://www.anvur.it/wp-content/uploads/2011/11/bando_vqr_def_07_11.pdf).

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2013). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area 11 (GEV11) (disponibile al link: https://www.anvur.it/rapporto/files/Area11/VQR2004-2010_Area11_RapportoFinale.pdf).

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2015). Valutazione della Qualità della Ricerca 2011-2014 (VQR 2011-2014). Bando di partecipazione, 11 novembre 2015 (disponibile al link: https://www.anvur.it/wp-content/uploads/2015/11/Bando%20VQR%202011-2014_secon~.pdf).

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2017). Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area 11b – Scienze Psicologiche (GEV11b) (disponibile al link: https://www.anvur.it/rapporto-2016/files/Area11b/VQR2011-2014_Area11b_RapportoFinale.pdf).

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2019). Nota di accompagnamento alla pubblicazione del Rapporto degli Alti Esperti VQR (disponibile al link: https://www.anvur.it/news/pubblicazione-rapporto-alti-esperti-vqr/).

Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (2020). Valutazione della Qualità della Ricerca 2015-2019 (VQR 2015-2019). Bando di partecipazione, 3 gennaio 2020 (disponibile al link: https://www.anvur.it/wp-content/uploads/2020/01/Bando-VQR-2015-2019_3_1_2020.pdf).

Baccini A., De Nicolao G., Petrovich E. (2019). Citation gaming induced by bibliometric evaluation: A country-level comparative analysis. *PloS One*, *14* (9).

Baldissera A. (2004). La valutazione nelle università italiane: processi organizzativi ed indicatori. *Quaderni di Sociologia*, *35*, 5-21.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Benedictus R., Miedema F., Ferguson M.W.J. (2016). Fewer numbers, better science. *Nature*, *538* (7626), 453-455.
Bonaccorsi A. (2018). *The Evaluation of research in social sciences and humanities: Lessons from the Italian experience*. Cham: Springer International Publishing AG.
Boring E.G. (1923). Intelligence as the tests test it. *New Republic*, *36*, 35-37.
Borrelli D., Stazio M. (2018). La «grande trasformazione» dell'università italiana. *Rivista Trimestrale di Scienza dell'Amministrazione: Studi di Teoria e Ricerca Sociale*, *issue 1/2018*.
Bruton S.V., Medlin M., Brown M., Sacco D.F. (2020). Personal motivations and systemic incentives: Scientists on questionable research practices. *Science and Engineering Ethics*, in corso di stampa.
Consiglio Universitario Nazionale (2020). Analisi e proposte su «Bando ANVUR per la valutazione della qualità della ricerca (VQR) 2015-2019». (disponibile al link: https://www.cun.it/uploads/7231/analisi_e_proposta_2020_01_16.pdf?v=).
Cubelli R. (2015). La psicologia italiana tra scienza della valutazione e valutazione della scienza. *Giornale Italiano di Psicologia*, *42* (4), 647-678,.
Cubelli R. (2018). Il farsi e il disfarsi dei Settori Scientifico-Disciplinari: storia di un'invenzione italiana. *Giornale Italiano di Psicologia*, *45* (3), 455-479.
Decataldo A. (2018). Valutare la didattica universitaria: considerazioni sui principi ispiratori e sui processi. *Rivista Trimestrale di Scienza dell'Amministrazione: Studi di Teoria e Ricerca Sociale*, *issue 1/2018*.
Deming W.E. (1982). *Quality productivity and competitive position*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.
Di Nuovo S., Cassibba R. (2018). Per una classificazione dei saperi e una «manutenzione» dei sistemi formativi relativi alla psicologia: alcune riflessioni sulle criticità attuali e sulle possibili alternative. *Giornale Italiano di Psicologia*, *45* (2), 265-278.
European Higher Education Area (1999). Bologna Declaration (disponibile al link: http://ehea.info/media.ehea.info/file/Ministerial_conferences/02/8/1999_Bologna_Declaration_English_553028.pdf).
European Higher Education Area (2001). Prague Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2001_Prague/44/2/2001_Prague_Communique_English_553442.pdf).
European Higher Education Area (2003). Berlin Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2003_Berlin/28/4/2003_Berlin_Communique_English_577284.pdf).
European Higher Education Area (2005). Bergen Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2005_Bergen/52/0/2005_Bergen_Communique_english_580520.pdf).
European Higher Education Area (2007). London Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2007_London/69/7/2007_London_Communique_English_588697.pdf).
European Higher Education Area (2009). Leuven Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2009_Leuven_Louvain-la-Neuve/06/1/Leuven_Louvain-la-Neuve_Communique_April_2009_595061.pdf).
European Higher Education Area (2015). Yerevan Communiqué (disponibile al link: http://www.ehea.info/media.ehea.info/file/2015_Yerevan/70/7/YerevanCommuniqueFinal_613707.pdf).

Estratti riservati ai soci AIP, © Società editrice il Mulino

European Higher Education Area (2018). Paris Communiqué (disponibile al link: http://ehea.info/Upload/document/ministerial_declarations/EHEA-Paris2018_Communique_final_952771.pdf).
Graff H.J. (2015). *Undisciplining knowledge: Interdisciplinarity in the twentieth century*. Baltimore, MD: Johns Hopkins University Press.
Hirsch J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 16569-16572.
Kano N., Nobuhiku S., Fumio T., Shinichi T. (1984). Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control*, *14*, 39-48.
Laudanna A. (1995). Problemi della valutazione della ricerca in psicologia. *Giornale Italiano di Psicologia*, *22* (1), 9-24.
Laudanna A. (1996). Proposte per la valutazione della ricerca in psicologia. *Giornale Italiano di Psicologia*, *23* (1), 9-26.
Laudanna A., Miceli M., Welin A.C. (2001). La valutazione della produzione scientifica: criteri e proposte dell'Istituto di Psicologia del CNR. *Giornale Italiano di Psicologia*, *28* (3), 611-631.
Legrenzi P., Malaguti D., Umiltà C. (2017). Formazione universitaria e professioni: il caso peculiare della psicologia. *Giornale Italiano di Psicologia*, *34* (3), 509-530.
Magnani A. (2019). Boom di auto-citazioni, ecco perché la ricerca italiana è «dopata». *Sole 24 Ore*, 11 settembre 2019.
Ministro dell'Istruzione, dell'Università e della Ricerca (2011). Decreto Ministeriale 15 luglio 2011 n. 17 «Linee guida valutazione qualità della ricerca (VQR) 2004-2010» (disponibile al link: https://www.anvur.it/wp-content/uploads/2011/11/vqr_d.m._n._17_del_15_07_2011_firmato.pdf).
Ministro dell'Istruzione, dell'Università e della Ricerca (2013). Decreto Ministeriale 30 gennaio 2013 n. 47 «Autovalutazione, accreditamento iniziale e periodico delle sedi e dei corsi di studio e valutazione periodica» (disponibile al link: http://attiministeriali.miur.it/anno-2013/gennaio/dm-30012013.aspx).
Ministro dell'Istruzione, dell'Università e della Ricerca (2015). Decreto Ministeriale 27 giugno 2015 n. 458 «Linee guida valutazione qualità della ricerca (VQR) 2011-2014» (disponibile al link: https://www.anvur.it/wp-content/uploads/2015/07/dm_27_06_2015_vqr_11-14.pdf).
Ministro dell'Istruzione, dell'Università e della Ricerca (2016). Decreto Ministeriale 12 dicembre 2016 n. 987 «Autovalutazione, valutazione, accreditamento iniziale e periodico delle sedi e dei corsi di studio» (disponibile al link: http://attiministeriali.miur.it/anno-2016/dicembre/dm-12122016.aspx).
Ministro dell'Istruzione, dell'Università e della Ricerca (2019). Decreto Ministeriale 7 gennaio 2019 n. 6 «Autovalutazione, valutazione, accreditamento iniziale e periodico delle sedi e dei corsi di studio» (disponibile al link: https://www.miur.gov.it/web/guest/-/autovalutazione-valutazione-accreditamento-iniziale-e-periodico-delle-sedi-e-dei-corsi-di-studio).
Ministro dell'Istruzione, dell'Università e della Ricerca (2019). Decreto Ministeriale 29 novembre 2019 n. 1110 «Linee guida per la valutazione della qualità della ricerca (VQR) 2015-2019» (disponibile al link: https://www.anvur.it/wp-content/uploads/2019/12/DM-n.1110_29.11.2019.pdf).
Moher D., Naudet F., Cristea I.A., Miedema F., Ioannidis J.P.A., Goodman S.N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology*, *16* (3), e2004089.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Nucci M., Boscolo L., Sartori G. (2012). H-Index: valori e distribuzioni nella psicologia e psichiatria accademica italiana. *Giornale Italiano di Psicologia*, *39* (1), 7-19.
Parlamento della Repubblica Italiana (2010). Legge 30 dicembre 2010, n. 240, «Norme in materia di organizzazione delle università, di personale accademico e reclutamento, nonché delega al Governo per incentivare la qualità e l›efficienza del sistema universitario». *Gazzetta Ufficiale n. 10 del 14 gennaio 2011*, Suppl. Ordinario n. 11 (disponibile al link: https://www.camera.it/parlam/leggi/10240l.htm).
Rizzi D., Silvestri P. (2002). La valutazione del sistema universitario italiano: una storia recente. Nota di lavoro 2002.01, Dipartimento di Scienze Economiche, Università Ca' Foscari – Venezia.
Salvatore S., Pazzagli C., Lucarelli L., Gelo O., Giannone F. (2015). Livelli e fattori della produttività scientifica della psicologia accademica. Un'analisi bibliometrica. *Giornale Italiano di Psicologia*, *42* (1-2), 263-288.
United Kingdom Research and Innovation (2017). Initial decisions on the Research Excellence Framework 2021 (REF 2017/01) (disponibile al link: https://www.ref.ac.uk/media/1050/ref2017_01.pdf).

The evaluation of psychology higher education and research: Considerations on a turbulent decade

Summary. Almost ten years have elapsed since the start of evaluation activities in Italian academia. Some of these activities (ASN, VQR) have already accumulated the experience of more than one exercise, often with relevant repercussions for personal careers, educational programs and the distribution of resources to the structures that host training and research. In this period of time, academic psychology has been an active actor, sometimes showing internal cohesion and sharing, other times expressing the need for more articulation and differentiation, but never appearing unprepared towards the commitment represented by evaluation. The aim of this article is to understand, by looking at this relatively short past, whether among the many aspects of evaluation activities (in particular AVA, VQR, ASN and PRIN) some could be deemed fundamental, or should instead be changed to improve future practices and, more in particular, whether peculiarities of psychology evaluation are there that deserve special consideration in the dialogue with institutions and policymakers.

Keywords: Evaluation systems, evaluation, psychology, recruitment, quality assessment, tenure systems.

La corrispondenza va inviata a Luca Tommasi, Dipartimento di Scienze Psicologiche, della Salute e del Territorio, Università di Chieti-Pescara, Via dei Vestini 29, 66100 Chieti. E-mail: luca.tommasi@unich.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Estratti riservati ai soci AIP, © Società editrice il Mulino

GIORNALE ITALIANO DI PSICOLOGIA
3/4-2020, pp. 1047-94

Estratti riservati ai soci AIP, © Società editrice il Mulino

LO SFONDO SCONTATO DELLA VALUTAZIONE

CHIARA BERTI

Università di Chieti-Pescara «G. d'Annunzio»

Riassunto. L'articolo di Tommasi (2020) mostra come il tema della valutazione continui a essere oggetto di riflessione e discussione all'interno della psicologia accademica. Eppure questo dibattito sembra a volte troppo rivolto al particolare e distratto rispetto al disegno complessivo del quale le valutazioni fanno parte. Ad alcuni aspetti di questo disegno – quelli dati per scontati e che mi sembra non siano stati sufficientemente problematizzati da Tommasi – mi riferirò in questo articolo.

1. LA VALUTAZIONE: CIÒ CHE È MESSO A FUOCO

A dieci anni dall'entrata in vigore della legge n. 240/2010 con cui la valutazione diventa lo strumento al quale si consegna l'incentivazione della qualità e dell'efficienza del sistema, l'articolo di Tommasi (2020) mostra come il tema della valutazione continui a essere oggetto di discussione all'interno della psicologia accademica ed esprime una necessità sempre avvertita di riflettere sulle debolezze e sulle criticità che hanno segnato il processo avviato per correggerne o invertirne la rotta.

L'autore si pone domande di carattere generale – «se vi siano aspetti dei vari livelli della valutazione [...] che possano essere considerati irrinunciabili o che al contrario dovrebbero essere modificati» – e specifiche per la valutazione in psicologia, chiedendosi «se vi siano delle peculiarità [...] che debbano essere tenute in particolare considerazione nell'interlocuzione con le istituzioni di governo e politiche» (p. 13): un quesito, quasi un filo conduttore della sua riflessione sui turbolenti anni trascorsi, che rinvia al tema più generale dell'uniformità e della differenziazione che autorizzino o richiedano i saperi quando sono valutati nella loro qualità e rilevanza scientifica.

Sui diversi momenti della valutazione la psicologia italiana ha da tempo avviato una discussione. Ricorda Tommasi come la non scontata adesione all'approccio bibliometrico sia scaturita da posizioni di non facile conciliazione.

Sui limiti dell'impiego di strumenti di analisi bibliometrica nelle scienze umane e sociali, e la dovuta cautela nell'applicare in modo più

Estratti riservati ai soci AIP, © Società editrice il Mulino

o meno meccanico in quest'ambito sistemi di analisi già sperimentati in contesti diversi, sulla necessità di essere consapevoli delle pratiche scientifiche che caratterizzano la psicologia e i suoi diversi settori disciplinari, sulla funzionalità dei modelli di valutazione rispetto agli obiettivi, sulla non generalizzabilità e gli usi impropri dei risultati di un esercizio di valutazione, e sui cambiamenti che hanno interessato nel tempo i singoli momenti valutativi, la comunità degli psicologi italiani non ha smesso di riflettere, interrogarsi ed elaborare proposte (Cubelli, 2015, 2019).

Tra i problemi dell'attuale modello, Tommasi osserva «che l'attenzione dedicata dai meccanismi di valutazione ai corsi di studio potrebbe apparire sottodimensionata rispetto a quella dedicata alle strutture [...], alla ricerca e alla terza missione [...] e alle persone [...], a maggior ragione se si considera che il servizio offerto dalle università dovrebbe attribuire alla formazione un peso almeno analogo a quello attribuito alla ricerca», a fronte di una «valutazione della qualità dei dipartimenti realizzata attraverso il sistema AVA [...] sovrabbondante e probabilmente un po' sterile» (pp. 19-20).

Il dibattito sulla valutazione ha anche considerato i suoi effetti sulla qualità della vita quotidiana dei docenti e sull'organizzazione universitaria nel suo complesso. Afferma a questo proposito Tommasi che questa ha interessato in modo pervasivo la vita degli accademici, costituito una «pesante interferenza [...] nella vita quotidiana di ciascuno di noi» (p. 18) e rivoluzionato la mentalità e le prassi quotidiane degli accademici italiani.

Molto si è discusso in questo decennio sulla capacità conformativa di regole, di natura sostanzialmente normativa, che si esprime nella sottoposizione dei prodotti della ricerca a «criteri e parametri che cercano la propria oggettività in facili omogeneizzazioni e in altrettanto facili automatismi valutativi» (Barbati, 2019, p. 43), sui comportamenti adattivi ed elusivi (Rossi, 2018), sulla manipolazione dei dati citazionali e di impatto – un problema di primaria importanza nell'analisi bibliometrica – e sugli effetti distorsivi prodotti dalla valutazione sul suo stesso oggetto, ossia la ricerca scientifica[1]. Per Barbati (2012)

[1] Privilegiare la ricerca *mainstream* assicura una buona «resa» bibliometrica, ma al contempo la impoverisce e la irrigidisce caratterizzandola per un comportamento inerziale. La ricerca *agenda driven*, a differenza di quella *curiosity driven*, produce i suoi effetti in tempi solitamente più rapidi, a scapito però dell'innovatività (Banfi, 2012). Per quanto riguarda gli impatti conformativi, è stata evocata la legge Goodhart: nata in ambito economico, quando applicata all'ambito scientometrico, stabilisce che ogni volta che un parametro quantitativo viene elevato a indicatore normativo discriminante di qualità, efficienza o performance, il parametro stesso smette di essere un valido strumento di misura perché gli agenti intelligenti tendono a sviluppare strategie per soddisfare il criterio con modalità che vanno a detrimento proprio del livello di qualità che si intendeva certificare con tale parametro. In questo senso le uniche tornate di

Estratti riservati ai soci AIP, © Società editrice il Mulino

gli esempi delle distorsioni dell'oggetto della ricerca sono espressi in particolare dalle discipline appartenenti alle *Human and Social Science*, distanti più di altre dagli strumenti di valutazione quantitativa e che offrono perciò il banco di prova della sovrapposizione e confusione tra valutazione e misurazione. Non sono più i criteri e tantomeno i parametri a doversi adeguare ai saperi ma sono i saperi a doversi adeguare a essi. S'inverte dunque il rapporto tra la regola e il suo oggetto. Lo stesso Consiglio Universitario Nazionale, nelle sue analisi e pronunce, si è più volte espresso in riferimento ai caratteri assunti dalla valutazione esterna, al suo carico regolativo, agli impatti, anche conformativi, da essa prodotti sui diversi ambiti interessati e per i diversi fini che ne orientano l'esercizio.

2. DARE PER SCONTATO

Mi chiedo se, a fronte della pervasività e dell'impatto delle valutazioni nella vita dei singoli e delle organizzazioni, il rapporto con il sistema della valutazione non si sia sviluppato, almeno per alcuni e in relazione ad alcuni aspetti, in maniera scontata. L'attenzione appare essere stata troppo rivolta al particolare e distratta rispetto al disegno complessivo del quale ogni esercizio valutativo parte. Ad alcuni aspetti di questo disegno – quelli che più raramente in questi anni mi sembra siano stati problematizzati, andando così a costituire una struttura tacita e implicita – mi riferirò nel terzo paragrafo.

In un loro volume Legrenzi e Umiltà (2018) usano spesso la locuzione «dare per scontato» per indicare quei processi mentali di cui sfugge il funzionamento in quanto non se ne è consapevoli. Anche settori della sociologia adottano oggi approcci individualisti e cognitivi. Di recente si è sviluppata una corrente cognitivista della sociologia focalizzata sullo studio dell'*unmarked*, del *taken for granted* proprio come sfondo a ciò che viene messo a fuoco, illuminato dall'interesse e dal significato. Zerubavel (2019) usa il modello figura-sfondo della Gestalt per spiegare che quando si presta attenzione a qualcosa necessariamente molto altro rimane sullo sfondo. Prestare attenzione, mettere a fuoco qualcosa è sempre un processo selettivo che implica lasciare in ombra altro. Lo scontato non è solo ciò che non è sotto l'attenzione di tutti, riassume anche ciò che tutti sanno e quindi è banale e ovvio. Osserva inoltre Zerubavel che il dato per scontato di oggi può essere lo straordinario di domani.

ASN veramente significative dal punto di vista dei PI (Indicatori di Performance) potrebbero state le prime due, quando nessuno aveva avuto ancora il tempo di organizzarsi per fare un *gaming* dei parametri sufficiente a distorcerli e viziarli.

Estratti riservati ai soci AIP, © Società editrice il Mulino

L'aspetto che non viene messo in luce è che «dare per scontato» rimanda ad una credenza condivisa da tutti e non messa in discussione, una credenza fondante un certo modo di stare al mondo e, fino a prova contraria, non discutibile. La conoscenza può diventare scontata solo se lo è per tutti coloro che fanno parte dello stesso gruppo sociale. Notiamo o ignoriamo non solo in virtù delle capacità individuali di prestare attenzione, ma anche in quanto esseri sociali, come parti di una collettività che in base alle sue norme e ai suoi riferimenti culturali dirige i suoi processi attentivi.

Quanto più scontata, implicita e anonima è la conoscenza, tanto più acquista il carattere di concretezza che la contraddistingue. La funzione principale della condivisione sociale sta nella costruzione partecipata e in larga misura implicita della realtà in cui viviamo. La possibilità di dare per scontato e di rendere implicite e automatiche le conoscenze e i gesti necessari per l'esecuzione di gran parte delle azioni quotidiane è il requisito base per avviare un processo di oggettivazione, mediante il quale si sedimenta la parte più stabile delle «interpretazioni condivise» della realtà, quella che più difficilmente si problematizza e va a costituire una struttura «tacita e pertanto inconsapevole» valida per tutti. Soltanto quando diventa scontata per tutti, essa assume il carattere di oggettività «quasi naturale», reale al di fuori di noi; questo è presumibilmente l'aspetto cruciale della condizione umana: percepire e trattare con il carattere di realtà «naturale» una realtà costruita socialmente (Emiliani, 2008).

3. IL TACITO E L'IMPLICITO DELLA VALUTAZIONE

a) Il significato di *qualità tra rimandi, mutuazioni e nozioni di senso comune, e laconicità del legislatore.* Benché i concetti di qualità ed efficienza rinviino a modelli aziendali di organizzazione e funzionamento di soggetti e sistemi anche pubblici, essi possono essere riempiti di contenuti differenti a seconda dell'ambito interessato e degli obiettivi che s'intendono conseguire. Tommasi parla di una «rappresentazione intuitiva di ciò che tale concetto possa significare in ambito accademico» (p. 14) e legata al buon senso, anche tra coloro che hanno avviato e gestito il Processo di Bologna, e ricorda che «nessuna definizione di qualità o, eventualmente, dei requisiti ai quali essa potrebbe corrispondere, è mai stata proposta» (p. 15).

b) *La valutazione esterna come strumento di governo del sistema.* La «grande riforma» avviata nel 2010 ha cambiato il «chi» decide e insieme il «quanto», il «cosa» e il «come» decide, sia presso l'amministrazione statale del settore sia presso le Istituzioni universitarie. L'intervento del 2010 si connota non solo «per avere sottratto molte

Estratti riservati ai soci AIP, © Società editrice il Mulino

scelte al loro potere di autodeterminazione, ma anche e soprattutto per avere previsto un altro «centro» di regolazione. Con la legge n. 240/2010 entra, infatti, in scena la nuova «valutazione» esterna. Una sorta di «misura delle misure» con la quale si confrontano perciò tutti i soggetti e tutte le dimensioni dell'autonomia universitaria» (Barbati, 2019, pp. 11-12).

c) Una questione di *giustizia procedurale: la rappresentatività del soggetto che fissa le regole e la contrazione dell'autonomia.* La scelta del legislatore del 2010[2] avrebbe dovuto accrescere l'attenzione delle comunità per la legalità procedurale e per i processi decisionali. Tra le varie dimensioni della legalità procedurale, quella della rappresentatività si pone quando il soggetto regolatore non sia direttamente espressione dei soggetti interessati. La necessità di regole che siano correttamente percepite dai destinatari e dunque garantite nella loro applicabilità ed effettività suggerisce, in questi casi, il loro coinvolgimento. Ciò è tanto più necessario quando le regole che si tratta di definire possiedono un'elevata capacità di conformare scelte e comportamenti e non poggiano su esperienze né su pratiche consolidate, condivise o anche solo riconosciute (Barbati, 2012). Al contrario, osserva sempre Barbati (2019), la valutazione «cerca la propria forza nel farsi distante dalle esperienze delle comunità scientifiche quasi trasmutandole in interessi di parte, sino a sottrarre alla loro autodeterminazione lo stesso riconoscimento del presupposto della valutazione, ossia della scientificità delle pubblicazioni» (pp. 45-46). È invece proprio la distanza dai plurimi interessi espressi dalle tante autonomie universitarie che potrebbe diventare lo strumento che conduce a favorire scelte particolari, esito del dialogo che «chi decide» intrattiene solo con alcuni interessi, in particolare con quelli che più riescono ad avvicinarsi al decisore, ovvero al soggetto della valutazione esterna (Barbati, 2015).

Il Consiglio Universitario Nazionale (CUN) è l'organo elettivo di rappresentanza del sistema universitario e al suo interno le comunità scientifiche e disciplinari trovano l'unico luogo istituzionale, consultivo e propositivo, che consenta loro di parlare «con chi decide». L'assenza sia di un riordino strutturale, capace di fare del CUN la sede propria delle comunità scientifiche, sia di un riordino funzionale ne ha limitato «la possibilità di farsi interprete delle esigenze di cui le comunità sono portatrici, a fronte di un apparato regolatorio che

[2] È in realtà il D.P.R. 14 settembre 2011, n. 222 con cui si disciplinano le modalità di accesso alla prima e seconda fascia dei professori universitari a ridefinire il potere di autodeterminazione delle comunità scientifiche e lo stesso statuto della loro autonomia. L'art. 6, comma 5, colmando il silenzio del legislatore in merito a chi e come accertare la qualificazione degli aspiranti Commissari, sottrae alle comunità scientifiche il potere di scegliere chi sceglie e lo assegna all'ANVUR che con esse non intrattiene alcun rapporto organico o di rappresentanza (Barbati, 2019).

Estratti riservati ai soci AIP, © Società editrice il Mulino

interviene sui compiti e sulle funzioni alle quali sono chiamate, sulle modalità in cui devono e possono essere assolte, e ne depotenzia perciò anche la capacità di partecipare ai processi decisionali per il governo del sistema, consegnandola all'iniziativa spontanea ovvero "dal basso" dello stesso organo, lasciato "solo" a cercare o a reinterpretare un possibile ruolo» (Barbati, 2014, p. 345).

d) *Questioni di giustizia distributiva*. L'esito della misurazione non ha conseguenze automatiche e obbligatorie (Cubelli, 2015). I risultati della valutazione potrebbero servire per distribuire fondi e quote premiali privilegiando le aree più forti all'interno di un processo di sviluppo selettivo e specialistico, oppure potrebbero servire per potenziare le aree più deboli, in un contesto di offerta generalista ed equilibrata svolgendo una funzione incentivante, di miglioramento. A proposito delle logiche premiali che danneggiano le realtà più deboli, togliendo loro di fatto ogni possibilità di recupero, si è parlato di effetto S. Matteo – 25, 29: «Perché a chiunque ha sarà dato e sarà nell'abbondanza; ma a chi non ha sarà tolto anche quello che ha».

La consapevolezza auspicata dei problemi della valutazione e della scientometria (Banfi, 2012) da sola non basta. La riflessione sulla valutazione non può infatti limitarsi agli aspetti tecnici, ma essere attenta alle implicazioni delle scelte effettuate e prima ancora agli obiettivi che s'intendono conseguire e ai quali tendere tramite la qualità e l'efficienza attraverso i processi di riforma avviati nel 2010. Il legislatore, lungi dall'esplicitare questi obiettivi, «li ha consegnati alle diverse misure previste e all'interpretazione che di esse ne hanno dato i tanti attori chiamate ad applicarle, fra ripensamenti e rafforzamenti di un disegno che investe innanzi tutto la concezione della funzione assolta dall'Università» (Barbati, 2019, p. XI) e il ruolo da riconoscerle oggi al di là dell'identificazione dei compiti istituzionali che storicamente la connotano.

BIBLIOGRAFIA

Banfi A. (2012). Apples and Oranges. Spunti per una discussione sulla valutazione della ricerca nelle scienze umane e sociali. *Astrid Rassegna*, *12*, http://www.astrid-online.it/rassegna/2012/28062012.html.

Barbati C. (2014). Il governo del sistema universitario: soggetti in cerca di un ruolo. *Trimestrale di Diritto Pubblico*, *2*, 337-360.

Barbati C. (2015). Alla ricerca dei criteri di scientificità delle pubblicazioni. In G. Conte (a cura di), *Evoluzione e valutazione della ricerca giuridica*. Napoli: ESI.

Barbati C. (2019). *Il sistema delle autonomie universitarie*. Torino: Giappichelli.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Consiglio Universitario Nazionale (2019). Il sistema universitario nelle analisi del Consiglio Universitario Nazionale 2017-2019, https://www.cun.it/uploads/7019/Cunbiennio20172019.pdf?v=.
Cubelli R. (2015). La psicologia italiana tra scienza della valutazione e valutazione della scienza. *Giornale Italiano di Psicologia*, *42* (4), 647-675.
Cubelli R. (2019). Editoriale. *Giornale Italiano di Psicologia*, *46* (3), 399-408.
Emiliani F. (2008). *La realtà delle piccole cose. Psicologia del quotidiano.* Bologna: Il Mulino.
Legrenzi P., Umiltà C. (2018). *Molti inconsci per un cervello. Perché crediamo di sapere quello che non sappiamo*. Bologna: Il Mulino.
Rossi P. (2018). Esercizi di valutazione. Un'analisi di impatto dei criteri. *Quaderni di «Articolo 33»*, *3*, 63-74.
Tommasi L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.
Zerubavel E. (2019). *Dato per scontato. La costruzione sociale dell'ovvietà.* Milano: Meltemi.

The unmarked ground of evaluation

Summary. Tommasi's paper (2020) shows how evaluation continues to be an issue of debate for the psychological academic community. Yet, this debate sometimes appears too focused on particular aspects and doesn't pay due attention to the overall design of which the evaluations are part. This paper addresses some processes underlying evaluation which seem to be taken for granted and have not been sufficiently emphasized in Tommasi's article.

Keywords: Evaluation, psychology, policies, taken for granted, procedural justice.

La corrispondenza va inviata a Chiara Berti, Dipartimento di Scienze Psicologiche, della Salute e del Territorio, Università di Chieti-Pescara «G. D'Annunzio», Via dei Vestini 31, 66100 Chieti. E-mail: chiara.berti@unich.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Estratti riservati ai soci AIP, © Società editrice il Mulino

VALUTARE LA QUALITÀ DI PERSONE, STRUTTURE, DIDATTICA, RICERCA... ALLA RICERCA DEL KIT DI SOPRAVVIVENZA

SANTO DI NUOVO

Università di Catania

Riassunto. La valutazione della qualità nel mondo universitario assume aspetti e valenze diverse, che riguardano il contenuto di ciò che si valuta (ad esempio, i prodotti scientifici o la didattica o l'impegno organizzativo) e la finalità: abilitazione scientifica individuale, VQR che riguarda le strutture, richieste di finanziamenti. Vengono evidenziate alcune criticità e l'esigenza di risolverle nell'interesse delle giovani generazioni di studiosi e docenti.

L'articolo di Tommasi (2020) che ha aperto la discussione promossa dall'Associazione Italiana di Psicologia sulla valutazione della psicologia, riporta considerazioni e commenti molto condivisibili. Sono convinto che, evitando la sensazione di fastidio e la tentazione di rinunciare ad una missione che molti ritengono impossibile, «valga la pena sacrificare un po' di tempo come comunità scientifica per metterci d'accordo su che cosa funzioni e che cosa non funzioni della valutazione, come possa essere definita e misurata la qualità, che cosa dovrebbe caratterizzare in modo irrinunciabile l'identità scientifica della psicologia agli occhi dei sistemi di valutazione» (Tommasi, 2020, p. 18).

Il problema, come è stato giustamente sottolineato, è che la *qualità* resta difficile da definire in generale perché si declina in tanti aspetti non omogenei e può essere vista da prospettive diverse.

Una cosa è valutare la qualità della didattica e della formazione erogata da un corso di studi, per l'accreditamento e l'autovalutazione delle strutture, un'altra la qualità di un progetto che chiede di essere finanziato; i criteri per valutare la carriera e l'impegno scientifico di una persona non sono gli stessi che servono per valutarne la capacità e l'impegno nel trasmettere la propria scienza agli studenti o divulgarla e/o applicarla al territorio. I criteri delineati per ciascuna di queste tipologie di valutazione non sono – e forse non possono essere – sempre congruenti tra loro e questo finisce col disorientare i «soggetti» *della* (forse bisognerebbe dire *alla*) valutazione: che in realtà sono sempre gli stessi, vale a dire i docenti-ricercatori. Infatti

Estratti riservati ai soci AIP, © Società editrice il Mulino

è attraverso le valutazioni dei singoli che passano anche le valutazioni delle strutture e dei progetti collettivi. Riguardo a questi ultimi, abbiamo visto come per i finanziamenti PRIN la valutazione del curriculum del *principal investigator* è stata addirittura preliminare, e preclusiva, rispetto alla considerazione di ciò che nel progetto c'era scritto. Al contrario, l'impegno e l'abnegazione di un presidente di corso di studi per migliorare l'autovalutazione ai fini dell'accreditamento possono essere vanificati dalla ridotta collaborazione dei docenti del corso che preferiscono impegnarsi nella produzione scientifica o nelle visite all'estero, indispensabili per la loro valutazione personale: in questo caso la valutazione del responsabile conta poco ai fini della valutazione complessiva del corso.

Quando si parla di qualità dei «prodotti» (vale a dire, quello che un ricercatore produce nella sua giornata lavorativa) si parla della ricerca pubblicata in articoli ad alto impatto che possono essere considerati «eccellenti» (utili per la VQR), ma anche dei manuali che sono adatti per consentire una buona didattica e la soddisfazione degli studenti (AVA) o per una ampia e corretta divulgazione, o dei report delle ricerche-intervento, utili per la «terza missione» – adesso pure valutata nella VQR – e per portare utili e prestigio sociale al proprio dipartimento. Ma i prodotti non «indicizzati» sono inutili per l'ASN: la quale invece richiede, ai fini del raggiungimento della soglia quantitativa, *tante* pubblicazioni su riviste indicizzate (non *una «eccellente» all'anno* come richiesto dalla VQR) e che siano capaci di attirare *tante citazioni*, visto che due soglie su tre derivano proprio dalle citazioni ottenute. Cosa c'entra la qualità di quello che si scrive con l'essere tanto citati, e col modo in cui le citazioni si ottengono, resta dubbio: ma tanto si è detto su questo e non voglio tornarci in questa sede. Le teorie sull'adattamento ci dicono che bisogna imparare ad adeguarsi a quello che non si può cambiare, per cui se ci diranno che i prodotti di qualità sono quelli scritti in esperanto, tutti impareremo a scrivere in questa lingua e pure le riviste *open access* si adegueranno.

Però a qualche aspetto di troppo palese discrepanza tra norme, prassi e logica nel mettere insieme valutazioni diverse non riusciamo ad abituarci, per cui alcune proposte andrebbero fatte.

Prendo ad esempio il palese conflitto tra criteri della abilitazione scientifica nazionale e quelli della successiva chiamata dell'abilitato nella sede, dove si deve tenere conto anche della didattica.

Sappiamo che l'assunzione negli Atenei – diversamente che negli Enti di ricerca – avviene come docenti. Solo alcuni fra i tanti abilitati in base a criteri *scientifici* verranno chiamati ad entrare nel ruolo di *docenti*, o ad assumere funzioni docenti superiori a quelle precedenti. Ma il problema non è solo quello quantitativo da Tommasi bene evidenziato, di una eccedenza di abilitati che produce «troppa disponi-

Estratti riservati ai soci AIP, © Società editrice il Mulino

bilità qualificata», per cui tanti nominati *cavalieri* (della scienza) non potranno godere del *cavallo* (il posto di docente). La discrepanza tra i due momenti valutativi riguarda anche la pertinenza del settore in cui ci si abilita (SC) con quello in cui si viene chiamati poi ad insegnare (SSD). Un abilitato nel Settore Concorsuale può essere chiamato ad insegnare in qualunque dei Settori Disciplinari che lo compongono, pur essendo scientificamente specialista di un altro dei SSD che compongono il SC. È paradossale che al momento della ratifica della domanda – nei «semafori» preliminari alla valutazione vera e propria – la commissione possa dire al candidato che la sua candidatura non è adatta per un SSD tra quelli del SC, ma non può più farlo alla fine della valutazione: per cui quel candidato, una volta abilitato, potrà essere chiamato da un ateneo compiacente (o insipiente) proprio nel SSD per il quale si è accertato che non è adatto. Perché la commissione non potrebbe concludere che un candidato è abilitato – cioè scientificamente compatibile – esclusivamente in uno dei SSD, per cui potrà insegnare clinica ma non dinamica, sociale ma non organizzazione, psicobiologia ma non metodologia? Si eviterebbero così a monte rischi di incongruenza che non giovano certo alla qualità della didattica. Certo questo rischio potrebbe essere evitato dalle commissioni che poi valutano localmente, ma i fatti ci dicono che la compiacenza/insipienza di cui sopra non è affatto assente nel panorama della psicologia (e non solo).

Questa che ho citato è solo una delle tante incongruenze tra due tipi di valutazione nelle diverse fasi di giudizio di una stessa persona – tanto evidente che alcune proposte di legge vorrebbero superarla parlando di abilitazione scientifica *e didattica* – ma non certo l'unica.

In conclusione, qualche riflessione in forma forzatamente semplificata e volutamente provocatoria.

Abbiamo detto che tutto il sistema valutativo passa attraverso la valutazione di persone, anche quando poi queste valutazioni si considerano in modo aggregato: atenei, dipartimenti, corsi di studio, dottorati, gruppi di ricerca; e che sono le stesse persone che producono le pubblicazioni, insegnano (le prime due missioni dell'Università), promuovono e attuano la terza missione, chiedono i finanziamenti. Quindi, di là della discussione generale e di «sistema», vorrei considerare il punto di vista di un singolo docente/ricercatore (o aspirante tale), magari di un giovane che deve costruire la sua carriera oltre che contribuire alla struttura in cui si trova. Mettendomi nei suoi panni, vorrei che i criteri sui quali essere valutato fossero tali da non farmi procedere schizofrenicamente nella valorizzazione della mia carriera scientifica e didattica, come un novello Jekyll-Hyde costretto ad alternare maschere diverse nel corso della sua vita lavorativa.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Alcuni esempi di «dilemmi del ricercatore»:

1) *Produzione scientifica*:
a) pubblicare *pochi lavori all'anno ma di eccellenza* (utili per la valutazione della VQR e quindi al dipartimento che mi sostiene), oppure
b) *tanti e che attraggano tante citazioni* (per le mie soglie ASN)?
E a proposito di citazioni, scegliere *temi di mainstream e ampie collaborazioni* che li favoriscono, o *argomenti di nicchia* scientificamente più originali e innovativi – quindi di più alta qualità, ai fini della soddisfazione personale e anche della VQR – che però isolano dal contesto, rendono più difficile accedere alle riviste che privilegiano il *mainstream,* e diminuiscono la potenzialità di essere citati?

2) *Orientamento disciplinare*:
a) *focalizzato e specialistico*: facilita il raggiungimento sia dell'eccellenza che delle soglie, ma sul piano didattico è inadeguato per la copertura di un ampio *range* di aspetti del settore di insegnamento, oppure
b) *ampio e interdisciplinare*: utile per le collaborazioni e i finanziamenti di ricerca specie se internazionali, ma col rischio di essere valutato «eccellente ma non pertinente» per uno specifico settore nell'ASN.

3) *Impegno didattico e organizzativo nel dipartimento*:
a) assumerne lo stretto indispensabile, quindi il meno possibile, così da avere più tempo per la ricerca (dato che le giornate lavorative non sono elastiche...) e da godere ampi periodi liberi per fare esperienza all'estero, che viene valutata nell'ASN, oppure
b) contribuire, con la più ampia disponibilità che le forze consentono, alla crescita del proprio dipartimento (ai fini AVA), guadagnarsene la fiducia, e quindi la priorità ai fini di una possibile progressione di carriera?

Come giovane ricercatore, sarei davvero molto perplesso e confuso tra queste alternative.
Come presidente di una Associazione che ha metà dei soci giovani e in posizione precaria, e tre quarti nella posizione di valutato o di valutatore, vorrei che fosse data qualche risposta a queste perplessità in modo organico e condiviso. La direzione dovrebbe essere quella di una maggiore uniformità dei criteri, su cui ognuno conosca – in anticipo, e non in un «a posteriori» continuamente mutevole – *perché* viene valutato e *come*.
Le modifiche ai sistemi di valutazione da elaborare, concordare e proporre, nelle sedi opportune e in collaborazione con altre associa-

Estratti riservati ai soci AIP, © Società editrice il Mulino

zioni disciplinari del variegato mondo accademico, dovrebbero sostenere la fiducia nelle istituzioni accademiche da parte di chi ci trova già o e di chi vuole accedervi. E soprattutto dovrebbero accendere, e mantenere accesa, la speranza per un futuro personale e collettivo di ricerca e insegnamento di qualità nella propria disciplina.

BIBLIOGRAFIA

TOMMASI L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

Evaluating the quality of persons, structures, teaching, research ... in search of a survival kit

Summary. The quality assessment in the academic settings requires different aspects and criteria, concerning the content of what is evaluated (e.g., scientific products, or teaching, or organizational commitment) and the aim: individual scientific qualification, evaluation of the structures, requests for funding. Some critical issues are highlighted, and the need to cope with them in the interest of the younger generations of scholars and teachers.

Keywords: Evaluation, research, teaching, quality assessment, psychology.

La corrispondenza va inviata a Santo Di Nuovo, Dipartimento di Scienze della Formazione, Università di Catania, Via Biblioteca 4, Palazzo Ingrassia, 95131 Catania. E-mail: s.dinuovo@unict.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Estratti riservati ai soci AIP, © Società editrice il Mulino

PER UNA VALUTAZIONE DELLA DIDATTICA OLTRE I PROCESSI

RAFFAELLA IDA RUMIATI

SISSA, Trieste

Riassunto. Nel mio commento spiegherò perché nell'articolo di Tommasi (2020), altrimenti molto stimolante e accurato, la distribuzione delle responsabilità tra le diverse istituzioni coinvolte nei processi di valutazione non sia sempre condivisibile. In secondo luogo, proporrò come utilizzare proficuamente lo strumento dei test delle competenze trasversali e disciplinari anche in psicologia.

L'articolo di Tommasi (2020) propone un'analisi molto approfondita dei vari esercizi di valutazione che hanno come oggetto le diverse attività di docenti e università, con particolare riferimento all'ambito della psicologia. Tommasi offre una lettura della valutazione universitaria che attribuisce all'ANVUR compiti e responsabilità di protagonista quasi assoluto, sottovalutando il ruolo giocato da altre istituzioni, in primo luogo il Ministero dell'Università e della Ricerca (MUR), e i fattori di contesto che inevitabilmente influenzano i rapporti tra esse. Le azioni valutative della qualità della ricerca (VQR), della terza missione e della didattica (AVA), come pure l'identificazione dei criteri per l'erogazione dei finanziamenti ai Dipartimenti di Eccellenza e del FABBR, in linea teorica, fanno capo all'ANVUR. Quest'ultima è un'Agenzia tecnica indipendente dal MUR ma da quest'ultimo vigilata e, pur nel rispetto formale dei rispettivi ruoli, le due istituzioni cercano di mettere in atto dei processi condivisi. Inoltre, l'attuazione di molti di questi esercizi si avvale della partecipazione attiva dell'accademia e non è imposta dall'alto. Infatti, come ricordato anche da Tommasi, i GEV della VQR e le CEV di AVA sono costituiti da docenti che godono di un ampio grado di autonomia. La ASN, invece, è di pertinenza esclusiva del MUR, mentre l'ANVUR svolge il compito meramente tecnico di calcolare le soglie dei SSD, applicando le disposizioni ministeriali. PRIN e FIRB sono stati storicamente gestiti dal MUR e, a partire dalla sua istituzione, prevista dalla Legge 240/2010 (Legge Gelmini), dal CNGR mentre l'ANVUR dovrebbe valutare gli esiti *ex post* dei progetti ma non mi risulta che l'Agenzia abbia mai ricevuto dal MUR il mandato per farlo. Così come è ancora di là da

Estratti riservati ai soci AIP, © Società editrice il Mulino

venire l'integrazione dell'accreditamento dei corsi di studio (AVA) con quello dei dottorati, che è regolato da un decreto diverso, o la valutazione *ex post* di questi ultimi. Del resto, il più che frequente avvicendamento dei Ministri di Viale Trastevere degli ultimi anni ha comportato cambiamenti bruschi quando non incoerenti, nell'agenda della valutazione.

Quindi, data la divisione di competenze e ruoli delle varie istituzioni, ritenere che le politiche valutative siano interamente responsabilità dell'ANVUR, non solo non riflette la realtà ma finisce, da un lato, con l'affievolire quel giusto senso critico, che invece andrebbe sempre esercitato nei riguardi delle politiche di un Governo, veicolate a mezzo del Ministero competente, dall'altro, con il sollevare la comunità accademica (i docenti) e la *governance* degli Atenei dai propri oneri. Un'analisi simile è stata proposta da Capano, Regini e Turri (2016) i quali, oltre a mettere in evidenza le responsabilità proprie della politica in materia di università e ricerca, affermano che anche i docenti e la *governance* delle Università non sono sempre stati in grado di interpretare al meglio quello che di buono c'è nelle leggi, come è avvenuto nel caso dell'applicazione dell'autonomia universitaria, prevista dalla Legge Gelmini. Taluni Direttori di Dipartimento hanno impropriamente utilizzato degli indicatori bibliometrici – messi a punto dall'ANVUR per tutt'altro scopo (VQR) e a livello dipartimentale – per assegnare a un singolo docente assegni di ricerca o altre forme di finanziamento locale. Anche di queste scelte improvvide si è ritenuta spesso responsabile l'Agenzia.

L'articolo di Luca Tommasi lascia intendere che la tendenza dell'accademia italiana ad applicare un *lucus of control* esterno, in particolare quando si affrontano i nodi relativi alla valutazione, non ha risparmiato nemmeno la psicologia.

Non essendoci spazio per discutere tutti gli istituti di valutazione passati in rassegna nell'articolo-bersaglio, nel mio intervento mi soffermerò esclusivamente sulla valutazione della didattica. Normativamente essa è inclusa in AVA, uno strumento finora utilizzato principalmente per monitorare i processi di assicurazione della qualità dell'offerta formativa, che per quanto riguarda la verifica degli esiti di tale offerta è ancora nella sua infanzia. Nel quadro delle attività di valutazione della didattica previste in AVA, il DPR 76/2010 di istituzione dell'ANVUR e i successivi decreti (Allegato E del DM 987/2016 in seguito superato dal DM 6/2019) emanati dal MIUR, hanno conferito all'ANVUR il mandato di sviluppare indicatori relativi agli esiti degli apprendimenti e ai tassi di occupazione degli studenti. Questi interventi normativi sono in linea con le raccomandazioni dell'Unione Europea che promuovono le otto competenze-chiave per la cittadinanza attiva e l'apprendimento permanente, annunciate per la prima volta nel 2006,

Estratti riservati ai soci AIP, © Società editrice il Mulino

dove le competenze vengono definite come la «combinazione di conoscenze, abilità e attitudini adeguate per affrontare una situazione particolare». Tali norme hanno quindi fornito gli strumenti necessari per favorire una progressiva adesione del sistema universitario italiano a quello degli altri paesi dell'*European Higher Education Area* (EHEA), orientata verso un apprendimento sempre più centrato sulle competenze (si vedano le *European Standards and Guidelines for Quality Assurance* – ESG, Yerevan 2015).

È all'interno di questo quadro che, nel 2012, l'ANVUR ha avviato il progetto denominato TEst delle Competenze (TECO), al fine di valutare gli esiti degli apprendimenti degli studenti universitari, avvalendosi del CLA+, prodotto dal *Council for Aid to Education* (CAE). Le prove contenute nel CLA+ mirano a valutare l'analisi e la soluzione di problemi, il ragionamento scientifico e quantitativo, la lettura critica e la valutazione critica di una argomentazione, e la scrittura. Le due somministrazioni del CLA+ (nel 2013 e nel 2015) in varie università del nostro Paese hanno mostrato vari limiti, tra cui la scarsa partecipazione degli studenti, la bassa correlazione tra risposte aperte e chiuse, disaccordo tra *scorers*, l'autoselezione degli studenti, reclutati in base al superamento di un numero elevato di CFU, e l'elevato costo dell'intera procedura. L'ANVUR aveva avviato anche un tavolo di esperti, portatori di interesse delle varie macro-aree accademiche, cui era stato affidato il mandato di realizzare delle prove disciplinari. Tuttavia, non tutti gli esperti coinvolti hanno avuto l'opportunità di condividere il progetto con le proprie comunità accademiche, e le sperimentazioni effettivamente portate a termine hanno interessato solo le scienze dure e della vita.

Per superare i limiti delle prime esperienze TECO, dal 2016 l'Agenzia ha lanciato un nuovo progetto, organizzato in due filoni, uno per le competenze trasversali (TECO-T) e uno per le competenze disciplinari (TECO-D). La prima innovazione del nuovo TECO è metodologica e consiste nell'adozione del *valore aggiunto*, un approccio grazie al quale vengono rilevate le competenze sia degli studenti del primo anno della laurea triennale, sia di quelli del terzo anno, per quantificare il ruolo dell'Università nel loro sviluppo e/o mantenimento. Il secondo cambiamento riguarda la realizzazione delle prove. Quelle trasversali per gli ambiti di *literacy*, *numeracy*, *problem solving* e *civics* sono state costruite dall'Agenzia con la collaborazione di vari docenti, e sono state somministrate al computer nel corso di successivi esercizi, in aule informatiche delle università che hanno partecipato al progetto, utilizzando una piattaforma CINECA. L'ANVUR ha anche promosso e coadiuvato la realizzazione di TECO disciplinari le cui prove sono state interamente realizzate da gruppi di lavoro costituiti da docenti esperti, con la supervisione dell'Agenzia. I gruppi di lavoro delle professioni

Estratti riservati ai soci AIP, © Società editrice il Mulino

sanitarie sono stati tra i primi ad adottare il TECO disciplinare, avendolo ricavato dal *progress test* utilizzato per anni come strumento di autovalutazione. I loro test sono stati somministrati già a più coorti successive di studenti, evidenziando un loro significativo apprendimento disciplinare nel corso del triennio, nonché un ruolo dell'Università nel mantenimento e potenziamento delle competenze trasversali (cfr. Rumiati *et al.*, 2019). Anche il progress test di medicina è stato adattato al TECO, mentre le prove realizzate dai gruppi di lavoro di filosofia e pedagogia sono state standardizzate alla fine del 2019. I gruppi di lavoro di psicologia hanno appena completato le proprie prove (si veda il commento di Lucidi), mentre le prove di lettere, giurisprudenza e veterinaria sono ancora in via di definizione.

I TECO si prestano a diversi scopi. Il TECO trasversale potrebbe, per esempio, essere utilizzato come test di ingresso a psicologia. Questa scelta comporterebbe alcuni vantaggi: gli aspiranti studenti di tutti i CdS verrebbero sottoposti a un unico test, utilizzando domande equivalenti pescate da un ampio database, e gli Atenei risparmierebbero il denaro che ora viene speso per l'acquisto di test di ingresso (o di orientamento), spesso diversi da Ateneo ad Ateneo, che non sempre valutano adeguatamente le competenze trasversali. Il TECO disciplinare di psicologia offre informazioni utilissime al fine dell'autovalutazione, della progettazione e della revisione dei CdS. Somministrato agli studenti al terzo anno della laurea di primo livello potrebbe permettere di verificare le competenze teorico-professionali necessarie per l'eventuale inserimento nel mondo del lavoro (una scelta divenuta poco attraente stando all'esiguo numero di iscritti alla sezione B dell'Albo degli Psicologi), o per l'accesso ai corsi di Laurea Magistrale di Psicologia. Infine, il TECO disciplinare potrebbe contribuire alla costruzione di un percorso che porti alla laurea abilitante, come è stato recentemente realizzato dalla medicina.

In conclusione, sebbene in futuro si possa certamente migliorare, non vi è dubbio che l'università è il settore della pubblica amministrazione che nel nostro Paese si è distinto maggiormente per la capacità valutare se stesso. Nell'ambito della valutazione degli esiti degli apprendimenti, i TECO potrebbero servire a più scopi, come illustrato sopra, portando al centro della didattica le competenze necessarie per la formazione in ambito psicologico.

BIBLIOGRAFIA

Capano G., Regini M., Turri M. (2016). *Changing governance in Universities. Italian higher education in comparative perspective*. London: Palgrave Macmillan.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Rumiati R.I., Ciolfi A., Di Benedetto A., Sabella M., Infurna M.R., Ancaiani A., Checchi D. (2019). Key-competences in higher education as a tool for democracy. *Form@re*, *3*, 7-18.
Tommasi L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

For an evaluation of teaching beyond the processes

Summary. In my comment I will explain why I do not always share the way in which the responsibilities in the evaluation processes are attributed to the different institutions in the Tommasi's (2020) paper, otherwise very inspiring and accurate. Second, I will propose how to profitably use the tests of the transversal and disciplinary competences also in psychology.

Keywords: Learning outcomes, competences, quality assurance, evaluation system, psychology.

La corrispondenza va inviata a Raffaella Ida Rumiati, SISSA – Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, 34136 Trieste. E-mail: rumiati@sissa.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Estratti riservati ai soci AIP, © Società editrice il Mulino

UN SISTEMA DI VALUTAZIONE DELLA PSICOLOGIA ATTENTO AL GENERALE E AL PARTICOLARE

REMO JOB

Università di Trento

Riassunto. Nel mio commento allo stimolante intervento di Luca Tommasi (2020) propongo che la valutazione tenga conto degli aspetti sia generali, trasversali, sia specifici; che l'identificazione delle specificità sia compito dell'area stessa, non solo in funzione valutativa ma come tensione al miglioramento della qualità in sé; che sia improprio chiedere all'ANVUR di svolgere funzioni vicarianti rispetto a quelle proprie delle strutture disciplinari e le associazioni che le rappresentano.

L'intervento di Luca Tommasi (2020) tocca molti temi relativi al sistema di valutazione degli Atenei assumendo il punto di vista della psicologia e proponendo che sia necessario tener conto della sua specificità. Nel mio commento mi limito a discutere due nodi critici: la specificità deve essere considerata all'interno di una cornice generale ben definita; all'ANVUR non può essere attribuita una funzione vicariante di compiti e azioni proprie della nostra area disciplinare.

La riflessione di Tommasi si basa su una convinta adesione al sistema della valutazione, rivela cognizione di causa e competenza e si avvale dello studio della documentazione pertinente. Inoltre, il suo sforzo di offrire spunti di riflessione direttamente ancorati alla psicologia è encomiabile. Molte delle osservazioni sono condivisibili, altre meno. Nel mio breve commento vorrei sollevare alcuni punti che pertengono a quella che Tommasi indica come mancanza di specificità del sistema di valutazione ANVUR.

ANVUR offre un'analisi di processo e di prodotto, che è necessario ed è possibile migliorare dopo accurato monitoraggio, ma lo fa a un livello generale affinché possa essere applicata a tutte le aree di ricerca e didattica proprie dell'Università. All'interno di questo quadro generale sono identificate alcune specificità, come la distinzione fra aree bibliometriche e non-bibliometriche per la ricerca, ma concordo con Tommasi che ciò non avviene per la didattica. Tommasi considera la mancanza di specificità un limite in quanto «La prassi valutativa, schiacciando tutti sotto il peso dell'adeguamento a prerequisiti universali, fiacca la riflessione sulle esigenze, sulle tipicità e sui prerequisiti

Estratti riservati ai soci AIP, © Società editrice il Mulino

irrinunciabili di area». Vi sono due possibili letture di questa asserzione, che riguarda sia la ricerca sia la didattica (ed è quest'ultima che io prenderò in considerazione): la valutazione non deve essere generale ma specifica; la valutazione deve essere pluri-livello, generale e specifica. Io, da sempre, sostengo la seconda perché sono convinto che una valutazione del sistema universitario debba basarsi su criteri che, nello stesso tempo, tengano conto sia dei processi che caratterizzano ogni processo formativo sia delle specificità legate alle aree disciplinari che si differenziano per contenuti, metodi e strumenti. Non mi è chiaro quale delle due letture Tommasi prediliga: l'oculata scelta semantica dei verbi «schiacciare» e «fiaccare» sembra indicare una direzione, ma poiché non amo attribuire ad altri posizioni sulla base di «indizi» preferisco che Tommasi espliciti la sua, se lo desidera.

Gli strumenti utilizzati da ANVUR per la valutazione degli aspetti trasversali della didattica mi sembrano adeguati in quanto considerano, fra l'altro, la coerenza tra profilo in uscita e attività didattiche, il rapporto nei CdS tra didattica e ricerca, il ruolo dei diversi attori (docenti, studenti, personale TA, parti sociali) nel processo di assicurazione della qualità, e pertanto sono in grado di rilevare sia le eventuali carenze sia le buone pratiche. Questa valutazione di carattere generale è essenziale e insostituibile, anche se è necessario discutere se le modalità con cui essa è attualmente resa operativa siano del tutto accettabile o suscettibile di miglioramenti. Per certi aspetti, infatti, è meno generale di quanto si vorrebbe in quanto molto focalizzata sulle singole azioni di valutazione e meno sull'assicurazione di qualità a livello di sistema.

Nell'attuale modello ANVUR rimane carente, non so se per scelta, per dimenticanza, o per necessità, l'analisi degli aspetti specifici che dovrebbe integrare quella degli aspetti trasversali. Di fatto, una valutazione pluri-livello permetterebbe confronti sensati fra CdS di area diversa per quel che riguarda gli aspetti trasversali e fra CdS della medesima area sulla base di criteri più legati alle specificità della disciplina. Questa è una proposta che merita di essere approfondita perché potrebbe permettere una visione complessiva più articolata di ogni area.

Va rilevato che nella documentazione attuale, in fase di istituzione e in fase di riesame dei CdS, vi sono già spazi che richiedono analisi che si declinano per area sia in modo esplicito (ad esempio, la tipologia delle parti interessate consultate, il confronto con i CdS della stessa classe a livello nazionale e a livello territoriale), sia in modo implicito (ad esempio, l'impatto della ricerca sulla didattica, che in genere riguarda maggiormente i laboratori per le aree scientifiche e le biblioteche nelle discipline umanistiche), ma è pur vero che mancano quesiti che consentano analisi mirate su aspetti specifici. Tommasi evidenzia il tema delle classi di numerosità; io aggiungo l'uscita professionalizzante della laurea che costringe i CdS a conciliare una

Estratti riservati ai soci AIP, © Società editrice il Mulino

solida preparazione di base finalizzata alla laurea Magistrale (preferita da più del 90% di studenti e studentesse) con la necessità di offrire competenze immediatamente spendibili nel mercato del lavoro a chi desidera sostenere l'esame per l'accesso alla Sezione B dell'Albo degli Psicologi. Forse non è una specificità solo della Psicologia ma nella nostra area pesa molto. Il problema è complesso in quanto vede in gioco la normativa vigente, l'identità della Psicologia come disciplina dotata di un apparato teorico/metodologico complesso, la trasparenza nei confronti di studenti e studentesse e, *last but not least*, le risorse, anche di docenza, a disposizione delle singole istituzioni che offrono i CdS. La difficoltà a livello di valutazione nasce dal fatto che, mentre nella SUA sono esplicitati i profili professionali legati alla Laurea, con i relativi codici ISTAT, la quasi totalità dei Corsi di laurea in Psicologia offre, secondo me correttamente, percorsi ricchi di contenuti teorici e metodologici di base finalizzati all'ingresso alla Laurea Magistrale, limitando però le attività didattiche funzionali all'acquisizione di competenze coerenti con il profilo professionalizzante per la sezione B dell'Albo. CdS con obiettivi plurimi possono garantire alta qualità e trasparenza nei confronti dei potenziali iscritti/e?

Quale sia l'approccio adottato (analisi delle specificità o modello pluri-livello), il problema di fondo è sempre uguale, ovverosia chi debba identificare le specificità della psicologia. A tratti sembra che Tommasi postuli l'esistenza di una specificità della psicologia che l'ANVUR deve identificare e sottoporre a valutazione. Ma, a questo proposito, mi chiedo se non si stia qui chiedendo ad ANVUR di svolgere una funzione vicariante di quanto potrebbe o, meglio, dovrebbe, fare ciascuna area, compresa la nostra, ovvero identificare e farsi carico di elevare gli standard, di migliorare l'offerta didattica, di operare una *moral suasion* nei confronti di coloro che compongono l'area e le sue diverse articolazioni affinché assumano come proprio l'obiettivo e mettano in atto le azioni adeguate. L'identificazione delle specificità richiede, fra l'altro, capacità di analisi del contesto, creazione di consenso, scelta degli interlocutori, e degli alleati, con cui interfacciarsi. Un esempio di successo è stata la collocazione della psicologia nell'area delle discipline bibliometriche che ha visto una fase iniziale di accesa discussione, la creazione di una alleanza preziosa con le società scientifiche vicine e un'interlocuzione autorevole con ANVUR che ha giudicato convincente la proposta e l'ha adottata. Un esempio di mancato successo, invece, è quello indicato da Tommasi: «il problema della numerosità di classe resta un problema irrisolto nonostante diverse siano state le occasioni per trovare nei valutatori degli "alleati disciplinari" in grado di retroagire sul sistema circa la gravità di questo problema». Il tema delle numerosità di classe che ci vede in una classe di numerosità inaccettabile richiederebbe un lungo

Estratti riservati ai soci AIP, © Società editrice il Mulino

discorso. Una parte di responsabilità iniziale pesa forse sulla psicologia stessa se il Ministero ha creato le quattro classi basandosi sul numero (molto alto per la psicologia) di iscritti dell'anno accademico precedente alla determinazione, come Cubelli disse in un'assemblea AIP. Negli anni novanta e, in alcune sedi, negli anni recenti e ancor oggi, non sono stati posti limiti al numero delle immatricolazioni. A lungo si è pensato che quanto più alto fosse il numero di studenti e studentesse di psicologia tanto maggiore sarebbero stati la visibilità e il prestigio della disciplina all'interno degli Atenei e della società. Forse per questo non ricordo una risoluzione della CRUI su questo tema (ma la mia memoria può ingannarmi). Ovviamente, è inaccettabile che una decisione basata su criteri operativi già così dubbi non sia stata modificata a seguito dei tanti documenti (congiunti o singoli della Conferenza dei Presidi, di AIP, CPA e CNOP), ricchi di dati e argomentazioni, presentati dalla Psicologia negli ultimi 20 anni circa, in tutte le sedi possibili, dal Ministero al CUN. La mancata soluzione del problema segnala, secondo me, la debolezza politica della psicologia. Si può discutere se sia produttivo coinvolgere l'ANVUR come interlocutore su questo tema, come, a suo tempo, è stato fatto con il CUN, ma allora suggerisco sia fatto in modo esplicito e ufficiale, non attraverso l'azione di pochi valutatori nel ruolo di «alleati disciplinari»; tuttavia, è opportuno ricordare che interlocuzione non significa delega: sicuramente non possiamo chiedere all'ANVUR di sopperire alla mancanza di peso politico della nostra area disciplinare.

BIBLIOGRAFIA

TOMMASI L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

A general-to-specific quality assurance model for psychology

Summary. In this note on Luca Tommasi's stimulating commentary I argue that *a*) the assessment and quality assurance procedures about universities should take into account both general, i.e. across-the-board, and specific aspects; *b*) the analyses of specificities should be entrusted to the scientific area of interest (in this case, psychology) since this would allow to perform a fine-detailed assessment while, at the same time, favouring the spread and acceptance of the quality assurance process; *c*) it is improper for academic structures and relevant associations to delegate to ANVUR some of their own responsibilities.

Keywords: Quality assurance, assessment of universities, psychology.

La corrispondenza va inviata a Remo Job, Dipartimento di Psicologia e Scienze Cognitive, Università di Trento, Corso Bettini 31, 38068 Rovereto. E-mail: remo-job@unitn.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

VALUTARE PROCESSI E VALUTARE PRODOTTI: LA COMPLESSITÀ NELLA VALUTAZIONE DELLA QUALITÀ IN UNIVERSITÀ

LORENZO AVANZI[1], AMELIA MANUTI[2], MONICA MOLINO[3] GUENDALINA GRAFFIGNA[4] E DINA GUGLIELMI[5],

[1] Università di Trento, [2] Università di Bari, [3] Università di Torino, [4] Università Cattolica del Sacro Cuore, Piacenza, [5] Università di Bologna

Riassunto. Questa breve nota intende approfondire alcuni spunti emersi nella lettura dell'articolo di Luca Tommasi sul tema complesso della valutazione nel contesto universitario alla luce dei grandi cambiamenti dei sistemi di qualità avvenuti nel mondo accademico italiano in questo «decennio turbolento». L'obiettivo di questo contributo è quello di discutere le implicazioni che tali cambiamenti hanno generato per la Psicologia e nello specifico per il settore della Psicologia del Lavoro e delle Organizzazioni, proponendo al contempo alcuni possibili ambiti di sviluppo.

Negli ultimi anni, il tema della valutazione della psicologia è diventato un ambito di discussione accesa e complessa, su cui si confrontano posizioni talvolta anche molto distanti tra loro, frutto di differenti tradizioni di ricerca e dell'appartenenza a diverse generazioni di studiosi, con percorsi di carriera molto variegati. In particolare, nell'ambito della sezione AIP di Psicologia per le Organizzazioni questo tema è stato spesso oggetto di riflessione collettiva in seno alle assemblee dei soci o nel contesto di iniziative scientifiche monotematiche.

L'articolo di Tommasi (2020) offre una preziosa occasione per esaminare la rilevanza di questo tema alla luce dei grandi cambiamenti dei sistemi di valutazione avvenuti nel mondo accademico italiano in questo «decennio turbolento». Questa breve nota intende approfondire alcuni spunti emersi nella lettura dell'articolo soprattutto alla luce delle implicazioni che tali cambiamenti hanno generato per la Psicologia e nello specifico per il settore della Psicologia del Lavoro e delle Organizzazioni, proponendo al contempo alcuni possibili ambiti di sviluppo.

La crescente attenzione che in quest'ultimo decennio il sistema universitario ha dedicato alla ridefinizione dei sistemi di valutazione delle performance (siano esse legate alla didattica, alla ricerca e/o alla terza missione) segnala innanzitutto il bisogno di presidiare processi divenuti

Estratti riservati ai soci AIP, © Società editrice il Mulino

sempre più complessi e stratificati, verificandone l'andamento rispetto a criteri di successo definiti a monte. Questa è infatti la funzione principale dei sistemi di valutazione. In questo sforzo di adeguamento, l'introduzione di nuovi sistemi di valutazione opera sui comportamenti dei membri di un'organizzazione, modificandoli ed orientandoli.

Alla luce di questa riflessione, numerose sono le evidenze di cambiamenti positivi introdotti dai sistemi di valutazione anche nel nostro SSD: una più ampia internazionalizzazione della ricerca, in termini di divulgazione su riviste scientifiche diffuse in contesti spesso multidisciplinari, una maggiore partecipazione a gruppi di ricerca nazionali ed internazionali, anche con ruoli di leadership, una crescita delle citazioni da parte della nostra comunità, una riduzione delle tendenze alla autoreferenzialità e delle pratiche di ricerca senza sistemi di referaggio. Dunque, da questa prospettiva, anche interna al nostro settore, questi elementi sono stati letti come propulsori di un processo di crescita e arricchimento delle pratiche di ricerca.

D'altra parte, come ben evidenziato nell'articolo, numerosi sono anche i limiti generati da questi cambiamenti. Ci limitiamo a segnalare i più rilevanti per il nostro settore: scelte opportunistiche relativamente ad oggetti di ricerca e sedi di pubblicazione orientate a soddisfare gli indicatori piuttosto che la qualità della ricerca, disegni di ricerca e metodologie di analisi dei dati molto raffinate e sofisticate, ma spesso poco efficaci nel dare risposta ai reali bisogni delle organizzazioni e della società, con il rischio di perdita di innovatività e di diminuzione del valore applicativo dei risultati della ricerca.

Indubbiamente, la valutazione ha rappresentato una spinta al miglioramento, tuttavia necessita di un apprendimento dall'esperienza che non sempre trova spazio nel nostro sistema. La sfida, non facile, sarebbe di ottimizzare i vantaggi cercando di colmare le lacune evidenti. Se, da un lato, a noi sembra un elemento di positività il fatto che i metodi di valutazione si evolvano, cercando di migliorarsi in base all'esperienza, dall'altro riteniamo utile una riflessione che concorra a definire in maniera chiara i tempi e le forme del cambiamento, al fine di garantire processi di valutazione delle diverse colonne portanti dell'attività accademica (la ricerca, la didattica – come suggerito da Tommasi – ma anche sempre più a pieno titolo la terza missione), capaci di valorizzare le specificità delle differenti tradizioni disciplinari della psicologia e favorire la crescita scientifica dei singoli e delle comunità. Per raggiungere questi obiettivi appare importante sviluppare un sistema che utilizzi gli esiti della valutazione per un miglioramento riflessivo continuo, processo che richiede di conoscere in anticipo quali sono i comportamenti premiati.

All'interno di questo dibattito complesso, un utile suggerimento proposto da Tommasi riguarda la necessità di definire cosa è per noi

Estratti riservati ai soci AIP, © Società editrice il Mulino

qualità. Infatti, se non ne diamo una definizione operativa e attuale, «per la psicologia» diventa particolarmente difficile valutare la qualità.

Seguendo la tradizione dei sistemi organizzativi di cui ci occupiamo come Psicologi del Lavoro, la qualità è «dentro e fuori», riguarda il processo ed il prodotto: da un lato le procedure organizzative che concorrono a «creare qualità» (*Total Quality Management*), dall'altro la valutazione che il destinatario/cliente/utente esprime della qualità in termini sia di prodotto sia di servizio (*Customer Satisfaction*). In tal senso, secondo una visione antropocentrica, quella che la letteratura di settore etichetta come l'approccio della «Big Q» (Cole e Scott, 2000; Tarì e Sabater, 2006), la qualità può essere intesa come l'area di interazione tra il soggetto, con i suoi bisogni e le sue aspettative e l'oggetto con le sue caratteristiche intrinseche ed estrinseche. Pertanto, la qualità del prodotto/processo è implicita nella qualità del progetto organizzativo, nel senso che dovrebbe essere parte di una cultura (della qualità), di un «modo di fare le cose» che regola il funzionamento globale dell'organizzazione, i suoi processi ed i comportamenti dei suoi membri e che si esprime nell'output finale e nel modo in cui viene erogato.

Se applichiamo questa definizione di qualità al contesto universitario, è innanzitutto possibile distinguere la qualità dei processi dalla qualità dei prodotti. Tuttavia, tra i due piani, tra il dentro ed il fuori, esiste uno scambio continuo per quanto riguarda la qualità sia della ricerca sia della didattica.

La qualità dei nostri prodotti scientifici e degli interventi che programmiamo ed eroghiamo nei contesti organizzativi non può essere svincolata dalla loro capacità di operare secondo un metodo scientifico orientato a spiegare o ad interpretare la realtà (a seconda della prospettiva epistemologica entro cui si situa l'intervento). Dunque, la qualità della ricerca sia essa di base o applicata, deve avere come priorità la produzione di conoscenza, ma deve essere anche ricerca del senso, ricerca dell'utilità e dell'applicabilità dei risultati cui perveniamo. Similmente, la qualità del sistema didattico si fonda sulla capacità della formazione di trasferire non solo contenuti (i prodotti) ma anche quel metodo (i processi) che consenta di leggere le organizzazioni, di interpretarne i bisogni e di proporre soluzioni. Formare nuovi psicologi non solo al rigore scientifico e metodologico del sapere psicologico (aspetto imprescindibile) ma anche alla capacità di dialogare con il mondo degli stakeholders, alla capacità di intercettare problemi della comunità di riferimento e di contribuire alla loro risoluzione dovrebbe essere un obiettivo primario della nostra attività didattica. Per sostenere questo processo virtuoso inevitabilmente anche l'attività di ricerca scientifica dovrebbe essere caratterizzata dal bilanciamento tra rigore scientifico-metodologico e capacità di interrogarsi rispetto ai problemi del mondo reale. I processi di valutazione della didattica, della ricerca e della terza

Estratti riservati ai soci AIP, © Società editrice il Mulino

missione potrebbero costituire un elemento propositivo fondamentale per stimolare il ripensamento critico sulle pratiche accademiche e per alimentare questo virtuoso bilanciamento nel dialogo tra «dentro e fuori» l'accademia. Di qui l'importanza di processi di valutazione che dovrebbero tener conto, accanto agli indicatori oggettivi, indicatori suscettibili di cogliere il cambiamento che la ricerca e la didattica operano nel contesto. In quest'ottica, la vocazione applicativa del nostro settore consente alla ricerca di orientare le proprie pratiche verso l'utilizzo del sapere psicologico ed il metodo empirico a vantaggio delle concrete esigenze dei sistemi organizzativi oggetto di studio e dunque operando cambiamenti sui comportamenti e sugli atteggiamenti dei propri membri. Allo stesso modo, la didattica dovrebbe essere ripensata e valutata rispetto alla sua capacità di sollecitare la riflessione critica, l'apprendimento per esperienza, il confronto tra pari, l'impiegabilità dei laureati e la loro capacità (a cascata) di generare cambiamento positivo nel contesto sociale di riferimento. Tutte dimensioni che la letteratura internazionale indica come rilevanti nella definizione di un modello di qualità della didattica, orientata alla costruzione ed alla trasformazione di competenze piuttosto che al trasferimento ed alla tesaurizzazione delle conoscenze (Gibbs, 2006, 2010; Marshall, 2016; Williams, 2016).

Queste riflessioni richiederebbero chiaramente una controparte operativa in termini di valutazione non immediata e, infatti, il tema delle proposte è forse il punto più spinoso del dibattito e meriterebbe una disanima della stessa lunghezza dell'articolo, qui ci limitiamo a segnalare alcuni punti particolarmente importanti per il nostro settore.

Se il metodo bibliometrico è una base di partenza non modificabile, può essere migliorato il ricorso a criteri qualitativi negli ambiti dove utili, si pensi ad esempio all'ASN o alla VQR. Nonostante i punti di forza prima richiamati, gli indicatori bibliometrici non possono infatti essere esaustivi della qualità per varie ragioni già citate in precedenza, dai sistemi distorsivi (come le autocitazioni) alle diverse tradizioni di ricerca che ancora persistono nel nostro raggruppamento non tutte con la stessa accessibilità a riviste con indicatori elevati anche quando presentano ricerche rigorose, originali e innovative. Su questa linea, basandosi sugli apprendimenti derivanti dalle valutazioni passate, si potrebbero avanzare micro-proposte di miglioramento. La maggiore valorizzazione delle esperienze didattiche e di terza missione, a fianco della produttività in termini di ricerca, potrebbe essere un primo correttivo del sistema di valutazione in atto. Inoltre, la possibilità di approfondire la dimensione di impatto sociale delle attività accademiche dei singoli dovrebbe essere un ulteriore punto di riflessione nell'ambito dei processi di valutazione delle carriere.

Un altro punto riguarda l'assenza di indicatori di valutazione in relazione alla nostra capacità di sistema di entrare in circuiti di ricerca

Estratti riservati ai soci AIP, © Società editrice il Mulino

di più ampia portata (finanziamenti, progetti innovativi), tema poco presente nei diversi sistemi discussi nell'articolo. Così come l'assenza di indicatori relativi alla capacità di intercettare e rispondere alle istanze del mondo sociale reale (aspetto più volte ribadito in questo scritto come importante per gli Psicologi per le Organizzazioni, e potenzialmente per tutte le discipline psicologiche) e il costruire reti multidisciplinari e istituzionali.

Infine, se l'oggetto di valutazione è «qualità» della performance accademica merita maggiore attenzione il tema della didattica. Da un lato, infatti, come ampiamente discusso dall'autore nell'articolo bersaglio, la centratura è sul processo di *quality assurance* dei corsi di studio, processo prettamente di tipo organizzativo che coinvolge coordinatore e gruppo qualità nel monitoraggio del CdS a un primo livello e Dipartimenti e Ateneo negli altri, con un'assenza importante della «qualità» della didattica universitaria, che abbia come scopo fondamentale quello di garantire agli studenti un apprendimento significativo ed efficace e di passare loro pensiero critico e la capacità di intercettare e rispondere ai bisogni del contesto reale. Pensiamo, ad esempio, all'impegno dei docenti di migliorare costantemente la proposta didattica, anche nei metodi e nella interazione con studenti, aspetti non considerati dai diversi sistemi di valutazione in essere. La sottostima della didattica, come segnala anche Tommasi, è evidente anche dal fatto che tra tutte le attività di valutazione solo il sistema AVA si occupa di valutare gli aspetti didattici (dedicando tra l'altro spazio all'AVA di sede e di Dipartimento) ed è del tutto assente dai meccanismi di valutazione delle persone (ASN e FFABR). Questo tema probabilmente risente di resistenze al cambiamento che contraddistinguono il mondo accademico, perché, mentre sulla valutazione della ricerca il processo ha preso avvio diversi anni fa e molti elementi sono dati per acquisiti (e si può quindi ragionare su cosa ha funzionato e cosa va modificato), sul versante della didattica anche i processi di valutazione interni agli Atenei hanno in diverse situazioni trovato ostacoli. Se da un lato, questo è comprensibile per le difficoltà intrinseche a questo tipo di valutazione tra cui le differenze tra discipline, ciclo di studio e dimensioni del corso, dall'altro resta in secondo piano e non vengono valorizzati aspetti come le metodologie didattiche utilizzate, le interazioni con gli studenti, l'apprendimento e la professionalizzazione degli studenti, le interconnessioni e gli scambi con il mondo del lavoro proposti durate il percorso di studi. La scarsa valorizzazione della didattica va, inoltre, considerata anche nella prospettiva dell'impegno del docente e dei risultati negli altri ambiti in cui è coinvolto. Riflessione che in generale può essere allargata a un'ampia fetta di lavoro istituzionale.

Estratti riservati ai soci AIP, © Società editrice il Mulino

È complesso normare questi aspetti e occorre attenzione nel definire possibili criteri quali-quantitativi in maniera netta perché il rischio è di indirizzare i comportamenti verso i soli criteri valorizzati, ma è evidente la necessità di ampliare lo spazio dedicato alla valutazione della didattica, nello specifico per gli aspetti prima richiamati e ad oggi assenti.

Più in generale, la necessità di valorizzare la didattica e la terza missione, a fianco della ricerca, ci riporta a uno dei nodi cruciali della valutazione universitaria, ossia che i ruoli accademici si caratterizzano per un'attività molto variegata, difficilmente sintetizzabile con i soli indicatori, dove l'integrazione con il qualitativo può aiutare sapendo però che il meccanismo sottostante che ormai governa i processi di valutazione e le scelte dei singoli può diventare un'arma a doppio taglio.

BIBLIOGRAFIA

COLE R., SCOTT W. (2000). *The quality movement and organization Theory.* Thousand Oaks: Sage.

GIBBS G. (2006). Why assessment is changing. In C. Bryan, K. Clegg (eds.), *Innovative assessment in higher education*. London: Taylor and Francis, pp. 23-36.

GIBBS G. (2010). *Dimensions of quality*. York: Higher Education Academy.

MARSHALL S. (2016). Quality as sense-making. *Quality in Higher Education*, *22* (3), 213-227.

TARÌ J., SABATER V. (2006). Human aspects in a quality management context and their effects on performance. *International Journal of Human Resource Management*, *17* (3), 484-503.

TOMMASI L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

WILLIAMS J. (2016). Quality assurance and quality enhancement: Is there a relationship? *Quality Assurance in Education*, *22* (2), 97-102.

Assessing processes and assessing products: The complexity of quality assessment in the academic context

Summary. This short note aims to analyze some of the suggestions coming from the reading of the paper by Luca Tommasi focused on the complex issue of quality assessment in the academic context in light with the great changes that have invested the Italian Universities in this «turbulent decade». The main aim is to discuss about the implications brought about by these changes for Psychology and specifically for Work and Organizational Psychology as disciplinary domain. In the meantime, some possible future developments are proposed.

Keywords: Assessment, quality, research, teaching, third mission.

La corrispondenza va inviata a Dina Guglielmi, Dipartimento di Scienze dell'Educazione, Università di Bologna, Via Filippo Re 6, 40126 Bologna. E-mail: dina.guglielmi@unibo.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

AUTOVALUTAZIONE, VALUTAZIONE PERIODICA E ACCREDITAMENTO DEI PROCESSI O DEGLI ESITI?

FABIO LUCIDI

Sapienza Università di Roma

Riassunto. Ci si focalizza su due questioni entrambe concernenti AVA. La prima è di natura generale e si riferisce alla mancata corrispondenza tra la valutazione di qualità e valutazione dei processi di assicurazione di qualità. La seconda fa riferimento a un tema specifico che è quello della valutazione degli esiti dell'apprendimento attraverso i test sulle competenze disciplinari.

L'articolo di Luca Tommasi (LT, 2020) affronta al ruolo dei numerosi organismi di valutazione con cui l'università è chiamata a confrontarsi, anche i processi legati all'Abilitazione Scientifica Nazionale), al percorso di Autovalutazione, Valutazione periodica e Accreditamento di sedi e CdS, gli esercizi Valutazione della Qualità della Ricerca.

Nessun organo della pubblica amministrazione in Italia è stato capace e disponibile a collaborare alla messa a punto di meccanismi e processi per la propria valutazione come l'università. Anche se questo non è stato adeguatamente riconosciuto e valorizzato dalla opinione pubblica e dalla politica che tendono tuttora ad accusare l'accademia di essere autoreferenziale, personalmente trovo questo del tutto comprensibile e giusto. L'Università è per sua natura contesto del tutto avvezzo ai processi di valutazione. Inoltre l'università rivendica costantemente e giustamente la propria autonomia responsabile e la valutazione, quando funziona, è uno strumento dell'autonomia. Naturalmente quando non funziona essa rischia di tradire obiettivi e valori per cui è stata introdotta e accolta favorevolmente da molti di noi, trasformandosi, al contrario, in strumento di controllo e limitazione. Se e quali processi di valutazione introdotti nel sistema universitario funzionino e come migliorare quelli con le maggiori criticità è domanda complessa su cui molti e illustri colleghi si confrontano continuamente. Non sarò quindi io a contribuire a questo dibattito limitandomi, nello spazio che ho a disposizione, a concentrarmi esclusivamente su due questioni entrambe concernenti AVA. La prima è di natura generale e si riferisce alla mancata corrispondenza tra la valutazione della qualità e la valutazione dei processi di assicurazione di qualità. La seconda fa riferimento a un tema

Estratti riservati ai soci AIP, © Società editrice il Mulino

specifico che è quello della valutazione degli esiti dell'apprendimento attraverso i test sulle competenze disciplinari.

Come giustamente scrive LT, nel lessico comune la parola «qualità» indica un valore tanto confuso quanto positivo, un prodotto o un servizio di qualità ha le caratteristiche desiderabili e promette di soddisfare chi ne fruisce; è adeguato a uno scopo o utile per una data funzione. Una università di qualità può essere definita come dice LT, riferendosi «... ai *propri* più alti ideali di verità, giustizia e bellezza...». Già questo avverbio relativizza il concetto di qualità in ambito universitario, contesto nel quale ogni valutatore deve fortunatamente fare i conti con un ulteriore valore positivo che ho già citato: quello dell'autonomia. Ogni ateneo, in autonomia e coerenza con il proprio statuto, ha il diritto di esprimere (esplicitamente) una propria visione della qualità della didattica, della ricerca e della terza missione, dandosi obiettivi propri e mettendo in atto specifiche strategie per raggiungerli. Questa concezione di autonomia, naturalmente, corrisponde a responsabilità. A partire dalla legge 240/2010 il questo concetto è operazionalizzato nel d.lgs. 19/2012 attraverso il sistema di Autovalutazione, Valutazione e Accreditamento (AVA) finalizzato a responsabilizzare gli atenei nelle scelte effettuate non solo ai fini dell'accreditamento dei corsi di laurea e di dottorato, ma anche delle politiche di bilancio di ateneo e dell'accesso alla quota premiale del FFO. In sostanza, l'idea del legislatore era di incentivare le università a definire le proprie politiche e a selezionare (attraverso una autovalutazione) gli indicatori della loro realizzazione da sottoporre poi alla valutazione dell'ANVUR. Per dirla in parole povere, di fronte a una coperta corta, sono gli atenei a stabilire se daranno priorità alla ricerca dell'eccellenza o alla riduzione delle diseguaglianze, fermo restante l'esercizio da parte degli atenei di un'autonomia responsabile e affidabile nell'utilizzo delle risorse pubbliche (ovvero nel rispetto dei parametri della sostenibilità economica) e nei comportamenti collettivi e individuali relativi alle attività di formazione e ricerca. Una volta definiti i propri obiettivi di qualità – ovvero le politiche – alla luce della propria autonomia responsabile, gli atenei possono essere valutati sul modo in cui realizzano processi e strumenti adeguati per raggiungerli – ovvero il processo di Assicurazione di qualità (AQ). Si usa questo termine per descrivere l'insieme delle azioni, prima programmate e poi effettuate dalle sedi, per accertarsi che didattica, ricerca e terza missione siano efficaci ai fini stabiliti. Le azioni dell'AQ servono dunque agli atenei non per dimostrare che tutto va bene, ma per permettere i rilevare i problemi laddove esistenti, approfondirli, trovare possibili soluzioni ad essi. Non è quindi *strictu sensu* la qualità di sedi e CdS a dover essere valutata da AVA bensì i processi messi in atto da sedi e CdS per assicurarla. Come scrive LT, questo percorso logico (in linea con l'attuale normativa) probabilmente mette

Estratti riservati ai soci AIP, © Società editrice il Mulino

insieme una non-definizione di qualità con una attenzione concentrata prevalentemente su metodi e strumenti per valutarla. Però allo stesso tempo non comprime (o comprime solo in parte) l'autonomia universitaria. Differentemente da quanto Tommasi mi sembra suggerire, la difficoltà di reperire una definizione alta (assoluta?) di qualità non è a mio parere tra i limiti del sistema AVA. Al contrario, ribadisco che il rischio insito nel fornire una tale definizione sarebbe quello della riduzione nella autonomia funzionale delle università. I detrattori del sistema di valutazione ritengono che i meccanismi propri della VQR e dell'ASN già hanno questo effetto, quanto meno nel senso di spingere i ricercatori verso una standardizzazione e sovra-produzione quantitativa della propria produzione scientifica. L'accusa spesso mossa ad AVA è forse quello di essere farraginosa e di rappresentare più uno *stressor* che un opportunità per il sistema, ma difficilmente si può sostenere che essa limiti la libertà di insegnamento, concetto che ovviamente non va confuso con quello di scarsa trasparenza, discrezionalità, arbitrarietà nella proposta formativa.

Se dunque, e con una logica, le procedure valutative AVA della didattica poggiano sulla verifica dei *processi* di assicurazione della qualità, ci si potrebbe chiedere se il grande assente non sia proprio l'*esito* formativo. In realtà, da alcuni anni e con alterne vicende, sono stati messi in atto dei tentativi di fornire strumenti per l'autovalutazione dei livelli di competenze forniti dai CdS agli studenti universitari. In effetti anche le attività di valutazione della qualità degli esiti dell'offerta formativa sono normate dal DPR costitutivo dell'ANVUR (76/2010, art. 3). In particolare, sono oggetto di valutazione l'efficienza e l'efficacia dell'attività didattica sulla base di standard qualitativi, anche in riferimento agli esiti dell'apprendimento (art. 3, com. 2 lettera *a*). In effetti, un progetto di questo genere (progetto TECO – Test sulle Competenze) è stato avviato dall'ANVUR a partire dal 2012 all'interno di AVA e prevede la definizione e somministrazione di test in grado di rilevare le competenze trasversali (TECO-T) e disciplinari (TECO-D) acquisite dagli studenti durante i corsi di studio triennali. Le prove TECO-T vengono realizzate direttamente dall'Agenzia con la collaborazione di gruppi di esperti. La rilevazione delle competenze disciplinari, invece, si avvale dell'impegno di gruppi disciplinari chiamati a costruire le prove. Ho avuto, in questi ultimi mesi, il privilegio di coordinare il gruppo disciplinare TECO-D per la L-24 (la triennale in Psicologia), insieme con i colleghi Cassibba (Università di Bari), Di Nuovo (Università di Catania), Di Pellegrino (Università di Bologna), Fraccaroli (Università di Trento), Galati (Sapienza – Università di Roma), Ghisi (Università di Padova), Mosso (Università di Torino), Sacchi (Università di Milano-Bicocca), Salvatore (all'epoca, Università del Salento ora Sapienza), Vecchi (Pavia) e vorrei sfruttare

Estratti riservati ai soci AIP, © Società editrice il Mulino

questa occasione per ringraziare tutti loro per l'impegno e la disponibilità che hanno voluto dedicare a questo lavoro. Considerando che il TECO-D per la classe di laurea della Psicologia rappresenta una novità, vorrei sfruttare questa occasione per condividere alcune informazioni sul processo che ha portato alla preparazione della prova. Esso è partito, seguendo le linee guida dell'ANVUR dall'analisi delle Schede SUA dei Corsi di Studio (quadri del gruppo A4) allo scopo di poter definire un gruppo Obiettivi Formativi Finali (OFin) condivisi, ovvero una base di conoscenze e competenze che un laureando della L24 dovrebbe aver raggiunto, coerentemente con quanto dichiarato nelle SUA-CdS (quadri del gruppo A4) dei diversi Atenei. Partendo da questi materiali, dalle delibere e dalle prese di posizione della precedente Conferenza dei Presidi, di AIP, CPA e CNOP sulla formazione in Psicologia, anche sintetizzate su articolo pubblicato su questa rivista da Cassibba e Di Nuovo (2018), il gruppo di lavoro ha definito 4 Obiettivi formativi sostanzialmente condivisi in ciascuno di questi documento come parte della formazione di un laureato triennale della classe. Al fine di condividere con la comunità accademica l'esito di questa azione e di sottoporlo a una discussione pubblica, il gruppo ha presentato questi obiettivi formativi in un Workshop organizzato dall'ANVUR il giorno 8 Maggio 2018. Gli obiettivi formativi identificati sono quelli di:

1) Conoscere e comprendere i metodi empirici alla base della psicologia come scienza.

2) Conoscere e comprendere i processi del funzionamento della mente e del comportamento, inclusi i loro correlati psicobiologici, anche in un quadro di sviluppo nel ciclo di vita.

3) Conoscere e comprendere le relazioni interpersonali, intra e inter-gruppali, nei diversi contesti sociali, a livello di diade, famiglia, piccoli gruppi e organizzazioni.

4) Conoscere e comprendere le forme di alterazione del funzionamento mentale e del comportamento umano, anche in relazione a diverse fasce di età, livelli e contesti e le principali forme di intervento.

A partire dal Giugno dello stesso anno il gruppo ha iniziato a lavorare sulla declinazione degli OFin in Obiettivi Formativi Specifici (OFS) articolati secondo i Descrittori di Dublino. Nel farlo, come previsto dalle linee guida dell'ANVUR, il gruppo di lavoro si è di nuovo confrontato con le schede Sua-CdS delle diverse L24 dei diversi Atenei, partendo dai Quadri A4.b.2 e A4.b.3, con l'intento esplicito di ricercare competenze e competenze il più possibile trasversali alle proposte formative da essi erogate.

Sulla base di questo lavoro, i 4 Obiettivi formativi generali sono stati declinati in 26 Obiettivi formativi specifici che, a loro volta sono

Estratti riservati ai soci AIP, © Società editrice il Mulino

stati articolati in 209 aree di possibile domanda. Per avere una esempio del passaggio dal generale al particolare si consideri il seguente: Tra i diversi Obiettivi Specifici dell'Obiettivo Formativo (1) *Conoscere e comprendere i metodi empirici alla base della psicologia come scienza* vi sono gli *Elementi di Psicometria*; tra le molte aree di possibili domande entro gli elementi di psicometria vi è l'*Attendibilità.*

Terminato questo processo il gruppo ha cercato un ulteriore confronto con i rappresentanti della comunità accademica (Membri della Giunta CPA, Direttori dei Dipartimenti, presidenti dei Corsi di Studio) che si è svolto il 14 Febbraio 2019, per recepire commenti o suggerimenti. Da quel momento il gruppo ha assunto il compito di produrre diverse domande su ciascuna delle 209 aree generative, per che sono state dapprima a un processo interno di analisi e miglioramento al quale seguirà, nei prossimi mesi un test empirico per valutarne attendibilità e validità. Speriamo sinceramente che questo sforzo possa servire alla nostra comunità per disporre di un ulteriore strumento di autovalutazione. Abbiamo lavorato nella che esso non debba in nessun modo essere usato per inutili (più probabilmente dannose) classifiche tra le sedi, tra i corsi di studio. La logica di TECO dovrebbe garantirci da tutto questo, visto che i risultati verranno forniti ai Corsi di Studio stessi per permettere una riflessione al proprio interno circa le aree di migliore successo formativo o di possibile miglioramento.

BIBLIOGRAFIA

Di Nuovo S., Cassibba S. (2018). Per una classificazione dei saperi e una «manutenzione» dei sistemi formativi relativi alla psicologia: alcune riflessioni sulle criticità attuali e sulle possibili alternative. *Giornale Italiano di Psicologia*, *45* (2), 265-278.

Tommasi L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

Self-evaluation, periodic evaluation and accreditation of learning processes or learning outcomes?

Summary. This comment focuses on two issues both concerning AVA. The first is of a general nature and refers to the mismatch between quality assessment and quality assurance process assessment. The second refers to a specific issue which is the assessment of learning outcomes through disciplinary competence testing.

Keywords: Evaluation systems, psychology, learning outcomes, quality assessment.

La corrispondenza va inviata a Fabio Lucidi, Dipartimento di Psicologia dei Processi di Sviluppo e Socializzazione, Sapienza Università di Roma, Via dei Marsi 78, 00185 Roma. E-mail: fabio.lucidi@uniroma1.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Estratti riservati ai soci AIP, © Società editrice il Mulino

LA VALUTAZIONE DELLA QUALITÀ: TRA FORMALISMO, AZIENDALISMO, *COMPLIANCE* E «MERITO»

TERRI MANNARINI

Università del Salento

Riassunto. La riflessione sui sistemi di valutazione del mondo accademico Italiano porta alla luce diversi nodi, concettuali e metodologici: *a*) una concezione di «qualità», nell'ambito della didattica, che oscilla tra formalismo e aziendalismo, occultando le differenze e le disparità tra sedi e territori; *b*) la capacità dei sistemi di valutazione di orientare le pratiche della ricerca, con un inevitabile portato di conformismo soprattutto tra i ricercatori più giovani; *c*) una concezione individualistica del merito che trascura le condizioni di contesto e sul medio-lungo termine genera effetti disgregativi per la comunità accademica.

1. INTRODUZIONE

L'articolo di Luca Tommasi «La valutazione della psicologia: considerazioni su un decennio turbolento» (2020) offre uno spaccato articolato di come sia stato introdotto e sia andato configurandosi in Italia il sistema della valutazione accademica negli ambiti della didattica (AVA), della ricerca (VQR) e delle carriere (ASN), con un'appendice sul tradizionale, ma sempre più residuale, finanziamento della ricerca rappresentato dai PRIN. Il quadro tratteggiato nell'articolo, più che fornire una risposta alla domanda implicita che lo guida – «se vi siano aspetti della valutazione che possano essere considerati irrinunciabili o che al contrario dovrebbero essere modificati per migliorare prassi e metodologie operative» – mette in evidenza le numerose problematiche (e le molteplici possibili risposte) sollevate dal sistema di valutazione in atto, sia a monte, in termini di principi e relative traduzioni metodologico-operative, sia a valle, in termini di effetti. In questo commento, per natura breve, cercherò di mettere a fuoco tre punti su

Il contributo è da considerarsi espressione di un pensiero condiviso tra i membri del Comitato Esecutivo della Sezione Psicologia Sociale dell'AIP: Elvira Cicognani, Laura Ferrari, Maria Giuseppina Pacilli e Stefano Pagliaro.

Estratti riservati ai soci AIP, © Società editrice il Mulino

cui continuare il ragionamento aperto da Luca Tommasi, dando per acquisita l'importanza e le ragioni di una cultura della valutazione.

2. AVA: TRA FORMALISMO E AZIENDALISMO

Se c'è un'area della valutazione in cui la specificità della psicologia viene penalizzata più che in altre aree, a mio parere, dall'impiego di indicatori generali e di ordine formale, questa è quella dei corsi di studio. Della numerosità della classe L-24 si è discusso in diverse occasioni (Di Nuovo e Cassibba, 2018) e per questa ragione non mi dilungo: è opinione unanime nella psicologia accademica e professionale che un'elevata numerosità di studenti difficilmente possa coniugarsi con un'alta qualità della formazione, dal momento che impedisce una didattica mirata a sviluppare le *core competencies* del lavoro psicologico. Sotto questo profilo, l'indicatore AVA relativo al rapporto studenti/docenti non consente di portare alla luce in modo evidente il problema, mascherando attraverso i numeri situazioni molto differenziate.

La numerosità degli iscritti, che gli Atenei fanno fatica a ridurre in presenza di una domanda elevata sull'intero territorio nazionale, gioca a sfavore dell'occupabilità, come documentato dai dati sul mercato del lavoro. Ma l'occupabilità dei laureati è un indicatore che azzera differenze territoriali molto profonde, essendo la capacità dei contesti di assorbire i laureati in psicologia estremamente differenziata nelle regioni del paese. Da questo punto di vista, i corsi di studio e gli Atenei delle aree economicamente meno avanzate sono destinati a risultare, sulla base di questo indicatore, sempre di minor «qualità» rispetto agli altri. La natura formale e universale dei requisiti sui quali si valuta la qualità dell'offerta formativa fa dunque sì che non solo non emergano le specificità della psicologia, ma autorizza comparazioni che sarebbero legittime solo se tutti i corsi, le sedi, gli atenei e i territori avessero le stesse caratteristiche strutturali, culturali ed economiche. Si tratta evidentemente di una finzione strategica, necessaria all'implementazione operativa di un concetto formale di qualità.

La torsione aziendalistica della qualità emerge, invece, in forma evidente nel ruolo che il sistema AVA attribuisce alla valutazione che della didattica fanno gli studenti e le studentesse, anzi per la precisione alla loro «soddisfazione», che anche terminologicamente ricalca la logica della «customer satisfaction». Non si vuole mettere in discussione il principio che chi studia possa o debba essere coinvolto nel processo valutativo; tuttavia, c'è da dubitare che al centro della valutazione ci sia effettivamente la qualità dei corsi. Per esempio, si può obiettare che la valutazione non correla con l'apprendimento, e che quindi i/le docenti i cui corsi vengono valutati meglio non siano ne-

Estratti riservati ai soci AIP, © Società editrice il Mulino

cessariamente anche coloro che raggiungono meglio gli obiettivi formativi (Uttl, White e Gonzalez, 2016); che le valutazioni studentesche dicono qualcosa più su chi valuta che sulla qualità dei corsi valutati (Feistauer e Richter, 2017); che le valutazione sono soggette a distorsioni, come dimostrano gli studi sul *gender* and *race bias*, secondo cui le accademiche donne e gli accademici appartenenti a minoranze etnico-culturali ricevono valutazioni più negative dei colleghi maschi e di gruppi dominanti, a parità di altre condizioni (Boring, 2017; Chavez e Mitchell, 2020; MacNell, Driscoll e Hunt, 2015). La centralità, almeno retorica, che il sistema AVA assegna alla valutazione del corpo studentesco appare francamente sovradimensionata.

3. LA RICERCA: *COMPLIANCE* E CONFORMITÀ

Nel commentare gli esiti relativi alle tornate ASN, Luca Tommasi osserva che la comunità scientifica ha risposto in maniera molto *compliant* alla sfida valutativa, adattandosi ai criteri via via più richiestivi definiti dall'elevamento delle soglie bibliometriche. Tale allineamento sembra riguardare tutti i settori della psicologia, come già suggeriva l'analisi analisi bibliometrica della produttività condotta da Salvatore e colleghi (2015) qualche anno fa. L'osservazione si accompagna a quella relativa alla VQR, secondo cui si è registrato, nel passaggio dalla prima alla seconda VQR – parole del GEV di Area 11b della VQR2 – uno «straordinario progresso della psicologia», identificabile nella tendenza a privilegiare le pubblicazioni in lingua inglese nella forma dell'articolo scientifico su riviste indicizzate. Non so dire se si tratti di un «progresso», dal momento che il termine ha una connotazione semi-ideologica, e non mi spingerei a dire, sulla base di questi indicatori, che la qualità della ricerca in psicologia sia oggi più alta rispetto a 10-15 anni fa, ma certamente si tratta di un cambiamento. Che non è brusco e quindi non sorprende, ma che dice molto della capacità dei sistemi di valutazione di orientare le pratiche della ricerca.

Il tema è dibattuto: se i criteri per la valutazione definiscono gli standard di qualità per la disciplina, la pressione a conformarsi agli standard è forte, soprattutto se a questi standard si agganciano gli avanzamenti di carriera. È un dato di fatto che i/le nuovi/e docenti di psicologia si sono mossi/e, e continueranno a muoversi, dentro i parametri premianti della valutazione. Al di là dei correttivi tecnici che si possono apportare (e sono stati apportati, cfr. Cubelli, 2015) per non penalizzare le specificità dei singoli settori e dei diversi approcci, non si può negare che aver definito per la psicologia *una* norma di riferimento per la ricerca «di qualità» abbia immediatamente definito

Estratti riservati ai soci AIP, © Società editrice il Mulino

come «non abbastanza di qualità» la ricerca non standard. In termini più generali, non ha incoraggiato il pluralismo.

Non è recentissimo, ma a questo proposito vale la pena ancora citare lo studio di Gillies (2012) sugli effetti del *Research Assessment Exercise* (RAE, il sistema britannico di valutazione della ricerca) nel periodo 1986-2008 sulla ricerca in campo economico, un settore, come la psicologia, in cui si confrontano scuole e approcci diversi. Gillies sostiene, sulla base dei dati analizzati, che l'effetto del RAE è quello di rafforzare il *mainstream* e di indebolire i paradigmi minoritari. In pratica, c'è un rapporto proporzionale tra la valutazione che ricevono gli accademici che aderiscono a un approccio e la loro numerosità: maggiore numero di aderenti, migliore la valutazione.

È questo un vantaggio per una disciplina? Un «progresso», o un cambiamento auspicabile? In linea di principio credo che si possa concordare sul fatto che il pluralismo scientifico sia una fonte di innovazione, un elemento di dinamizzazione del campo delle idee. Dovendo fare il punto, si può quindi perlomeno affermare, astenendosi da un giudizio di merito, che il sistema di valutazione della ricerca, così come è andato configurandosi, indirizza oggi la psicologia verso uno stato di conservazione, uno stato – storicamente determinato, transitorio – di «scienza normale».

4. IL MERITO NELLE CARRIERE

L'idea, condivisa nella comunità accademica come nel resto della società, che il merito debba essere il perno del sistema di reclutamento e di avanzamento delle carriere si è talmente incistata nell'identità professionale accademica e nel comune sentire da aver assunto le caratteristiche della naturalità, cioè dell'indiscusso. La ASN incarna, nei principi ispiratori, questo principio di giustizia. Non intendo qui aprire una discussione sulla meritocrazia (per una lettura critica rimando a Boarelli, 2019), credo però che la psicologia dovrebbe interrogarsi di più sia sul «merito», sia sugli effetti che l'adozione di un sistema basato sul *puro* merito genera all'interno della comunità accademica.

Sul merito. Mi limito ad una sola osservazione, di ordine pragmatico. La valutazione del merito viene fatta a valle, sulla base di ciò che si è prodotto/raggiunto, omettendo di includere nella valutazione le condizioni a monte entro le quali il processo si è realizzato. Un esempio: due pubblicazioni scientifiche sulla stessa rivista *open access* ottengono (più o meno) la stessa valutazione in termini di qualità, indipendentemente dal fatto che gli autori e le autrici: abbiano o meno fondi ai quali accedere per poter pagare la quota di pubblicazione; abbiano nella loro sede laboratori e strumenti applicativi che consentano

Estratti riservati ai soci AIP, © Società editrice il Mulino

loro di fare ricerca in un certo modo; se giovani, abbiano un/a tutor in grado di consentire loro di fare delle esperienze all'estero, frequentare le scuole di metodologia, trasferire loro un «capitale accademico». Tutte queste (e altre) condizioni di contesto rendono il merito qualcosa di molto più complesso, e molto meno individuale, della capacità dei singoli. Dietro la democraticità del merito, che applica a tutti/e lo stesso principio, ci sono differenze nelle condizioni di lavoro e di vita che se non vengono riconosciute e pesate rischiano di rendere il merito il contrario di un principio di giustizia. Non si tratta di abbandonare il criterio del merito, ma di riconoscere i fattori che concorrono – al di là delle capacità individuali – al risultato meritevole.

Sugli effetti. Sul piano osservativo, credo si possa convenire sul fatto che il sistema di valutazione abbia incentivato la competitività e l'individualismo (e in alcuni casi anche l'opportunismo), introducendo elementi di tipo disgregativo all'interno della comunità accademica. Se a questo si aggiungono la frustrazione e la demotivazione diffusi, derivanti dalla sproporzione tra numero di accademici/che abilitati/e e la possibilità di assorbimento del sistema, si comprende come il clima possa diventare esplosivo. Questo, al di là di ogni altra considerazione, non è positivo per la comunità accademica: non lo è dal punto di vista etico-relazionale né dal punto di vista del funzionamento organizzativo, e neanche, per stare alla logica aziendalistica della qualità, dal punto di vista della produttività complessiva.

BIBLIOGRAFIA

Boarelli M. (2019). *Contro l'ideologia del merito*. Roma-Bari: Laterza.

Boring A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145, 27-41.

Chavez K., Mitchell K.M.W. (2020). Exploring bias in student evaluation: Gender, race and ethnicity. *Political Sciences*, *53* (2), 270-274.

Cubelli R. (2015). La psicologia italiana tra scienza della valutazione e valutazione della scienza. *Giornale Italiano di Psicologia*, *42* (4), 647-675.

Di Nuovo S., Cassibba R. (2018). Per una classificazione dei saperi e una «manutenzione» dei sistemi formativi relativi alla psicologia: alcune riflessioni sulle criticità attuali e sulle possibili alternative. *Giornale Italiano di Psicologia*, *45* (2), 265-278.

Gillies D. (2012). Economics and research assessment systems. *Economic Thought*, *1*, 23-47, http://et.worldeconomicsassociation.org/papers/economics-and-research-assessment-systems/.

Feistauer D., Richter T. (2017). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education*, *42* (8), 1263-1279.

MacNell L., Driscoll A., Hunt A.N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, *40* (4), 291-303.

Estratti riservati ai soci AIP, © Società editrice il Mulino

Salvatore S., Pazzagli C., Lucarelli L., Gelo, O. Giannone F. (2015). Livelli e fattori della produttività scientifica della psicologia accademica. Un'analisi bibliometrica. *Giornale Italiano di Psicologia*, *42* (1-2), 263-288.
Tommasi L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.
Uttl B., White C.A., Gonzalez D.W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22-24.

Quality evaluation: In between formal standards, business-like models, compliance and «merit»

Summary. The evaluation systems of the Italian academia raise several conceptual and methodological issues: *a*) as for undergraduate and graduate psychology programs, a conception of «quality» that oscillates between formalism and business-like models, hiding the differences and disparities between universities and territories; *b*) the ability of evaluation systems to drive research practices, fostering compliance and conformity especially among the younger researchers; *c*) an individualistic conception of merit that overlooks the influence of context and in the medium-long run has disruptive impacts on the academic community.

Keywords: Evaluation, formalism, compliance, conformity, meritocracy.

La corrispondenza va inviata a Terri Mannarini, Dipartimento di Storia e Studi sull'Uomo, Università del Salento, Complesso Studium 2000, Edificio 5, Via di Valesio angolo Viale San Nicola, 73100 Lecce. E-mail: terri.mannarini@unisalento.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

LA PSICOLOGIA DEL FUTURO DIPENDE DA NOI?

TOMASO VECCHI[1,2]

[1] *Università di Pavia,* [2] *IRCCS Fondazione Mondino*

Riassunto. Discutere sui meccanismi o i criteri adottati nella valutazione accademica, sia didattica che scientifica, è molto difficile. Un aiuto può venire dalla distinzione tra ciò che è realmente in discussione e ciò che invece è indipendente dai nostri pensieri o desideri. La valutazione accademica, la necessità di criteri efficaci e condivisi a livello internazionale per la valutazione degli atenei, non può essere messa in discussione. Ciò porta necessariamente a riflessioni sui metodi e gli ambiti di studio della psicologia, su quale sia il suo futuro, e quale il rapporto tra la psicologia e le altre discipline. L'ambito di azione di una disciplina deriva, operativamente, dalle sue scoperte, dall'adozione o dalla modifica di un determinato paradigma teorico, dagli avanzamenti tecnologici e dal suo riconoscimento nel contesto socioculturale. È inevitabile e auspicabile che in questa evoluzione vengano ridefiniti in modo flessibile i rapporti tra la psicologia e le discipline tradizionalmente ad essa più vicine così come quelle verso cui l'avvicinamento è avvenuto negli ultimi decenni, come le scienze della vita e le tecnologie dell'informazione.

L'articolo di Luca Tommasi (2020) riassume con grande efficacia tutte le complesse sfaccettature di un meccanismo di valutazione (della ricerca, della didattica, delle strutture) che si è sviluppato in Italia negli ultimi decenni e che è esploso, a partire dal 2010, con l'istituzione dell'ANVUR. Tanto è facile condividere molte delle criticità sollevate dall'autore e riconosciute da molti di noi nei diversi esercizi valutativi in cui siamo stati coinvolti in questi anni, tanto è difficile identificare e proporre soluzioni alternative. L'università è un sistema complesso ed arrivare a meccanismi condivisi tra aree umanistiche, sociali, scientifiche e tecnologiche è quasi impossibile. Questa eterogeneità di approcci si ritrova anche al nostro interno dove differenti aree di ricerca psicologica hanno storie scientifiche diverse, tradizioni professionali con maggiore o minore collegamento con la psicologia accademica, rapporti più o meno consolidati con le discipline affini.

Forse un buon punto di partenza può essere comprendere le valutazioni che sono effettivamente sotto il controllo della nostra comunità accademica e quelle che invece non lo sono. In altre parole, significa capire su quali temi sia importante uno sforzo di riflessione comune, e su quali un'eventuale discussione rimarrebbe sterile, non

Estratti riservati ai soci AIP, © Società editrice il Mulino

potendo in alcun modo modificare il contesto valutativo in cui tutti noi siamo inseriti.

Una prima considerazione riguarda l'importanza della valutazione o, meglio, la necessità stessa di essere valutati. Una ipotesi estrema, ad esempio, suggerisce un ritorno a situazioni in cui la valutazione è lasciata alla responsabilità dei singoli docenti/ricercatori, generando nei fatti una sorta di autovalutazione positiva generalizzata. L'articolo di Tommasi ha messo bene in luce i vantaggi, in termini di incremento di qualità, a cui la valutazione universitaria ha portato in questi anni. Tuttavia, il punto principale, a mio parere, è che la nostra presenza nel contesto universitario internazionale non ci permette di scegliere diversamente. Possiamo ragionare – in parte, e solo in parte – sui criteri di questa valutazione, su come effettuarla, sui parametri più adatti alla realtà italiana, ma non possiamo discutere la necessità di fare o meno una valutazione delle nostre strutture, della nostra didattica e della nostra produzione scientifica. Non è solo un problema di ranking internazionale, che purtroppo mal si adatta alla organizzazione delle università italiane, ma soprattutto di attrattività e rilevanza scientifica che le nostre università vogliono avere in un contesto globalizzato di riferimento. La possibilità di attrarre studenti stranieri dipende in buona parte dalla valutazione che otteniamo nei parametri di valutazione della didattica e da come essi sono percepiti all'estero. La possibilità di attrarre fondi e finanziamenti dipende dall'autorevolezza scientifica del singolo ricercatore ma sempre più anche dall'immagine che i parametri di valutazione danno del suo ateneo. Al tempo stesso, l'accessibilità alla pubblicazione su riviste prestigiose è funzione della qualità percepita dell'ateneo di appartenenza dell'autore. Se vogliamo che il nostro sistema universitario sia in grado di competere in ambito internazionale, l'esercizio della valutazione non può essere messo in discussione. Non accettare il principio della valutazione tra pari, della presenza di contesti di maggiore o minore qualità, della necessità di mettere in atto in modo continuativo e consapevole politiche di miglioramento della qualità reale e percepita dei nostri corsi di laurea e della nostra ricerca, porta necessariamente alla esclusione dal sistema universitario internazionale, alla impossibilità di reperire fondi di ricerca o di attrarre buoni ricercatori. La valutazione non è una scelta: la scelta è solo sui criteri con cui viene effettuata, sulla modalità che scegliamo di darci, sempre consapevoli che cedimenti sul piano delle qualità e dei criteri universalmente riconosciuti, porterebbe comunque a una lenta implosione del nostro sistema universitario.

Una seconda considerazione che si è posta con forza durante questi anni di valutazione, soprattutto attraverso la VQR, riguarda più nello specifico la realtà della psicologia italiana. Molti di noi ricordano la discussione, anche aspra e poco produttiva, che vi è stata

Estratti riservati ai soci AIP, © Società editrice il Mulino

nella comunità accademica riguardo all'inserimento della psicologia tra le discipline bibliometriche. Vi era realmente la possibilità di scegliere? La psicologia poteva rimanere in aree non bibliometriche? Queste domande hanno portato necessariamente a riflessioni sui metodi e gli ambiti di studio della psicologia, su quale sia il suo futuro, e quale il rapporto tra la psicologia e le altre discipline. Seppur interessanti, tali discussioni spesso non hanno tenuto conto di quello che a mio parere era il punto cardine: non siamo noi a decidere il ruolo della psicologia, a definirne il futuro o le aree di pertinenza rispetto alle discipline affini. Come spesso accade, l'ambito di azione di una disciplina deriva, operativamente, dalle sue scoperte, dall'adozione o dalla modifica di un determinato paradigma teorico, dagli avanzamenti tecnologici e dal suo riconoscimento nel contesto socioculturale. È inevitabile e auspicabile che in questa evoluzione vengano ridefiniti in modo flessibile i rapporti tra la psicologia e le discipline tradizionalmente ad essa più vicine così come quelle verso cui l'avvicinamento è avvenuto negli ultimi decenni, come le scienze della vita e le tecnologie dell'informazione (ICT).

Le critiche alla valutazione bibliometrica sono fin troppo facili: l'impossibilità di un confronto tra discipline diverse ma anche tra ambiti distinti all'interno della stessa disciplina; la proliferazione di autocitazioni e di una produttività fine a se stessa; l'aumento esponenziale del numero medio di autori nelle pubblicazioni di alcuni settori; l'impossibilità di valutare correttamente il lavoro di un singolo autore, soprattutto nel caso dei ricercatori più giovani; le criticità legate alla valutazione della forza di un gruppo e dei fondi a disposizione e non della qualità del singolo ricercatore; l'obiettivo, soprattutto da parte dei ricercatori più anziani, di un continuo innalzamento dei parametri citazionali che rende di fatto impossibile una credibile progressione dei giovani; la credenza che variazioni negli indici bibliometrici portino necessariamente a differenze nella qualità dei ricercatori; l'impatto sugli indici bibliometrici dell'avvento delle riviste *open access*; l'utilizzo esclusivo di questi parametri non associati ad altri di reale valutazione individuale; l'idea stessa che dalla semplice valutazione bibliometrica possa derivare una valutazione complessiva di idoneità accademica o scientifica, senza considerare l'impatto che l'acquisizione di impegni didattici e/o incarichi organizzativi e istituzionali hanno sulla produttività scientifica degli individui.

L'adozione di una prospettiva bibliometrica nazionale spesso conduce anche a errori nella percezione stessa del valore delle singole realtà e dei singoli ricercatori. Se non viene fatto un costante confronto con i parametri di ricerca internazionali, ma solo con i valori di riferimento italiani, si può arrivare a credere che tutte le aree psicologiche abbiano una uguale maturità scientifica. Si può pensare che il miglior

Estratti riservati ai soci AIP, © Società editrice il Mulino

ricercatore in un dato ambito rappresenti il medesimo livello di competitività del più qualificato ricercatore in un ambito apparentemente vicino. Non è così. Sarebbe come paragonare la nazionale italiana di calcio con quella di baseball e suggerire che la qualità internazionale di entrambe sia paragonabile. Il finanziamento delle diverse aree, sulla base di parametri di confronto unicamente nazionali va a supportare tra loro scopi molto diversi. In un caso si premia una produzione di eccellenza che va mantenuta, nell'altro caso si sostiene un settore scientificamente debole con l'obiettivo di farlo crescere e sviluppare su standard internazionali condivisi. Sono politiche del tutto diverse che non è semplice seguire con i medesimi criteri di valutazione.

Se ne potrebbero aggiungere molte altre, ma bastano queste gravi criticità a mettere in dubbio l'importanza dei parametri bibliometrici? Io credo di no. Non lo credo soprattutto perché in Italia siamo stati tra gli ultimi ad adottare questi parametri, a riconoscerne non solo i limiti ma anche il loro valore intrinseco che si associa ad un miglioramento della qualità scientifica di una comunità. Nella realtà italiana, questo valore non è ancora unanimemente condiviso rendendo necessaria una spinta obbligata verso il riconoscimento dei parametri bibliometrici quale valore internazionale di riferimento. Tali parametri, possono essere corretti, bilanciati, adattati alla luce dell'esperienza dei contesti dove questi metodi sono adottati da decenni ma non ancora abbandonati.

L'obiettivo di questo breve commento era di mettere in luce come i vincoli e i limiti decisionali del processo valutativo ci vengono dalla realtà internazionale, dalla definizione condivisa della disciplina e dai confini che essa ha: questi fattori non dipendono dalle opinioni di ognuno ma riflettono un contesto internazionale da cui siamo tutti, in egual modo, dipendenti. Comprendere la psicologia nella sua evoluzione globale e nei rapporti con le altre discipline è sicuramente difficile, non solo rispetto alla situazione attuale ma anche volendo fare un esercizio di previsione per i prossimi decenni. Cosa sarà la psicologia nel 2030 o nel 2050? I giovani che formiamo, i nostri studenti, si confronteranno con il contesto lavorativo in un arco temporale compreso tra il 2030 e il 2070. La psicologia del futuro è già parte del nostro lavoro, dato che decidiamo cosa insegnare ai giovani che stiamo formando. Il tema esula dal contesto della valutazione anche se ne diventa parte nel momento in cui dobbiamo dare il giusto valore a come la psicologia si rapporta con le altre discipline. La psicologia potrà avere un ruolo guida o sarà necessariamente marginale nelle direzioni che le ricerche sul comportamento avranno nel prossimo futuro? Questo tema si è spesso discusso analizzando il rapporto tra psicologia e neuroscienze, forse anche per la presenza al nostro interno di ricercatori che avevano pregiudizialmente una fascinazione/avversione per alcuni metodi o tematiche di ricerca. Difficile essere obiettivi in queste condi-

Estratti riservati ai soci AIP, © Società editrice il Mulino

zioni tanto che, solo per fare un esempio, proposte didattiche che prevedono una sostanziale esclusione del ruolo della psicologia dall'ambito delle neuroscienze hanno trovato un consenso, seppur minimo, all'interno della nostra comunità. Propongo allora una riflessione su di un ambito che sta avendo un grande sviluppo e che non è entrato finora esplicitamente nei nostri corsi di laurea. Uno dei due progetti *flagship* europei finanziati molto cospicuamente dalla EU per almeno un decennio – *Human Brain Project* – riguarda la simulazione del comportamento umano attraverso modelli computazionali. Si studia il comportamento, le sue basi neurali, la plasticità cognitiva, l'impatto delle emozioni e dei processi motivazionali, insieme naturalmente ad ambiti più strettamente informatici, ingegneristici e di *data science*. Il progetto utilizza per buona parte metodi, strumenti e conoscenze che sono parte del bagaglio della psicologia per come si è definito negli ultimi decenni: la psicologia italiana ha però in questa rete di ricerca un ruolo del tutto marginale. Indipendentemente dal valore di questo progetto, e dalla qualità di progetti simili in America e Cina, credo sia facile prevedere uno sviluppo esponenziale di questi temi e la possibilità per la psicologia di contribuire alle conoscenze che sono o saranno necessarie.

È una nostra decisione, e il frutto della nostra capacità previsionale, lasciare o meno questi promettenti ambiti di ricerca ai fisiologi ed ai biologi (nel caso delle neuroscienze) o agli ingegneri e agli informatici (nel caso dei modelli computazionali). Siamo davvero noi a decidere dove va la psicologia, quali saranno i parametri di valutazione e i metodi di finanziamento dei prossimi decenni? La nostra unica scelta è indirizzare le nostre riflessioni alla psicologia del futuro, affiancando ai rapporti con le discipline tradizionalmente affini – quali ad esempio le scienze sociali e dell'educazione – quelli con la biologia e la biochimica, con la medicina e le scienze della vita, con la fisica, l'ingegneria e la bioetica. Se le nostre previsioni sul futuro della psicologia saranno giuste, i criteri più corretti di valutazione del nostro lavoro emergeranno automaticamente.

BIBLIOGRAFIA

TOMMASI L. (2020). La valutazione della psicologia: considerazioni su un decennio turbolento. *Giornale Italiano di Psicologia*, *47* (1), 13-47.

Is the future of psychology open to discussion?

Summary. The assessment of teaching and scientific quality for universities and individuals is a very hard task. A first step is to recognize that no one is in the position to reject evaluation policies. Further, the adoption of internationally recognized indicators

Estratti riservati ai soci AIP, © Società editrice il Mulino

is a compelling condition for maintaining Italian psychology within the scientific community. Similarly, the definition of psychology, its methods, its relationships with other disciplines, are not defined by individual wishes or personal approaches. The present and the future of psychology are determined by the evolution of the discipline and its relationships with emergent areas of research such as life and data science, information technologies and bioethics. The adoption of the best indicators for assessing psychological teaching and research is a direct consequence of our capacity to understand what psychology is now and, above all, what psychology will be in the next decades.

Keywords: Assessment, scientific metrics, interdisciplinarity, scientific evaluation.

La corrispondenza va inviata a Tomaso Vecchi, Dipartimento di Scienze del Sistema Nervoso e del Comportamento, Università di Pavia, Piazza Botta 11, 27100 Pavia. E-mail: vecchi@unipv.it

Estratti riservati ai soci AIP, © Società editrice il Mulino

Giornale italiano di psicologia

Estratti

ISSN 0390-5349

ISBN 978-88-15-36909-3