

Per una “teoria del cambiamento”: riflessioni a margine della valutazione delle iniziative di Valorizzazione delle conoscenze degli Atenei Italiani

*Mauro Palumbo**, *Anna Fausta Scardigno***, *Antonio Felice Uricchio****

The text critically analyses the evaluation of the Third Mission and Knowledge Valorisation in Italian universities, with a specific focus on the VQR 2020-2024 exercise. The authors highlight differences between AN-VUR guidelines and the international scientific literature, emphasizing how terms such as impact and added value are often used ambiguously. To overcome a purely bureaucratic logic, the adoption of a theory of change that promotes institutional learning and reflexivity is proposed. The goal is to transform evaluation from a simple formal compliance into a strategic tool to strengthen the social and civic role of the Universities. A concentric circles model is also suggested to better monitor the effects of interventions on territories and beneficiaries. In summary, the essay promotes a vision of evaluation understood as a transformative process capable of generating a real improvement in the way by which Universities develop their civic and social responsibility.

Keywords: third mission; civic universities; evaluation; social responsibility of universities; quality evaluation of research.

Introduzione

La Terza Missione (d’ora in poi TM) che nell’ultimo esercizio VQR Anvur 2020-2024 è stata definita come Valorizzazione delle conoscenze (d’ora in poi VdC), è anche in Italia oggetto di riflessione e analisi scientifica,

* Mauro PALUMBO, Università di Genova (palumbo@unige.it)

** Anna Fausta SCARDIGNO, Università di Bari Aldo Moro (annafausta.scardigno@uniba.it)

*** Antonio Felice URICCHIO, Università di Bari Aldo Moro (antonio.uricchio@anvur.it)

Invio Proposta: 19-05-2026. Accettazione: 14-07-2026.

sempre più fondata non solo su considerazioni teoriche, ma anche su ricerche empiriche (Fontana e Valentini 2020, Reale, a cura di, 2025, Boffo, Moscati e Rostan, 2025, Palumbo, Scamuzzi e Borgna, 2025), spesso basate sulle esperienze maturate da parte dell’Agenzia nazionale di valutazione dell’Università e della ricerca (Celletti e Uricchio, 2025). L’Agenzia, infatti, ai sensi dell’Art. 3, comma 1 del DPR 76/2010, come modificato dal DM n. 12/2026, valuta la qualità complessiva delle attività didattiche, di ricerca e, anche su richiesta del Ministero, di valorizzazione delle conoscenze, ivi inclusa la terza missione, delle Università, delle istituzioni AFAM e degli enti di ricerca.¹

Questa specifica attenzione è almeno in parte ascrivibile all’inclusione dei casi di TM nei periodici esercizi valutativi dell’ANVUR, ma ci piace pensare che sia anche dovuta alla crescente consapevolezza da parte degli Atenei dell’importanza del loro contributo all’evoluzione sociale, economica e culturale dei territori sui quali insistono. In questo senso, anche il progressivo ampliamento delle tematiche e dei campi d’azione riconducibili alla TM che si è verificato tra i primi esercizi valutativi ANVUR e quello appena concluso ci pare testimoni la crescente importanza assegnata alle attività degli Atenei finalizzate all’accrescimento della coscienza civica e al rafforzamento della democrazia, come emerge dalla valorizzazione delle iniziative di social e civic engagement e di capacity building.

Mentre i primi due esercizi condotti dall’ANVUR hanno avuto un intento ricognitivo, anche al fine di ottenere una classificazione adeguata delle attività di TM e per cercare di arrivare a possibili indicatori valutativi comuni, nel caso in esame si è mantenuto il ricorso al metodo dei “casi di studio” scelti dagli Atenei, modificando tuttavia la denominazione dell’oggetto (non più Terza Missione, ma appunto di Valorizzazione delle conoscenze) ed ampliando la tipologia proposta, che raggiunge i venticinque campi d’azione. Non trascurabile poi il fatto dell’esplicita funzione premiale dell’attività valutativa.

¹ Con l’introduzione del sistema di Autovalutazione, Valutazione Periodica e Accreditamento (AVA), la qualità della Terza Missione viene introdotta tra i requisiti di qualità delle Sedi e dei Corsi di Studio (allegato C del DM 987/2016) e vengono identificati indicatori e parametri di valutazione della Terza Missione molto più fini e dettagliati, inerenti non solo la quantità di attività svolte, ma anche alcuni aspetti legati alla qualità e all’impatto di dette attività (allegato E del DM 47/2013). AVA, inoltre, dà avvio alla creazione di una base informativa unica (art. 9 Decreto Ministeriale 987/2016, già previsto nel DM 47/2013 e ss.mm.ii., in particolare Allegato E) sulla qualità della ricerca dipartimentale, la Scheda Unica Annuale della Ricerca Dipartimentale SUA-RD, insieme a quella dedicata ai Corsi di Studio e all’offerta formativa, la SUA-CdS.

Il passaggio semantico-concettuale dalla “Terza Missione” alla “Valorizzazione delle conoscenze” non sembra privo di effetti. Il primo concetto/termine aveva infatti un punto di forza nella residualità definitoria (analogamente al concetto/termine di settore terziario nell’economia) e questo aveva permesso ad ANVUR di identificare, nel bando TM precedente, due distinte “famiglie” di attività, denominate Valorizzazione della ricerca e Produzione di beni pubblici. Di contro, il concetto/termine “valorizzazione delle conoscenze”, utilizzato nell’esercizio più recente, sembra ricalcare l’approccio della RACCOMANDAZIONE (UE) 2022/2415 DEL CONSIGLIO del 2 dicembre 2022 sui principi guida per la valorizzazione delle conoscenze, che intende la «Valorizzazione delle conoscenze» come «il processo con cui si crea valore sociale ed economico a partire dalle conoscenze, collegando aree e settori diversi e trasformando i dati, le competenze tecniche e i risultati della ricerca in prodotti, servizi, soluzioni e politiche basate sulle conoscenze sostenibili che portano vantaggi alla società» (p. 143).

ANVUR si è infilata, in questo modo, in una camicia di Nesso, perché le definizioni della Raccomandazione Europea si sforzano, a nostro avviso senza un grande successo, di estendere ad attività svolte in collaborazione con/ o a beneficio della società civile un modello di “trasferimento” che mal si presta alla bisogna. Erogare formazione nei Poli Carcerari, favorire l’empowerment dei cittadini, svolgere attività a forte contenuto sociale, rientrano tutte in una concezione cooperativa o di servizio in cui dei beni pubblici sono co-prodotti dalle persone che lavorano negli Atenei e dai destinatari e solo con una (inutile) torsione semantico concettuale possono essere viste come valorizzazione di prodotti di ricerca². In breve, l’esame dei 25 campi di azione interni alle cinque aree tematiche consente di affermare che almeno una metà di questi sia lontana da una pur ampia concezione di trasferimento tecnologico, richiamando piuttosto la prospettiva della *civic university*, ovvero le attività precedentemente raggruppate sotto l’etichetta “Produzione di beni pubblici”.

² Infatti, nei “considerando” della raccomandazione troviamo acrobazie concettuali quali “La valorizzazione delle conoscenze è un cambiamento di paradigma che introduce nuovi aspetti in grado di massimizzare il valore della R&I esistenti e future e del capitale di conoscenze, compresa la conoscenza tacita” (16) ... «Le pratiche, i processi, le capacità e le competenze imprenditoriali e quelli che facilitano il dialogo con i cittadini, la società civile e la classe politica sono elementi cruciali per il successo delle iniziative di valorizzazione delle conoscenze. Per trasformare la conoscenza in valore aggiunto, che si tratti di innovazioni progressive o dirompenti, di politiche basate su dati concreti o del benessere dei cittadini, occorrono culture, pratiche o atteggiamenti proattivi/ intraprendenti e di co-creazione/impegno intersettoriale, combinati con sforzi imprenditoriali in alcune o in tutte le fasi del processo di valorizzazione.» (18)

1. L'Apprendimento Permanente come Campo di Valut-Azione

In questo saggio cerchiamo di evidenziare alcuni punti di criticità degli esercizi di valutazione anche in virtù del coinvolgimento degli stessi autori nel processo valutativo. In particolare, l'esperienza maturata come Presidente dell'ANVUR ha consentito, non solo di sovrintendere l'intero processo valutativo, ma soprattutto di mettere a fuoco il processo valutativo anche attraverso il confronto costante con i componenti del GEV – Gruppi di Esperti della Valutazione, le istituzioni valutate e le comunità scientifiche. Particolare rilevanza anche il contributo di Anna Fausta Scardigno che ha operato all'interno del GEV TM sia nel terzo che nel quarto esercizio della VQR e di Mauro Palumbo, impegnato scientificamente e nel proprio Ateneo sui temi della valorizzazione delle conoscenze. Da un lato, questo può evidentemente inficiare la “terzietà” delle nostre posizioni, ma dall'altro è in piena armonia con la posizione di autori che sottolineano la tensione etica e valoriale che anima il valutatore e il suo orientamento al miglioramento delle politiche (es. House e Howe, 1999; Patton, 2011; Fetterman, Kaftarian e Wandersman, 2015).

Come è noto, a livello europeo i criteri ESG (Environmental, Social and Governance) forniscono anche i criteri di valutazione delle agenzie di assicurazione della qualità e delle loro attività. Ciò garantisce che le agenzie di assicurazione della qualità aderenti all'EHEA (European Higher Education Area) aderiscano allo stesso insieme di principi, mentre processi e procedure sono organizzati in modo tale da adattarsi agli obiettivi e alle esigenze dei contesti in cui esse operano.

Come chiarito nel documento approvato nel 2015³, gli ESG hanno i seguenti obiettivi: 1) Definiscono un quadro comune per i sistemi di assicurazione della qualità dell'apprendimento e dell'insegnamento, a livello europeo, nazionale ed istituzionale. 2) Rendono possibile l'assicurazione ed il miglioramento della qualità dell'istruzione superiore nello Spazio Europeo dell'Istruzione Superiore. 3) Promuovono la fiducia reciproca, facilitando così il riconoscimento e la mobilità all'interno dei singoli Paesi e fra Paesi diversi. 4) Offrono informazioni in merito all'assicurazione della qualità nell'European higher education area.

³ European Association for Quality Assurance in Higher Education (ENQA), European Students' Union (ESU), European University Association (EUA), European Association of Institutions in Higher Education (EURASHE). In cooperation with: Education International (EI), BUSINESSEUROPE, European Quality Assurance Register for Higher Education (EQAR Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG). (2015). Brussels, Belgium.

Gli ESG sono fondati sui seguenti quattro principi di assicurazione della qualità nell'EHEA: 1) Le Istituzioni di istruzione superiore hanno la responsabilità primaria della qualità dei servizi didattici offerti e della sua assicurazione. 2) L'assicurazione della qualità tiene conto della diversità dei sistemi di istruzione superiore, delle Istituzioni, dei corsi di studio e degli studenti. 3) L'assicurazione della qualità promuove lo sviluppo di cultura della qualità. 4) L'assicurazione della qualità tiene conto delle esigenze e delle aspettative degli studenti, degli altri portatori di interesse e della società in generale.

Obiettivi e principi forniscono un quadro comune di riferimento entro il quale gli ESG possono esser utilizzati da parte di diverse Istituzioni, agenzie e Paesi, anche avendo riguardo ai propri modelli regolatori e organizzatori. L'EHEA è, peraltro, caratterizzato dalla sua diversità quanto a sistemi politici, sistemi di istruzione superiore, tradizioni socioculturali ed educative, lingue, aspirazioni ed aspettative. Questo rende del tutto inappropriato un approccio monolitico alla qualità ed all'assicurazione della qualità nell'istruzione superiore. Un'ampia accettazione di tutti gli standard è la condizione essenziale per creare una visione comune dell'assicurazione della qualità in Europa. Per questi motivi, gli ESG devono mantenersi ragionevolmente ad un livello di principi tale da garantire l'applicabilità degli standard e delle linee guida a tutte le tipologie di servizi educativi offerti. Anche il confronto in atto sulla revisione degli ESG (che dovrebbe concludersi nel 2027) non sposta il modello primigenio, venendo via via arricchito anche alla luce dell'evoluzione tecnologica e delle esperienze all'interno dei ENQA, EQAR e delle reti di agenzie extraeuropee.

Proprio nel corso dell'ultima Conferenza ministeriale di Tirana, intitolata significativamente "Evolution not revolution", di maggio 2024, le parti interessate hanno concordato che gli ESG non dovrebbero essere sovraccaricati, restando snelli, pur se adeguati ai cambiamenti intervenuti. Sulla base del lavoro preparatorio e dell'analisi di dieci anni di attuazione dell'ESG 2015, la revisione si è concentrata sull'aggiornamento del documento al contesto attuale, compreso l'allineamento con le altre politiche del Processo di Bologna e la riflessione sul contesto sociale e tecnologico in continua evoluzione.

Il protocollo di intesa tra Rete delle Università Italiane per l'Apprendimento (d'ora in poi RUIAP), l'Associazione Italiana di Valutazione (AIV) e ANVUR⁴ ha permesso di avviare un primo confronto, a partire dal gruppo di

⁴ Siglato il 29 maggio 2023, il Protocollo ha come obiettivo principale lo sviluppo di strategie valutative relative all'Apprendimento Permanente negli Atenei italiani. Le parti si impegnano a sviluppare iniziative di studio e ricerca sull'analisi degli effetti delle attività valutative, di

lavoro RUIAP sulla valutazione dell'apprendimento permanente in chiave di TM, sia di analisi dell'esercizio VQR TM 2015-2019, che su quello appena concluso (2020-2024). In particolare, si è lavorato per quattro obiettivi:

- migliorare la definizione dei confini dell'Apprendimento Permanente (AP),
- costruire domande riflessive sui criteri valutativi,
- sviluppare indicatori comuni per la valutazione,
- ricostruire a partire da casi critici “teorie del cambiamento” e “teorie del programma” che possano essere da guida nel prossimo esercizio valutativo e non solo.

I risultati relativi ai primi tre obiettivi sono stati pubblicati in un numero monografico di *EPale Journal* 2024⁵ e presentati anche in sede europea (Palumbo e Piazza, 2023). Qui proseguiamo la riflessione su criteri e indicatori di valutazione in parte anticipata altrove (Brambilla e Palumbo, 2024) e sulle possibili lezioni apprese dai due esercizi valutativi oltre che sulle nuove domande che essi pongono. In primo luogo, un fondamentale interrogativo riguarda il tema della Terza Missione o della Valorizzazione delle conoscenze in rapporto a quello più ampio della responsabilità sociale delle Università in termini di divulgazione dell'impatto delle loro attività sulla società.

La risposta non può prescindere dalla rilevanza degli Atenei come agenti di cambiamento democratico e di innovazione sociale; la letteratura più recente sembra orientata proprio in questo senso (cfr. ad es. Charles et al., 2026) e in generale da tempo analizza la cosiddetta Terza Missione in chiave di civic university (Scardigno et al., 2024; Palumbo e Scardigno, 2024) considerando quindi non solo l'impatto complessivo sul territorio e sulla società, ma anche la dimensione vocazionale del servizio reso ad esse (Goddard et al., 2021; Hazelkorn, 2020).

In secondo luogo, la riflessione si vuole soffermare sull'aspetto metodologico e in particolare su come tradurre l'impianto valutativo promosso da ANVUR, fondato sui quattro criteri del valore aggiunto, dell'impatto, dell'influenza sul contesto e del contributo della struttura proponente, in percorsi metodologici praticabili, condivisi e utili per la crescita della consapevolezza degli Atenei della rilevanza della loro presenza sul territorio. Come anticipato, i macrocriteri già definiti da ANVUR sono stati i punti di riferimento delle

predisposizione di modelli e strumenti di impatto delle attività universitarie, nella prospettiva della valorizzazione dell'impatto dell'AP nel sistema universitario e della ricerca, <https://www.ruiap.it/protocollo-di-intesa-anvur-aiv-ruiap/>

⁵ <https://www.indire.it/wp-content/uploads/2020/05/EPAL-JOURNAL-N-14-2023.pdf>

istituzioni per costruire domande valutative sugli impatti generati. E sono altresì coerenti con le prospettive delle civic universities sopra ricordate.

In questo contributo ci pare tuttavia opportuno segnalare la discordanza tra il dettaglio dei criteri valutativi contenuto nelle linee guida del GEV ANVUR (2024) e le definizioni correnti in letteratura (segnatamente di impatto e valore aggiunto), recentemente segnalate su questa Rivista (cfr. Carrera et al., 2025; ma si veda anche Brambilla e Palumbo, 2024).

Questa discordanza ha a nostro avviso creato problemi nella descrizione dei casi di studio da parte degli Atenei, privi di un riferimento puntuale cui affidarsi, e sembra anche inficiare il valore di apprendimento della valutazione e favorire invece un approccio “adempimentale”, pietra tombale di ambizioni migliorative della VQR⁶.

Il tema sarà ripreso nelle considerazioni conclusive, ma in questa sede è bene ricordare che delle due funzioni principali della valutazione, quella di accountability e quella di learning, viene di solito privilegiata la seconda e anche l’ANVUR si presenta come istituzionalmente deputata a favorire il miglioramento delle attività degli Atenei.

Questa ambizione confligge in verità innanzi tutto con la funzione essenzialmente premiale della valutazione, che di norma favorisce comportamenti adattivi ed “opportunistici”, non necessariamente seguiti da miglioramenti reali.

Questo vale sia per il ranking tra Atenei che deriva dalle valutazioni periodiche (e incide quindi sulla loro reputazione), sia per la quota di Fondo di Finanziamento Ordinario che viene ripartita sulla base di tale ranking. Le due cose non sono necessariamente connesse, ma riteniamo che una metodologia valutativa maggiormente riconducibile a modelli riconosciuti in letteratura avrebbe potuto favorire un apprendimento almeno metodologico da parte degli Atenei, al di là dell’adempimentalità opportunistica dalla funzione premiale.

Importante è quindi soffermarsi su come gli Atenei, nei processi di selezione dei casi e della loro descrizione, abbiano potuto tener conto delle Linee Guida e dell’applicazione delle stesse a cura del GEV TM, sottolineando l’importanza della valutazione come elemento di arricchimento della capacità riflessiva degli Atenei e delle loro strategie di raccordo con il territorio, anche attraverso un confronto tra le “logiche valutative” seguite dai diversi protagonisti della VQR/TM.

⁶ Il citato studio (Carrera et al., 2025), condotto sulla passata VQR/TM, evidenzia, in un campione di casi, che anche per quel che riguarda la scelta degli indicatori di impatto gli Atenei hanno optato per quelli proposti di ANVUR e che solo in rari casi hanno proposto indicatori diversi.

Potrebbe essere utile piuttosto (come si riprenderà nelle conclusioni del paper) un supporto nella fase valutativa curata dai GEV al processo riflessivo e autovalutativo (questa volta *ex ante* rispetto al secondo esercizio ed *ex post* rispetto al primo), che sarebbe sicuramente più efficace e condiviso se sostenuto da un’impostazione comune ai vari ambiti tematici e di agevole e flessibile applicazione. Com’è noto, nella VQR/TM (2020-24) sono stati mantenuti dal Bando i quattro criteri del precedente esercizio, con pochissime modifiche⁷, fatta salva l’esplosione dei campi di azione, che ha sicuramente accentuato l’attenzione verso la dimensione “civica” e dell’impatto sociale degli Atenei. Infatti, dei 25 campi di azione (5 per ogni area tematica), ben 15, corrispondenti a Produzione e gestione di beni pubblici (Area II), Public Engagement (Area IV) e Sostenibilità ambientale, inclusione e contrasto alle disuguaglianze (Area V) sono riconducibili a questa dimensione e anche alcuni campi d’azione che rientrano nelle due aree I e IV, come il campo e) dell’area I (Attività di formazione per promuovere la cultura dell’innovazione) sono fortemente connotati in questo senso.

Le osservazioni che seguono si appuntano in particolare sui concetti di impatto e di valore aggiunto utilizzati dal GEV ANVUR e ulteriormente dettagliati nelle singole aree tematiche e talvolta anche nei singoli campi d’azione.

In linea generale, con il termine impatto ci si riferisce «agli effetti a lungo termine, positivi e negativi, primari e secondari, previsti o imprevisti, prodotti direttamente o indirettamente da un intervento di sviluppo» (OECD, 2024 p. 17). Come ha osservato Stern (2016, p. 21), «i decisori politici hanno spostato l’attenzione da un obiettivo della valutazione principalmente orientato all’accountability a uno che considera prioritario anche l’apprendimento». Di qui il crescente interesse sul perché e sul come i programmi funzionino o meno, questioni appunto cruciali per il loro miglioramento.

Inoltre, molto spesso il concetto di impatto viene inserito in una sequenza logica, ispirata al Project Cycle Management (d’ora in avanti PCM), adottato stabilmente dalla Commissione europea e dalle principali agenzie di valutazione, che distingue tra: Input – risorse impiegate, Attività – azioni realizzate, Output – prodotti immediati (beni/servizi erogati), Risultati (outcome / outreach) – cambiamenti diretti sui destinatari, Impatto (impact) – effetti di

⁷ Ai sensi del Art. 9, comma 7, del Decreto n. 8 del 31/10/2023 del Presidente ANVUR si tratta di:

- dimensione sociale, economica e culturale dell’impatto, considerando la capacità di valorizzare le conoscenze anche collegando aree e settori diversi;
- rilevanza rispetto al contesto di riferimento;
- valore aggiunto per i beneficiari;
- contributo scientifico, organizzativo e/o gestionale della struttura proponente.

medio-lungo periodo sul contesto (cfr. European Commission, 2024; Huemann e Turner, R., eds., 2024; White, 2009).

Questo approccio, che ha ispirato il lavoro del gruppo “Valutazione dell’AP” della RUIAP, ripreso in Brambilla e Palumbo (2024), permette di affermare che il valore aggiunto per i destinatari andrebbe collocato al livello dei risultati /outcome, e dovrebbe misurare il cambiamento conseguito sui destinatari dell’intervento causalmente attribuibile a questo, mentre l’impatto riguarda le trasformazioni di medio-lungo periodo generate dagli interventi nei confronti dei destinatari e del contesto.

Naturalmente anche questo approccio ha dei limiti, innanzitutto perché presuppone nessi di causalità lineare che, come detto, non sono propri del mondo sociale in cui operano le attività di TM e, in secondo luogo, perché rimane “guidato dagli obiettivi” del decisore, mentre molti interventi di TM ricostruiscono i loro obiettivi a posteriori e dunque sono per entrambe le ragioni poco adatti al modello. Questo ha tuttavia una preliminare funzione di chiarificazione concettuale ed operativa per l’affollato mondo dei costruttori di azioni di TM e può quindi essere usato, in prima approssimazione, sia in sede di progettazione che di valutazione.

Confrontando con queste definizioni le declinazioni che ne ha dato il GEV TM (cfr. ANVUR, 2024), si può constatare che il consenso diffuso circa l’utilità di questi criteri rischia poi di ridursi per declinazioni degli stessi piuttosto lontane da quelle correnti. Nel PCM, separare risultati e impatti serve a distinguere benefici immediati da trasformazioni strutturali; problematizzare la sostenibilità nel tempo delle trasformazioni indotte e riflettere sui meccanismi di cambiamento: tutti aspetti di grande importanza in una concezione apprenditiva e trasformativa della valutazione (Weiss, 1997/2007; Patton, 2011).

Nell’approccio del GEV, l’impatto assume una valenza descrittiva del caso e viene articolato ed esemplificato con il rischio di fargli perdere il significato originale. Le specificazioni dei criteri nei diversi campi d’azione finiscono per complicare il quadro di riferimento per chi compila le schede dei casi di studio e per far perder loro di vista la necessità di dimostrare, se non un’attribuzione o un nesso causale, almeno che l’azione abbia messo in moto alcuni “meccanismi” che hanno prodotto certi “effetti” in un determinato “contesto”.

Questo per i casi “di successo” presentati dagli Atenei, ma come è noto il metodo funziona anche per capire quando e perché il meccanismo non ha funzionato, con effetti di apprendimento anche dai fallimenti (non riportati magari all’ANVUR, ma presenti nella realtà).

Già nella definizione iniziale delle Linee Guida del GEV, l’impatto si stempera in varie dimensioni non tutte coerenti con il significato che assume nella letteratura valutativa: «Criterio a) Dimensione sociale, economica e

culturale dell’impatto, considerando la capacità di valorizzare le conoscenze anche collegando aree e settori diversi.⁸ Tale criterio valuta se le attività presentate nel caso studio siano state capaci di avere impatto, come precedentemente definito, sociale, economico e culturale. Tale criterio inoltre apprezza la capacità del caso studio di collegare, in modo interdisciplinare e multidisciplinare, aree e settori diversi all’interno delle dimensioni sociali, economiche, culturali, della salute, istituzionali, etc. su cui è intervenuto.» (GEV ANVUR TM, 2024, p. 12).

Nel dettaglio delle Tematiche II e III, in particolare, le specificazioni e le digressioni sembrano far premio sulle qualificazioni degli impatti. Inoltre, la concezione di impatto proposta dal GEV considera doverosamente l’orizzonte temporale, coerente con quello previsto dall’esercizio valutativo come da bando, ma non sempre praticabile per interventi relativamente “recenti” (gli impatti di solito richiedono tempo per manifestarsi, mentre in un arco temporale breve sono rilevabili solo i risultati); inoltre, la definizione operativa del GEV sembra puntare quasi esclusivamente agli effetti coerenti con gli obiettivi dichiarati e non ha potuto considerare gli effetti eventuali (e inattesi) degli interventi, che possono sia controbilanciare quelli attesi, se di carattere negativo, sia evidenziare le “benedizioni nascoste”, se di carattere positivo, come messo in luce da Hirshman (Stame, 2015).

Dunque, la VQR/TM sembra utilizzare il termine impatto in modo semanticamente non sempre conforme rispetto allo standard internazionale, secondo il modello adempimentale proprio dell’esercizio periodico della VQR-TM. Ma su questo torneremo tra breve.

Del pari, anche il criterio del valore aggiunto sembra tener conto non solo dei destinatari diretti degli interventi, ma di una platea di crescente ampiezza di soggetti, che fa per l’appunto sfumare il concetto di valore aggiunto in quello di impatto, mentre nella letteratura valutativa (cfr. Scriven, 1991) il valore aggiunto rappresenta il contributo addizionale prodotto dall’intervento e richiama talvolta analisi di tipo controfattuale (e sotto questo aspetto non coincide con i risultati dello schema PCM, come forse sarebbe invece più logico attendersi dall’impianto della VQR).

Preme ancora sottolineare che nei documenti metodologici del GEV impatto, valore aggiunto e rilevanza rispetto al contesto risultano frequentemente sovrapposti, come si evince anche dal fatto che gli stessi indicatori proposti vengono utilizzati per più criteri e tendono quindi a essere trattati come concetti fungibili, con una ridondanza descrittiva e una compressione delle dimensioni valutative che, secondo Scriven, riduce la qualità analitica del giudizio.

⁸ Qui è evidente il rimando alle Raccomandazioni del Consiglio Europeo.

Infine, forse anche per l'esigenza prioritariamente descrittivo-ricognitiva dell'esercizio valutativo, si ravvisa l'assenza, nei documenti del GEV, della richiesta di esplicitare una qualche teoria del programma, ovvero una qualche spiegazione del perché l'intervento A, magari sotto determinate condizioni, ha prodotto l'effetto B e questo "autonomizza" risultati e impatti come elementi di giudizio invece che come passi di una sequenza di produzione del cambiamento. Non è dunque richiesta una teoria causale sottostante, sia di tipo "forte" che di tipo "debole", ossia né riconducibile alla Theory Based Evaluation (Chen, 2005), né alla Contribution Analysis (Mayne, 2001; Leeuw, 2023). Di contro, concetti come impatto e valore aggiunto, nella letteratura valutativa internazionale (OECD, 2024; Rossi, Lipsey e Freeman, 2018; White, 2009; Mayne, 2001), costituiscono categorie analitiche orientate all'analisi causale e all'apprendimento, più che semplici descrittori di effetti osservabili nel breve periodo, e assumere questa prospettiva da parte del GEV avrebbe potuto favorire la crescita delle capacità degli Atenei di riflettere in modo consapevole e produttivo sugli effetti delle loro iniziative.

Come hanno infatti evidenziato sia i lavori del Gruppo "Valutazione" istituito presso la RUIAP, sia quelli di alcuni GEV di Ateneo, non è frequente avere nelle singole realtà locali una chiarezza teorica e operativa circa i concetti di risultati e impatto e, di conseguenza, il contributo del GEV ANVUR nel praticare standard metodologici, percorsi e indicatori rilevanti per una valutazione utile sia per ANVUR che per i singoli Atenei avrebbe potuto essere maggiore di quanto effettivamente prodotto.

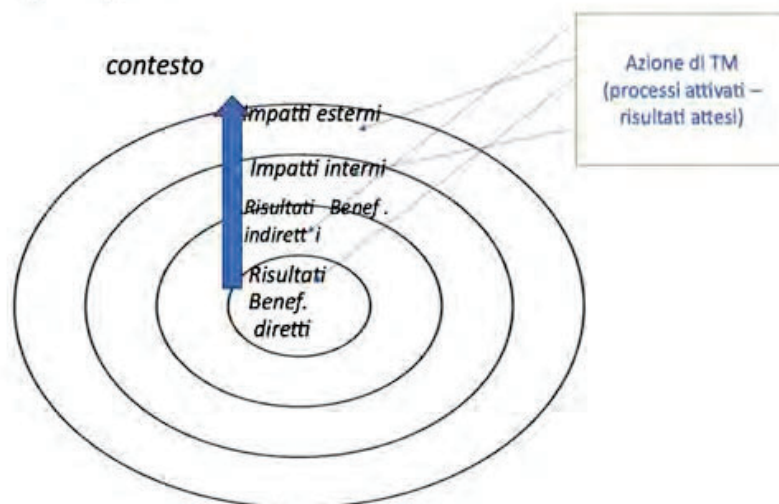
2. Una proposta per la valutazione

Come uscire da questa *impasse* magari nella prospettiva di possibili interventi migliorativi del prossimo esercizio del VQR? Nell'ambito del Gruppo di Lavoro RUIAP sulla valutazione dell'Apprendimento Permanente abbiamo proposto un percorso, ispirato al modello del PCM, che abbiamo chiamato "a cerchi concentrici", attraverso il quale gli effetti di un intervento vengono visti come l'esito del lancio di un sasso in uno stagno, che "colpisce" prima i destinatari diretti (gli *intended addressees*), poi quelli indiretti e infine il contesto, provocando un movimento che si sviluppa nel tempo e nello spazio dai primi agli ultimi: gli effetti, cioè, si verificano in tempi crescenti mano a mano che ci si allontana del "bersaglio" (e producono conseguenze via via meno intense) e concorrono con altri fattori a produrre mutamenti strutturali anche nel contesto.

Questo modello potrebbe rafforzare l'utilizzo in chiave di apprendimento da parte degli Atenei, che hanno ormai avvertito l'esigenza di valutare le attività di TM in chiave di learning e al di là delle periodiche occasioni della VQR (e dei criteri operativizzati in modi diversi dai valutatori di turno), in una prospettiva di rendicontazione sociale, di *Civic University* svincolata dalle periodiche esigenze adempimentali.

In questo modo si potrebbero aiutare i colleghi impegnati in azioni di TM a distinguere, nell'auspicabile monitoraggio e valutazione degli esiti dei loro sforzi, tra le *realizzazioni*, ossia le attività e i servizi messi in campo (quelli che rientrano, nel linguaggio del GEV, nel contributo della struttura proponente), i *risultati*, ossia gli effetti conseguiti sui destinatari diretti e indiretti degli interventi imputabili causalmente a questi (nel lessico anvruriano il valore aggiunto), gli *impatti*, magari distinti in *interni* (riferiti all'Università stessa) e *esterni*, ossia sul contesto entro il quale l'intervento agisce⁹.

Fig.1 – dai risultati agli impatti



Fonte: Brambilla e Palumbo (2024)

I vantaggi di questa impostazione potrebbero poi essere massimizzati, per le finalità interne di rendicontazione sociale, dalla descrizione della “teoria del cambiamento”, ossia della ragione per cui le attività intraprese avrebbero prodotto gli effetti (previsti o meno, intenzionali o meno) rilevati. Siamo

⁹ In breve, rispetto al modello del PCM, il valore aggiunto è assimilabile ai risultati, mentre paradossalmente il vero assente è l'impatto, che è spesso costituito dagli stessi esiti dell'intervento, senza un'adeguata distanza temporale e senza tener conto degli effetti inattesi.

consapevoli che le modalità di selezione dei casi di studio sono state molto diverse nei diversi Atenei, l'apprendimento però dovrebbe riguardare in primis il livello strategico che ha scelto i casi di studio e ne ha magari guidato la redazione, il livello intermedio dei Dipartimenti/Centri che hanno partecipato alla loro selezione e che spesso ne sono stati i gestori e una platea più ampia di persone impegnate nelle diverse azioni di TM, al di là del fatto che abbiano portato a presentare casi di studio o meno.

Le iniziative di formazione continua nei territori che la RUIAP e il Gruppo di lavoro sulla Valutazione intendono attivare nel prossimo anno accademico cercheranno di applicare questo modello a tutte le attività di TM campionate dagli Atenei Associati, provando ad attivare un processo di apprendimento strategico da parte degli Atenei, che potrebbero poi selezionare, con la cadenza della VQR, i casi di studio dotati dei migliori risultati e impatti e della “teoria del cambiamento” più affidabile, in relazione ai meccanismi “evidenti” e dei contesti di attuazione, in piena autonomia delle istituzioni e in una logica di miglioramento e learning che non risponde all'adempimento.

Le discontinuità emerse nel confronto tra la letteratura valutativa e l'impostazione del GEV ANVUR possono essere lette come il riflesso di un più generale scarto tra due concezioni della valutazione: da un lato, la valutazione come strumento di apprendimento riflessivo, verso la quale l'ANVUR nel tempo si è comunque orientata; dall'altro, la valutazione come dispositivo di giudizio comparativo *ex post*. In questo quadro, l'approccio delle *teorie del programma* offre una possibile chiave interpretativa – e correttiva – delle criticità rilevate e una strada potenziale generativa dell'autonomia della TM nei contesti di HE.

Nella valutazione *theory-based*, infatti, l'attenzione si sposta dalla mera misurazione degli effetti alla ricostruzione dei meccanismi causali che collegano risorse impiegate, azioni, ed esiti, con la mediazione dei contesti¹⁰. Impatto e valore aggiunto non sono categorie auto-evidenti, ma dipendono da una catena causale da esplicitare, che distingue tra output, outcome intermedi, effetti di medio-lungo periodo e condizioni contestuali abilitanti o ostacolanti, spingendo alla formalizzazione delle “teorie del cambiamento” soggiacenti all'intervento. Questa permette di affrancare la valutazione da una funzione prevalentemente rendicontativa, rafforzando la sua funzione di apprendimento. Secondo l'approccio di C. Weiss (1988), non è possibile un corretto “uso” della

¹⁰ E' appena il caso di citare il contributo di Pawson e Tilley (1997), che hanno proposto la nota formula $M+C=O$ (Meccanismo più Contesto = Outcome), per evidenziare che i meccanismi attivati da un intervento possono dare risultati diversi, anche opposti, a seconda dei contesti in cui operano. Questo si lega, in via più generale, alla concezione di causalità generativa proposta in particolare da Coleman (1990/2005), Elster (2007/2010) e Barbera (2004).

valutazione senza prendere in considerazione l’ambiente che circonda le organizzazioni; la valutazione, di conseguenza, perché sia utilizzabile appieno, potrebbe richiedere dei veri e propri processi di cambiamento anche a livello istituzionale che ne rendano possibili le “innovazioni” e che la stessa ricerca valutativa suggerisce in termini di “teorie del cambiamento”.

Trasferire i risultati della valutazione non è però un processo meccanico, indiscriminato e indipendente dal contesto e dai bisogni sociali, specie di quelli delle fasce deboli. Occorre implementare un approccio di “trasferimento meditato” (Schon, 1987), che consenta un processo ragionevole di trasferibilità degli esiti in contesti simili, ovvero non costruendo ex novo i disegni di ricerca, ma rendendo la valutazione in grado di rispondere a domande e bisogni socialmente rilevanti in termini di incidenza. I beneficiari del programma, ci insegna l’approccio teorico di Weiss, sono i primi ad avere un interesse in gioco e occorre prendere in carico i loro bisogni in quanto espressioni della società civile (in tal senso Carol Weiss rifiuta il concetto troppo astratto di comunità) e trasformarli in “domande di ricerca” in vista di una giustizia sociale (House, 1993). La valutazione, infine, è utile in quanto la logica e i processi della ricerca sono in grado di sostenere le organizzazioni ad insegnare a pensare in modo valutativo, ovvero *evidence based*. Michael Patton (1997), suggerisce che imparare come apprendere, ovvero imparare dal processo valutativo, significa andare oltre l’uso dei risultati e diventare “organizzazioni che apprendono”, ovvero oltre l’emivita dell’esercizio per agire in modo valutativo, il che permette di generare un impatto duraturo del cambiamento, sul modo di pensare degli stakeholder, a partire proprio dall’imparare a generare le “domande di valutazione”, su cui con Proietti (nel già citato numero EPale 2023) si è ragionato in merito proprio all’esercizio VQR 2015-2019 della TM.

Un orientamento alla valutazione come accompagnamento dei processi – tipico degli approcci basati sulle teorie del programma – consentirebbe dunque di interpretare l’impatto come una traiettoria in costruzione, e il valore aggiunto come una relazione causale situata, da discutere e argomentare piuttosto che da assumere implicitamente. Di qui la nostra proposta, di riconsiderare l’impianto metodologico e concettuale che soggiace alla valutazione della TM e della VdC, anche con un lavoro comune con gli studiosi del settore, che sostenga la funzione riflessiva della valutazione, i cui destinatari saranno in primis gli Atenei e solo in subordine e periodicamente l’ANVUR.

Detto per inciso e in conclusione di paragrafo, questo restituirebbe anche alla valutazione la dignità di disciplina orientata al miglioramento delle politiche e degli interventi e al rafforzamento della capacità degli attori sociali e degli stakeholder di governare i processi attivati, orientandone il percorso e gli effetti alle priorità e ai valori di riferimento degli Atenei che ne sono i

responsabili. E assegnerebbe all'ANVUR un ruolo propulsivo e proattivo nel favorire questa positiva evoluzione nei singoli Atenei. Una delle funzioni della ricerca valutativa, in quanto ricerca sociale applicata, è infatti proprio quella di innescare risultati di apprendimento collettivi che migliorano i processi di implementazione del programma in vista della sua riedizione.

Alla base vi è la considerazione teorica (Rossi, Freeman e Lipsey, 2018) che ogni valutazione debba essere cucita (e tagliata) su misura del proprio programma e che non esistano linee guida rigide per dirigere il processo di studio del contesto. In sostanza, ogni valutazione finalizzata al supporto decisionale e deliberativo deve quindi tener conto dei vincoli di tempo, di risorse, di varianti riconoscibili. Tutto il filone della cosiddetta “valutazione formativa” suggerisce che, oltre alla qualità eccellente della valutazione, essa debba necessariamente essere “utilizzabile”, ovvero fruibile e orientata ad un corretto uso dei risultati (Patton, 1997), rafforzando il legame tra le differenti politiche di Atenei e Dipartimenti e aprendo realisticamente il processo valutativo alle comunità che sono coinvolte nei processi e nelle iniziative (non solo come beneficiari ma come protagonisti del processo di valutazione).

3. Apprendere dalla Valutazione

Cosa abbiamo quindi imparato da questo doppio esercizio (2015-2019 – 2020-2024) di valutazione della TM di ANVUR, al di là delle classi assegnate dal GEV al caso di VdC sottomesso dalle istituzioni? Siamo in grado come istituzioni di Higher Education di rispondere a criteri di sostenibilità interna della valutazione, di essere indipendenti nel processo di riflessività valutativa e di innescare un processo di apprendimento?

In un framework orientato all'utilizzazione della ricerca valutativa, l'ascolto dei beneficiari diretti e indiretti (e non solo attraverso l'uso indicatori di valore aggiunto attribuiti spesso da persone e organizzazioni diverse dai beneficiari) nel processo di valutazione è assolutamente fondamentale. A nostro parere, l'introduzione nel processo di ricerca valutativa di “visite di servizio” da parte degli stessi membri del GEV capaci di applicare le loro competenze valutative in una logica di capacity building, finalizzata a mettere in evidenza l'intera gamma di effetti conseguiti dalle azioni di TM, potrebbe essere una delle innovazioni più strategiche per risolvere i “bias” delle differenze di contesto spesso complesse da evidenziare e mettere in luce nella lettura dei casi di studio.

A questo riguardo sarebbe utile predisporre, secondo le linee guida proposte da RUIAP, un modello di analisi degli effetti, interni ed esterni, delle

attività di TM nel campo specifico dell’AP capace, da un lato, di essere applicato anche alle altre missioni degli Atenei e, dall’altro lato, di fornire agli Atenei uno strumento di monitoraggio e valutazione in itinere sganciato dalle scadenze VQR e legato invece all’evoluzione della loro pianificazione strategica. L’adozione di un sistema valutativo della Vdc, autonomo rispetto ai cronoprogrammi ANVUR, consentirebbe alle istituzioni di alimentare costantemente (in relazione agli outcome delle iniziative che si intende misurare in termini di generazione di impatti ed effetti strutturali) il paniere dei casi di studio, moltiplicando gli apprendimenti e migliorando il processo decisionale in termini di condivisione con l’intero Ateneo. Sul piano della sostenibilità esterna si potrebbe lavorare ad una sorta di “spinta gentile”, sempre non collegata alla premialità, ma che sostenga i comportamenti virtuosi degli atenei, in termini di community engagement come *unica missione di valore pubblico e collettivo* (si vedano i lavori internazionali sul superamento della distinzione tra prima, seconda e terza missione). Provocatoriamente, si potrebbe anche pensare ad un sostegno agli Atenei in maggiore difficoltà, che preveda missioni di supporto su richiesta degli Atenei, rivolte a promuovere il miglioramento di criticità riscontrate e riconosciute, ovviamente in riferimento a tutte e tre le missioni.

A questo riguardo sarebbe utile predisporre, secondo le linee guida proposte da RUIAP, un modello di analisi degli effetti, interni ed esterni, delle attività di TM nel campo specifico dell’AP capace, da un lato, di essere applicato anche alle altre missioni degli Atenei e, dall’altro lato, di fornire agli Atenei uno strumento di monitoraggio e valutazione in itinere sganciato dalle scadenze VQR e legato invece all’evoluzione della loro pianificazione strategica. L’adozione di un sistema valutativo della Vdc, autonomo rispetto ai cronoprogrammi ANVUR, consentirebbe alle istituzioni di alimentare costantemente (in relazione agli outcome delle iniziative che si intende misurare in termini di generazione di impatti ed effetti strutturali) il paniere dei casi di studio, moltiplicando gli apprendimenti e migliorando il processo decisionale in termini di condivisione con l’intero Ateneo. Sul piano della sostenibilità esterna si potrebbe lavorare ad una sorta di “spinta gentile”, sempre non collegata alla premialità, ma che sostenga i comportamenti virtuosi degli atenei, in termini di community engagement come *unica missione di valore pubblico e collettivo* (si vedano i lavori internazionali sul superamento della distinzione tra prima, seconda e terza missione).

Provocatoriamente, si potrebbe anche pensare ad un sostegno agli Atenei in maggiore difficoltà, che preveda missioni di supporto su richiesta degli Atenei, rivolte a promuovere il miglioramento di criticità riscontrate e riconosciute, ovviamente in riferimento a tutte e tre le missioni. Questo potrebbe anche ampliare il raggio d’azione dell’Agenzia, responsabilizzandola

maggiormente sugli effetti della sua azione, per esempio analizzando se e come siano aumentati gli effetti di apprendimento dopo la diffusione dei ranking nel maggio 2026, chiedendosi se e come gli stakeholders (atenei e comunità locali) ne siano effettivamente informati e siano in grado di tenerne conto in modo costruttivo. Ogni esercizio valutativo dovrebbe portare ad un effetto di apprendimento chiaro e condiviso che dia la possibilità di supportare il processo deliberativo dell'organizzazione che ha valutato. La valutazione formativa non dovrebbe fermarsi all'analisi testuale del caso di TM, all'aspetto della scrittura e della identificazione di indicatori. Un buon caso è tale in ottica learning se alla base c'è stato un processo valutativo che ha permesso di descrivere bene gli impatti generati, che ha definito le baseline (possibilmente in modo partecipato), coinvolto nell'implementazione gli stakeholders e definito l'impatto atteso in termini di risultati da generare (misurando poi il saldo naturale tra generato e non). Oltre all'analisi documentale, la valutazione può utilizzare dati prodotti in altri ambiti valutativi secondo una logica per così dire emancipatoria e non lineare che si potrebbe applicare alla Terza Missione, anche con interviste e raccolte dati empirici non raccolti ma utilizzati ad hoc per il processo di costruzione delle teorie del cambiamento. Su questi aspetti infatti intende lavorare il gruppo RUIAP su Valutazione e apprendimento.

Inoltre, occorre fare una riflessione su quanto la VQR/TM abbia favorito l'apprendimento degli Atenei e, in particolare, quanto questo si sia tradotto in miglioramento rispetto all'esercizio precedente e comunque abbia prodotto una maggiore responsabilità sociale dell'istituzione. E come dimenticare - perché ci sia apprendimento- la funzione dei contesti: ovvero la realistica messa in evidenza delle differenze territoriali e la naturale differenziazione non solo in termini di grandi, medie e piccole università, ma anche di differenze collegate al processo decisionale di selezione dei casi e di condivisione istituzionale di esiti e di impatto che ne abbiano giustificato la scelta per le istituzioni?

La finalità formativa della valutazione implica anche la possibilità che, nel costruire le teorie del cambiamento nel corso di implementazione dell'iniziativa, l'agency degli attori sia accompagnata da una messa a valore del cambiamento generato in relazione alla stessa iniziativa oggetto di valutazione. In quest'ottica, la valutazione rientra nella sua dimensione di ricerca sul campo e applicata che sostiene il processo decisionale della stessa istituzione (per esempio il processo di selezione e scelta dei casi di studio). Sul piano dei contenuti, la disponibilità di un glossario condiviso e di un paniere minimo comune di indicatori sarebbero di grande aiuto agli Atenei per sviluppare autovalutazioni del proprio operato nelle varie aree e campi di azione e, nel caso sviluppato in questo saggio, in particolare per quel che concerne l'Apprendimento Permanente.

Infine, aspetto non meno rilevante riguarda il processo di miglioramento della stessa formazione degli esperti e delle esperte impegnati nei singoli Atenei, ovvero la crucialità delle iniziative *evidence based* (per esempio come dicevamo in precedenza con visite in loco) finalizzate a questo scopo. In una prospettiva di valutazione della TM sulla base di un modello di reclutamento per competenze del GEV si ridurrebbero probabilmente i *bias* dovuti alla numerosità di assegnazioni di casi a campi di azione particolarmente numerosi (p.e. per il public engagement; il trasferimento tecnologico, ecc.) a svantaggio di campi e aree tematiche di più recente acquisizione nel dominio della TM (p.e. i campi delle scienze della vita, della sostenibilità, ecc.).

In che misura il secondo esercizio di valutazione della TM attraverso la definizione di “casi” ha ottenuto un effetto di apprendimento? Che ruolo ha giocato “realisticamente” la dimensione contestuale micro-meso-macro (rilevanza interna al caso) nel generare tali effetti in termini di lezioni apprese? Il percorso delle istituzioni in riferimento al processo di selezione dei casi da sottoporre alla valutazione ANVUR e che porta alla scelta definitiva dei casi, a nostro parere, viene in larga parte delegato alle strutture dipartimentali degli atenei e questo è un elemento significativo indiretto per mostrare la scarsa condivisione interna del processo decisionale a vantaggio di una logica di ripartizione che isola ancora più la TM nel quadro interno dell’istituzione.

Inoltre, la possibilità che i casi coinvolgano fino a tre aree tematiche e fino a tre campi di azione richiederebbe una analisi interna della diffusività degli impatti generati che le istituzioni non riescono a mettere a valore, spesso, a nostro avviso, proprio per la conseguenza che comporta la ripartizione del numero dei casi in relazione al numero dei dipartimenti e non in relazione all’impatto generato (elemento su cui bisognerebbe intervenire anche indipendentemente dalle tempistiche di ANVUR).

Su questi aspetti non è più possibile derogare, occorre rimettere al centro della valutazione l’apprendimento reale delle istituzioni e comprendere in quale misura tale processo richieda una valutazione trasformativa che risponda ad un cambiamento misurabile e di impatto (da adempimento vs learning e miglioramento organizzativo). E sollecitare le istituzioni valutate a partecipare attivamente alla costruzione di modelli valutativi utilizzabili congiuntamente, nella rispettiva autonomia, dagli Atenei e dall’ANVUR.

Dichiarazioni

Gli autori dichiarano che il Software Copilot (Microsoft) è stato utilizzato esclusivamente per il controllo dei dati e dei riferimenti bibliografici. Gli autori si assumono la piena responsabilità del manoscritto finale.

Riferimenti bibliografici

ANVUR (2024). *Documento sulle modalità di valutazione dei casi studio Gruppo di Esperti della Valutazione Attività di valorizzazione delle conoscenze (GEV Interdisciplinare)*, Roma, 31 luglio VQR-2020-2024-Modalita-valutazione-GEV_VdC.pdf

Barbera, F. (2004). *Meccanismi sociali. Elementi di sociologia analitica*. Bologna: Il Mulino.

Boffo, S., Moscati, R., Rostan, M. (2025). *La Terza Missione nell'Università italiana. Politiche e attività in dodici atenei*. Milano: Guerini e Associati.

Brambilla, R., Palumbo, M. (2024). Valutare per valorizzare: l'Apprendimento permanente nella Valutazione della Terza Missione Accademica. *Epale Journal on Adult Learning and Continuing Education*, 14, pp. 26-35, Marzo.

Carrera, L., Castelnuovo, I., Cavaglia, G.M.C., Chimenti, S., De Bortoli, A., Pascuzzi, E., Piazza S., Sardi, M. (2025). Il Public Engagement delle università italiane nella VQR 2015-19. Uno studio di casi. *RIV. Rassegna Italiana di Valutazione*, XXIX, 92, pp. 123-143.

Celletti, A., Uricchio, A. F. (2025). Valutare per valorizzare e migliorare: l'esperienza della VQR. *Quaderni di Sociologia*, 95, pp. 67-80.

Chen, H.T. (2005). *Theory-Driven Evaluations / Practical Program Evaluation* Sage.

Charles, D., Nieth, L., Coenen, L. Cinar, R. (eds.) (2026). *Universities, Knowledge and Places*. London-New York: Routledge.

Coleman, J. (1990). *Foundations of Social Theory*, Harvard University Press, Cambridge, MA; trad. it. *Fondamenti di teoria sociale*. Bologna: Il Mulino, 2005.

Elster, J. (2007). *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.; trad.it. *La spiegazione del comportamento sociale*. Bologna: Il Mulino, 2010.

European Commission: Directorate-General for International Partnerships, *Evaluation handbook*, Publications Office of the European Union. (2024), <https://data.europa.eu/doi/10.2841/340793>, <https://op.europa.eu/en/publication-detail/-/publication/06350947-4d56-11ef-acbc-01aa75ed71a1/language-en>

Fetterman, D.M., Kaftarian, S.J., Wandersman, A. (2015). *Empowerment evaluation: Knowledge and tools for self-assessment, evaluation capacity building, and accountability*. Thousand Oaks, CA: Sage.

Fontana, R., Valentini, E. (a cura di) (2020). *Conseguenze della valutazione. Idee e pratiche dei docenti universitari nelle scienze sociali*. Milano: FrancoAngeli.

Goddard, J., Hazelkorn, E., Kempton, L., Vallance, P. (2021). *The civic university: The policy and leadership challenges*. Cheltenham–Northampton (MA): Edward Elgar.

Hazelkorn, E. (2020). *Research evaluation: Global trends and good practice*. Paris: OECD Publishing

House, E. (1993). *Professional Evaluation: Social Impact and Political Consequences*, Thousand Oaks, Ca., Sage.

House, E.R., Howe, K.R. (1999). *Values in evaluation and social research*, Thousand Oaks, Ca., Sage.

Huemann, M., Turner, R. (eds.) (2024). *The Handbook of Project Management* (6th ed.). London: Routledge.

Leeuw, F.L. (2023). John Mayne and Rules of Thumb for Contribution Analysis: A Comparison with Two Related Approaches. *Canadian Journal of Program Evaluation*, 37(3), 403–421.

Mayne, J. (2001). Addressing Attribution through Contribution Analysis. *Canadian Journal of Program Evaluation*, 16(1), pp. 1–24.

OECD/European Union (2024). *Measure, Manage and Maximise Your Impact: A Guide for the Social Economy*. Local Economic and Employment Development (LEED), OECD Publishing, Paris, <https://doi.org/10.1787/2238c1f1-en>.

Palumbo, M., Piazza, R. (2023). How to support Universities in strengthening the impact of ULLL and UNCE activities. The role of the Italian Network of UCL (RUIAP), eucen Studies eJournal of University Lifelong Learning. Vol 7, No 02. eucen Electronic Press. pp. 13-18. <https://doi.org/10.53807/070213Cq>

Palumbo, M., Scardigno, F. (2024). Lunga vita alla Valutazione (della Terza Missione) degli Atenei: tra rendicontazione sociale e apprendimento permanente, *RIV, Rassegna Italiana di Valutazione*, 88/89, pp. 17-32.

Palumbo, M., Scamuzzi, S., Borgna, P. (a cura di) (2025). *Valutazione della ricerca: dai progetti agli effetti*. Numero monografico di *Quaderni di Sociologia*, 95, pp. 7-23.

Patton, M.Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.

Patton, M.Q. (1997). *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, Ca., Sage.

RACCOMANDAZIONE (UE) 2022/2415 DEL CONSIGLIO del 2 dicembre 2022 sui principi guida per la valorizzazione delle conoscenze, GUUE L 317 del 9.12.2022.

Reale, E. (a cura di) (2025). *Gli effetti della valutazione esterna della ricerca sul lavoro accademico in Italia*. Milano: FrancoAngeli.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2018). *Evaluation: A Systematic Approach* (8th ed.). Thousand Oaks, Ca., Sage.

Schon, D.A. (1987). *Education for the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*, San Francisco: Jossey-Bass.

Scardigno A.F., Colosimo M., Manuti A., Pastore S. (2024). *Lifelong Learning and Higher Education serving for the inclusion of refugees and students with migratory backgrounds*, in “Contemporary Issues in University Lifelong Learning. Making diversity and inclusion strategy” in *European Journal of University Lifelong Learning* (EJULL), 8(1), <https://doi.org/10.53807/0801ZDRx>

Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park, CA.: Sage, 4^a ed.

Stame, N. (2015). Hirshman possibilismo e valutazione. *RIV. Rassegna Italiana di Valutazione*, XXII, 62, pp. 103-119.

Weiss, C. (1988). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice* 9(1), pp. 5-19.

Weiss, C. (1997). Theory-based evaluation: past, present and future. in Rog D. J. (ed.) *Progress ad Future Directions in Evaluation*, New Directions for Evaluation, 76, San Francisco, Jossey – Bass pp. 41-55.

White, H. (2009). Theory-based impact evaluation: principles and practice. *Journal of Development Effectiveness*, 1:3, 271-284, DOI: 10.1080/19439340903114628