

XXX - 95/2026

rív

Rassegna
Italiana di
Valutazione

Rassegna Italiana di Valutazione

XXX - 95/2026

rív

Rassegna
Italiana di
Valutazione

FrancoAngeli 

riV. Rassegna Italiana di Valutazione.

Quadrimestrale dell'Associazione Italiana di Valutazione

Direttore editoriale

Giuseppe Moro, *Università di Bari Aldo Moro*, giuseppe.moro@uniba.it

Comitato direttivo

Veronica Lo Presti, *Sapienza Università di Roma*, veronica.lopresti@uniroma1.it

Francesco Mazzeo, *Università di Catania*, fmazzeo@unict.it

Erica Melloni, *Università Mercatorum*, erica.melloni@unimercatorum.it

Comitato editoriale

Michela Freddano, *INVALSI*, michela.freddano@invalsi.it

Francesca Pia Scardigno, *Valutatrice indipendente*, francescapia.scardigno@univaq.it

Claudio Torrigiani, *Università di Genova*, claudio.torrigiani@unige.it

Giancarlo Vecchi, *Politecnico di Milano*, giancarlo.vecchi@polimi.it

Comitato scientifico

Maria Carmela Agodi, *Università di Napoli Federico II*, agodi@unina.it

Gianluca Argentin, *Università di Milano Bicocca*, gianluca.arginin@unimib.it

Ferruccio Biolcati Rinaldi, *Università di Milano*, ferruccio.biolcatiunimi.it

Simone Busetti, *Università di Teramo*, sbusetti@unite.it

Daniele Checchi, *Università di Milano*, danile.checchi@unimi.it

Antonio Fasanella, *Sapienza Università di Roma*, antonio.fasanella@uniroma1.it

Osvaldo Feinstein, *Valutatore indipendente*, ofeinstein@yahoo.com

Dora Gambardella, *Università di Napoli Federico II*, dora.gambardella@unina.it

Giampietro Gobo, *Università di Milano*, giampietro.gobo@unimi.it

Sebastian Lemire, *Strada Education Foundation*, sebastian.lemire@strada.org

Andrea Lippi, *Università di Firenze*, lippi@unifi.it

Andrea Mairate, *Commissione Europea – Institute for European Studies*,
andrea.mairate@ec.europa.eu

Mita Marra, *Università di Napoli Federico II*, mita.marra@unina.it

Katia Horber Papazian, *Université de Lusanne*, Katia.Horber-Papazian@unil.ch

Carlo Pennisi, *Università di Catania*, cpennisiunict.it

Donatella Poliandri, *INVALSI*, donatella.poliandri@invalsi.it

Laura Polverari, *Università di Padova*, laura.polverari@unipd.it

Stefania Profeti, *Università di Bologna*, stefania.profeti@unibo.it

Claudio Radaelli, *European University Institute*, claudio.radaelli@eui.eu

Emanuela Reale, *CNR Roma*, emanuela.reale@icres.cnr.it

Gloria Regonini, *Università di Milano*, gloria.regonini@unimi.it

Maria Stella Righettini, *Università di Padova*, mariastella.righettini@unipd.it

Robert Shepherd, *Carleton University (CA)*, RobertPShepherd@cunet.carleton.ca

Dominika Wojtowicz, *Kozminski University (PL)*, doma@kozminski.edu.pl

Impaginazione e cura redazionale Micaela Sciarra, micaelasciarra@gmail.com

Sommario

a. XXX – n. 95

Editoriale	pag. 5
F. Mazzeo Rinaldi. <i>Dall'output algoritmico al giudizio valutativo: governare l'evidenza nell'era dell'Intelligenza Artificiale.</i>	» 12
W.A. Canu. <i>La facilitazione valutativa nell'approccio partecipativo.</i>	» 38
E. Celardi. <i>Contesti di attuazione, mediazioni operative e traiettorie di cambiamento. Un'analisi valutativa del progetto Ri-Belli.</i>	» 56
A. Mastrolia, M. Palmieri, D. De Valeri. <i>L'infrastrutturazione sociale come leva di sviluppo: Fondazione con il Sud a Castelbuono.</i>	» 75
E. Pascuzzi, S. Chimenti. <i>Partecipazione e valutazione nelle politiche universitarie di Public Engagement: verso modelli dialogici e generativi.</i>	» 102
M. Palumbo, A.F. Scardigno, A.F. Uricchio. <i>Per una "teoria del cambiamento": riflessioni a margine della valutazione delle iniziative di Valorizzazione delle conoscenze degli Atenei Italiani</i>	» 122

Editoriale

La comunità dell'Associazione Italiana di Valutazione sa bene che non esiste un unico modo di intendere e praticare la valutazione. Alkin (2004) ha rappresentato le numerose tradizioni della valutazione attraverso un albero con tre grossi rami: il primo, quello dei metodi, interessato allo sviluppo metodologico e alla validità del disegno valutativo, rappresentato da studiosi quali Campbell, Cook e Rossi; il secondo ramo, quello dell'attribuzione di valore, incentrato sulla formulazione del giudizio valutativo, al quale sono riconducibili Scriven, Stake e altri studiosi; e il ramo dell'uso, focalizzato sul contributo della valutazione ai processi decisionali e rappresentato, tra gli altri, da Patton, lo stesso Alkin, e altre studiose e studiosi.

L'importanza del rigore metodologico, sottolineata dai rappresentanti inclusi nel primo filone, deriva dal fatto che la valutazione trae gran parte della propria legittimazione dal riconoscimento della competenza metodologica dei valutatori. La sua autorevolezza si fonda, di conseguenza, sulla solidità del disegno di ricerca e sulla sua adeguatezza rispetto all'oggetto valutato. L'affermazione di Rossi secondo cui «*getting it right is more important than getting along*» (Rossi, 2004, pp. 127) esprime efficacemente questa priorità. In altre parole, una valutazione che non si fondi su un metodo solido, argomentabile e discutibile, per questi autori, smette di essere valutazione. Il metodo, tuttavia, non basta. Come ricordano gli studiosi riconducibili agli altri 'rami', la valutazione comporta la costruzione di un giudizio sull'oggetto valutato e solleva quindi la questione di chi contribuisca a definire le aspettative e i criteri rispetto ai quali tale giudizio viene formulato. Questa prospettiva solleva importanti questioni rispetto a chi viene incluso, e chi invece escluso, dalla costruzione del giudizio valutativo. La valutazione non può infine trascurare l'utilizzabilità del processo e dei suoi risultati da parte dei destinatari, siano essi i committenti o, più in generale, la comunità dei portatori di interesse. Ciò richiede di interrogarsi sui processi, sui linguaggi e sulle soluzioni operative capaci di favorirne l'uso, e arginare così uno dei principali punti di

malcontento dei valutatori, ovvero il poco interesse rispetto ai risultati di valutazione da parte di chi dovrebbe invece farsene carico.

Gli articoli del numero 95/2026 della RIV affrontano tutti gli aspetti appena richiamati, sottolineando le tensioni tra le diverse prospettive. Alla valutazione si chiede di produrre evidenze solide e tempestive, anche avvalendosi del contributo potente, ma non privo di ambiguità, offerto oggi dall'intelligenza artificiale. Le si chiede, al tempo stesso, di rendere conto della complessità dei processi, coinvolgere una pluralità di attori, sostenere le decisioni e contribuire all'apprendimento delle organizzazioni e delle comunità, rimanendo sempre ancorata a metodi trasparenti e fondati. Queste esigenze non sono necessariamente convergenti. Possono anzi generare tensioni tra standardizzazione e attenzione ai contesti, rapidità e riflessività, attribuzione degli effetti e comprensione dei processi, indipendenza del giudizio e partecipazione dei portatori di interesse. E inoltre, l'esperienza della valutazione genera essa stessa apprendimento rispetto alle modalità che possono aiutare a superare impasse e nodi critici affrontati sul campo.

I contributi raccolti nel fascicolo si collocano all'interno di questo spazio problematico. Pur affrontando oggetti molto diversi – dall'impiego dell'intelligenza artificiale nella valutazione, al miglioramento dei servizi socio-sanitari, ai progetti di contrasto alla povertà educativa, dai programmi di sviluppo locale alla valutazione nelle università – condividono una concezione della valutazione che non può essere ridotta all'assessment dei risultati raggiunti. Nei diversi articoli, valutare significa rendere espliciti i passaggi attraverso cui si costruisce un giudizio; interrogare l'interazione tra interventi e contesti; creare le condizioni per il confronto tra conoscenze e punti di vista; accompagnare processi di cambiamento; produrre apprendimenti capaci di orientare l'azione futura.

Il filo che unisce i contributi può essere ricostruito attraverso tre questioni. La prima riguarda la costruzione dell'evidenza e il rapporto tra informazione, interpretazione e giudizio. La seconda concerne la partecipazione, intesa non soltanto come coinvolgimento dei portatori di interesse, ma come dimensione epistemologica, metodologica e politica della valutazione. La terza riguarda il nesso tra valutazione e cambiamento: non un cambiamento lineare, generato in modo teleologico dall'intervento-causa, ma un processo situato, mediato dagli attori e dai contesti e capace di produrre apprendimento individuale, collettivo e istituzionale secondo percorsi non sempre lineari e spesso da ricostruire prestando attenzione alla dimensione di caso.

Il fascicolo si apre con il contributo di Francesco Mazzeo Rinaldi, dedicato al rapporto tra intelligenza artificiale ed evidenza valutativa. La diffusione di sistemi algoritmici in grado di analizzare grandi quantità di dati,

riconoscere configurazioni, formulare previsioni e generare testi ha alimentato aspettative rilevanti anche nel campo della valutazione. L'articolo pone una questione epistemica cruciale: come governare i processi di inferenza, interpretazione e giudizio che definiscono la valutazione nell'epoca dell'intelligenza artificiale? Attraverso una *critical narrative review* della letteratura internazionale, Mazzeo Rinaldi individua tre principali modi di concettualizzare l'intelligenza artificiale: come potenziamento della capacità analitica; come fonte di rischi epistemici ed etici dai quali proteggersi; come oggetto di valutazione e di governance. Queste prospettive lasciano tuttavia relativamente inesplorata la catena inferenziale che collega il prodotto dell'algoritmo al giudizio valutativo, del quale chi valuta mantiene la responsabilità ultima. Per rendere osservabile tale catena, l'autore propone un framework articolato in cinque stadi: output algoritmico, validazione tecnica, interpretazione informata dalla teoria, triangolazione e giudizio valutativo. Si tratta di una proposta analitica per riconoscere le forme di «compressione epistemica», ossia i punti nei quali inferenze, interpretazioni e responsabilità possono essere delegate, opacizzate o date per scontate. Si tratta di un nodo che merita ulteriore approfondimento da parte della comunità dei valutatori e delle valutatrici, affinché le pratiche valutative informate dall'intelligenza artificiale possano essere governate in modo metodologicamente difendibile e responsabile rispetto ai giudizi che ne derivano.

Il successivo contributo di Walter Antonio Canu affronta gli aspetti metodologici della facilitazione nell'ambito della valutazione partecipativa. Canu, richiamando ampiamente l'opera di Claudio Bezzi, distingue due principali giustificazioni della partecipazione. La prima è di carattere etico e politico: partecipare significa riconoscere ai soggetti interessati, comprese le componenti meno forti o meno rappresentate, la possibilità di contribuire alla definizione di ciò che conta e dei criteri in base ai quali un intervento viene giudicato. La seconda è utilitaristico-metodologica: la partecipazione mira ad accrescere la pertinenza della valutazione, migliorare la qualità delle informazioni disponibili e favorire la comprensione e l'uso dei risultati. Per ottenere tali benefici serve tuttavia un adeguato apparato metodologico a supporto di chi svolge la funzione di facilitazione dei processi valutativi, sostenendo nelle diverse fasi del percorso di valutazione il coinvolgimento, l'elicitazione delle informazioni, la costruzione di significati e di senso condiviso, la presa di decisioni e la creazione di soluzioni. La facilitazione viene così presentata come una risorsa metodologica capace di rendere effettiva, e non soltanto simbolica, la partecipazione e di collegare la produzione di conoscenza alla formulazione dei giudizi e al loro possibile uso. Un punto non trattato, ma che certamente emerge dal contributo di Canu, riguarda le

competenze necessarie per svolgere il ruolo di valutatore-facilitatore e come si costruiscono: un ambito su cui AIV può proporre approfondimenti e proposte operative.

L'articolo di Elvira Celardi si interroga su come la scelta metodologica possa condizionare l'appropriatezza e la profondità del giudizio valutativo. L'autrice mette a confronto le differenti rappresentazioni del successo di un progetto che possono derivare, da un lato, dagli indicatori sintetici di risultato e, dall'altro, dall'analisi ravvicinata dei processi di implementazione. L'occasione è offerta dalla valutazione del progetto Ri-Belli, un'iniziativa di contrasto alla povertà educativa minorile realizzata in diverse aree del territorio ragusano e finanziata dall'impresa sociale Con i Bambini. Adottando una prospettiva possibilista – intesa come ricerca aperta delle possibilità di cambiamento effettivamente presenti nelle situazioni concrete – insieme a un approccio theory-based e realista, Celardi si interroga sul contributo del progetto all'attivazione di specifici processi di cambiamento. Combinando dati di monitoraggio, documentazione progettuale, focus group, interviste e osservazione diretta, l'autrice individua differenti configurazioni territoriali e familiari che incidono sull'attivazione dei processi di cambiamento. Lo stesso intervento – il laboratorio di educazione all'affettività – può così generare traiettorie di convergenza, disallineamento o adattamento trasformativo che possono essere spiegati solo con un'analisi ravvicinata dei processi di implementazione. In altre parole, Celardi mostra che il raggiungimento dei target non sia sufficiente a comprendere la capacità del progetto di generare cambiamento. Scarti, resistenze e adattamenti non costituiscono residui dell'analisi, ma elementi conoscitivi essenziali per ricostruirne il funzionamento. L'utilità della valutazione non consiste allora nell'individuare una soluzione da riprodurre altrove in modo invariato, ma nel riconoscere le condizioni, le soglie di accesso e le forme di mediazione delle quali tenere conto nella progettazione di interventi futuri.

L'articolo di Alfonsina Mastrolia, Marco Palmieri e Davide De Valeri presenta l'architettura metodologica e i risultati della valutazione ex post del programma "Sviluppo locale", promosso da Fondazione con il Sud a Castelbuono, in provincia di Palermo. In questo caso, il tema della partecipazione appare sia come uno strumento operativo sia come un risultato atteso della strategia di sviluppo locale. Il programma, mirante a rafforzare l'architettura sociale del territorio, ha previsto un ampio percorso partecipativo di coprogettazione delle azioni. La valutazione ex post mira a comprendere se esso abbia effettivamente sostenuto lo sviluppo delle reti e delle capacità collaborative presenti nel territorio. A tale scopo, il disegno metodologico integra la *Social Network Analysis*, utilizzata per ricostruire la struttura e l'evoluzione

delle relazioni tra le organizzazioni, e la *Nominal Group Technique*, attraverso la quale i portatori di interesse interpretano i risultati e concorrono alla formulazione del giudizio. L'integrazione delle due tecniche combina così l'analisi strutturale delle reti con l'interpretazione degli attori, facendo emergere un giudizio fondato su forme di conoscenza differenti e complementari. La valutazione conclude che il finanziamento ha consolidato risorse relazionali preesistenti, senza tuttavia riuscire a trasformare pienamente i diversi partenariati in una rete di comunità. Sul piano dell'uso, l'articolo mette inoltre in evidenza il limite di una valutazione condotta esclusivamente a posteriori. Da qui deriva la proposta di affiancare alla valutazione finale occasioni intermedie di monitoraggio partecipato, riflessione e capacity building. La valutazione potrebbe così diventare parte dell'accompagnamento del programma, rafforzandone l'utilità e accrescendo la capacità degli attori locali di riflettere sulle proprie pratiche e contribuire consapevolmente al suo sviluppo.

Il contributo di Stefania Chimenti ed Emanuela Pascuzzi discute il ruolo della valutazione partecipativa nella costruzione della politica di Public Engagement dell'Università della Calabria. In questo caso il processo valutativo, di impostazione costruttivista, serve contemporaneamente a ricostruire le pratiche esistenti, far emergere le rappresentazioni della comunità accademica e co-ideare priorità e strumenti per il futuro. L'assunto di fondo è che la produzione di conoscenza sul Public Engagement non possa essere separata dall'attivazione di processi di cambiamento: la ricerca è già intervento e la valutazione è già trasformazione. Attraverso il Confronto Creativo e l'*Appreciative Inquiry*, il percorso ha favorito la costruzione di un linguaggio comune, di una rappresentazione condivisa del Public Engagement e di nuove relazioni organizzative capaci di sostenere la policy. Le autrici osservano che, in organizzazioni complesse come le università, la valutazione partecipativa permette di valorizzare conoscenze e pratiche diffuse, evitando gli adattamenti passivi che possono risultare da processi decisionali eccessivamente centralizzati. Il suo sviluppo richiede tuttavia tempo, risorse, competenze di facilitazione e condizioni istituzionali adeguate. Il caso dell'Università della Calabria mostra come il coinvolgimento della comunità accademica possa rafforzare la legittimità delle scelte e sostenerne l'attuazione, facendo emergere istanze capaci di alimentare anche il confronto nazionale su questi temi.

È proprio sul piano del sistema nazionale che si colloca il contributo conclusivo di Mauro Palumbo, Anna Fausta Scardigno e Antonio Felice Uricchio. Gli autori – tutti in diversa misura coinvolti direttamente nei processi di valutazione ANVUR – discutono criticamente la valutazione della Terza

Missione, ridefinita nell'ultimo esercizio della VQR – valutazione della qualità della ricerca come Valorizzazione della Conoscenza, mostrando come le categorie e i dispositivi valutativi non si limitino a misurare un oggetto, ma contribuiscano a definirlo. La scelta di virare da Terza Missione a “Valorizzazione della Conoscenza” sembra derivare da un recepimento forse troppo pedissequo della Raccomandazione (UE) 2022/2415, che ha definito la valorizzazione della conoscenza come il processo attraverso il quale si crea valore sociale ed economico trasformando dati, competenze e risultati della ricerca in prodotti, servizi, soluzioni e politiche capaci di produrre benefici per la società. Secondo gli autori, questa definizione rischia di rappresentare una “camicia di Nesso” per l'Università, poiché una concezione modellata sul trasferimento e sulla produzione di valore tende a escludere dal perimetro le pur importanti attività fondate sulla collaborazione con la società civile, sulla coproduzione di beni pubblici e sulla responsabilità civica delle università. Più in generale, gli autori rivendicano la funzione della valutazione come processo capace di generare apprendimento e miglioramento negli atenei, ma pone anche una questione fondamentale relativa alla costruzione del giudizio e alla costruzione di definizioni condivise, distinzioni analitiche e una ricostruzione plausibile dei nessi che collegano attività ed i loro effetti. La proposta di utilizzare le teorie del cambiamento e un modello “a cerchi concentrici” mira precisamente a rendere esplicita la propagazione degli effetti, dalle realizzazioni agli effetti sul contesto esterno. La proposta degli autori non consiste nell'abbandonare la valutazione nazionale, ma nel modificare il rapporto tra valutazione esterna e capacità valutativa interna, in cui gli atenei dovrebbero diventare i primi utilizzatori delle evidenze prodotte, sviluppando sistemi continuativi di monitoraggio, autovalutazione e apprendimento non subordinati alle scadenze della VQR.

Considerati nel loro insieme, i contributi mostrano che i tre rami dell'albero di Alkin non descrivono territori separati. Il metodo è la condizione che rende evidenze e giudizi argomentabili, ma la sua scelta dipende dalle domande, dagli oggetti e dagli usi attesi. Il giudizio costituisce la responsabilità specifica della valutazione, ma implica sempre una decisione su chi sia legittimato ed effettivamente coinvolto nella definizione dei criteri e nella interpretazione delle evidenze. L'uso, infine, non rappresenta soltanto il momento conclusivo della comunicazione dei risultati, perché può influenzare fin dall'inizio la scelta delle domande, il disegno del processo e il coinvolgimento degli altri attori.

Le tensioni che attraversano gli articoli sono costitutive della valutazione, che devono essere riconosciute e governate attraverso disegni trasparenti, argomentabili e adeguati agli oggetti e alle finalità. È forse proprio questa

capacità di rendere visibili e discutibili le proprie scelte a tenere insieme la pluralità delle valutazioni senza rinunciare all'identità della valutazione.

Il numero 95 della RIV è il secondo pubblicato in formato open access nella storia della rivista e l'ultimo curato dal Comitato editoriale 2023-2026, composto da Giuseppe Moro (direttore), Veronica Lo Presti, Francesco Mazzeo Rinaldi, Erica Melloni, Francesca Pia Scardigno, Claudio Torrigiani e Giancarlo Vecchi. Cogliamo questa occasione per ringraziare i soci e le socie dell'AIV, il Direttivo AIV e l'editore FrancoAngeli per avere sostenuto in questi anni la crescita della Rassegna Italiana di Valutazione. Rivolgiamo i nostri migliori auguri di buon lavoro al nuovo Direttivo dell'Associazione e al prossimo Comitato editoriale della RIV, nella convinzione che sapranno non solo valorizzare ma anche estendere il lavoro svolto fin qui.

Erica Melloni*

Riferimenti bibliografici

- Alkin, M. (ed.) (2004). *Evaluation Roots*. Sage.
Rossi, P.H. (2004), *My Views of Evaluation and Their Origins*, in Alkin, M. (ed), 2004, *Evaluation Roots*, pp.122-131, Sage.

* Erica Melloni, Università Mercatorum (erica.melloni@unimercatorum.it)

Dall'output algoritmico al giudizio valutativo: governare l'evidenza nell'era dell'Intelligenza Artificiale

*Francesco Mazzeo Rinaldi**

This article critically examines the role of Artificial Intelligence (AI) in policy and program evaluation, focusing on the processes through which algorithmic outputs become evaluative evidence. Drawing on a critical narrative review of international literature published between 2016 and 2025, the paper identifies three main ways AI is conceptualized within evaluation: as an extension of analytical capacity, as an interpretive and ethical risk, and as an object of evaluation and governance. The analysis shows that these perspectives often leave implicit the inferential chain linking outputs, interpretation, and evaluative judgment. To address this gap, the article proposes an analytical framework structured around five stages - algorithmic output, technical validation, theory-informed interpretation, triangulation, and evaluative judgment - aimed at making visible potential forms of “epistemic compression” in AI-mediated evaluation processes.

Keywords: artificial Intelligence; evaluative evidence; algorithmic governance; critical narrative review.

Introduzione

Negli ultimi anni l'Intelligenza Artificiale (IA) ha acquisito una crescente centralità nel dibattito sulle scienze sociali applicate e, più in generale, sulle pratiche fondate sull'uso dell'evidenza (Athey & Berg, 2023; Sharma, 2025). Nel campo della valutazione di politiche e programmi, questa attenzione si è inizialmente concentrata sulle opportunità offerte dalla disponibilità di grandi volumi di dati e da strumenti computazionali sempre più sofisticati:

* Francesco MAZZEO RINALDI, Università di Catania (fmazzeo@unict.it)
Invio Proposta: 20-05-2026. Accettazione: 15-07-2026.

capacità di analisi su larga scala, automazione di compiti routinari, possibilità di osservare fenomeni sociali attraverso nuove fonti informative (Petersson & Breul, 2017). I Big Data e, successivamente, l'IA venivano interpretati come risorse potenzialmente capaci di rafforzare processi valutativi evidence-based e practice-based, a condizione di saper trasformare i dati grezzi in informazioni rilevanti per le diverse fasi del ciclo delle politiche (Mazzeo Rinaldi & Occhipinti, 2023). Tuttavia, rimaneva sullo sfondo una questione cruciale: quali sono i passaggi attraverso cui l'informazione prodotta mediante strumenti computazionali diventa evidenza valutativa? Il rapido sviluppo dell'IA, e in particolare dei modelli di apprendimento automatico e dei sistemi generativi, ha reso questa domanda non più eludibile. Oggi, una quota crescente di analisi, classificazioni, predizioni e persino testi interpretativi può essere prodotta da sistemi algoritmici con un livello di autonomia e complessità senza precedenti. In questo scenario, il problema centrale non riguarda tanto se utilizzare l'IA nella valutazione, quanto come governare il processo che conduce dagli output algoritmici al giudizio valutativo.

La letteratura internazionale più recente riflette chiaramente questa tensione (Nielsen et al, 2025; Leeuw, & Bamberger, 2025). Accanto a contributi che enfatizzano le potenzialità dell'IA in termini di efficienza, scala e tempestività (Shapiro & Lam, 2025), si moltiplicano le riflessioni critiche sui rischi epistemici ed etici associati all'uso di sistemi opachi, potenzialmente distorti e difficilmente auditabili (Azzam, 2023; Greenstein & Cho, 2025). Recentemente, anche la rivista *Evaluation* ha dedicato un numero speciale alle sfide della digitalizzazione nella pratica valutativa, affrontano questioni metodologiche, di ownership ed etiche nell'uso dell'IA e Big Data (Potluka et al., 2025). Ciò che appare meno sviluppato è un quadro concettuale integrato che chiarisca il ruolo specifico della valutazione in questo contesto: non come semplice utilizzatrice di nuovi strumenti, ma come campo di sapere responsabile della trasformazione dell'informazione in evidenza e dell'evidenza in giudizio.

È proprio su questo passaggio che si colloca il presente contributo. L'obiettivo non è valutare l'efficacia dell'IA né proporre linee guida tecniche per il suo impiego, ma interrogare il modo in cui gli output algoritmici vengono trasformati in evidenza valutativa. La tesi di fondo è che la sfida posta dall'IA non sia primariamente tecnica, bensì epistemica: riguarda il governo dei processi di inferenza, interpretazione e giudizio che definiscono la pratica valutativa. Attraverso una sintesi critica della letteratura valutativa recente, l'articolo mostra che le diverse modalità di concettualizzare l'IA condividono un punto cieco: i passaggi attraverso cui l'output computazionale diventa evidenza restano in larga misura impliciti. Per affrontare questo nodo,

viene proposto un framework analitico che distingue cinque stadi – dall’output algoritmico al giudizio valutativo – rendendo visibili le tensioni che attraversano ciascun passaggio.

Il contributo dell’articolo non consiste nell’offrire una nuova “buona pratica” per l’uso dell’IA, ma nel fornire una lente interpretativa per analizzare come, dove e a quali condizioni la catena inferenziale viene compressa, delegata o opacizzata. In un contesto in cui la produzione dell’evidenza è sempre più mediata da sistemi algoritmici, il futuro della valutazione dipende meno dalla capacità di adottare nuovi strumenti che dalla capacità di preservare e governare i processi che fondano la legittimità del giudizio.

1. Metodo: una *critical narrative review*

L’articolo adotta una *critical narrative review* della letteratura su IA e valutazione (2016–2025). Questo approccio non mira a una ricognizione esaustiva né a una sintesi quantitativa degli effetti, ma alla chiarificazione concettuale e alla problematizzazione delle assunzioni epistemiche che attraversano un campo emergente. Nel dibattito metodologico sulle revisioni della letteratura, le *critical narrative reviews* sono state distinte dalle review sistematiche e meta-analitiche per la loro funzione interpretativa e teorica. Grant e Booth (2009), nella loro tipologia delle revisioni, collocano le narrative e critical reviews tra gli approcci maggiormente orientati alla costruzione di senso (*sense-making*), alla messa in discussione delle assunzioni dominanti e allo sviluppo concettuale, piuttosto che alla sola aggregazione delle evidenze empiriche. In modo analogo, Saunders e Rojon (2011) sottolineano come una *critical literature review* non si limiti a descrivere lo stato dell’arte, ma si caratterizzi per la valutazione argomentata delle posizioni esistenti, l’identificazione di lacune concettuali e la riformulazione dei problemi di ricerca. In questo senso, l’elemento “critico” non riguarda la quantità delle fonti analizzate, bensì la profondità dell’analisi e la capacità di esplicitare le assunzioni epistemologiche che attraversano un determinato campo di studi.

Questa impostazione risulta particolarmente appropriata nel caso dell’IA applicata alla valutazione, un ambito in cui la letteratura è ancora fortemente eterogenea, interdisciplinare e caratterizzata da rapide evoluzioni tecnologiche. In tali contesti, come osservano Greenhalgh e colleghi (2018), l’idea di una gerarchia rigida che pone le review sistematiche al vertice e le narrative review a un livello inferiore risulta fuorviante. Gli autori argomentano che la scelta del tipo di review dovrebbe dipendere dalla natura della domanda di ricerca e dagli obiettivi conoscitivi perseguiti, riconoscendo il valore delle

narrative e critical reviews quando si tratta di concettualizzare campi emergenti, chiarire concetti ambigui o problematizzare quadri teorici consolidati. La rapida evoluzione delle applicazioni di IA in ambito valutativo e la natura ancora sperimentale di molte pratiche rendono inoltre necessario considerare le evidenze disponibili come provvisorie e contestuali, rafforzando l'opportunità di un approccio critico e interpretativo alla sintesi della letteratura (Greenhalgh *et al.*, 2018). Coerentemente con questa impostazione, la sintesi proposta in questo articolo non intende offrire una mappatura esaustiva di tutte le applicazioni dell'IA nella valutazione, né una valutazione comparativa delle loro prestazioni. L'obiettivo è piuttosto quello di:

- identificare i principali modi in cui l'IA viene concettualizzata nella letteratura valutativa;
- mettere in luce convergenze, tensioni e punti ciechi ricorrenti;
- interrogare criticamente il rapporto tra output algoritmici, produzione dell'evidenza e giudizio valutativo.

La selezione dei contributi analizzati è pertanto di tipo intenzionale (purposive) e si concentra su lavori pubblicati prevalentemente tra il 2016 e il 2025 che hanno svolto un ruolo significativo nello strutturare il dibattito internazionale. La strategia di selezione dettagliata, articolata in tre fasi, è descritta nella Sezione 1.1. Il corpus selezionato è stato analizzato attraverso alcune lenti critiche ricorrenti, coerenti con la tradizione metodologica della valutazione:

- lo statuto epistemico degli output algoritmici;
- le implicazioni per validità, affidabilità e trasparenza dell'evidenza prodotta;
- le conseguenze etiche e distributive dell'uso dell'IA;
- le ricadute sulla governance della valutazione e sull'identità professionale dei valutatori.

Su queste basi, la sintesi critica si concentra sull'analisi dei principali modi in cui l'IA viene concettualizzata nella letteratura valutativa. La sezione che segue esplicita la strategia di selezione delle fonti e il corpus analizzato.

1.1 Strategia di selezione e corpus analizzato

La selezione dei contributi è avvenuta attraverso una strategia di campionamento intenzionale progressivo (Paterson *et al.*, 2001), articolata in tre fasi:

- Identificazione di anchor texts. Sono stati assunti come testi di riferimento tre lavori che hanno strutturato il dibattito internazionale post-

ChatGPT: (a) lo special issue di *New Directions for Evaluation* curato da Mason e Montrose-Moorhead (2023), comprendente 10 saggi; (b) il volume Routledge curato da Nielsen, Mazzeo Rinaldi e Petersson (2025), composto da 14 capitoli; (c) e lo special issue della rivista *Evaluation* curato da Potluka, Harten, Kocks e Dvorak che raccoglie 8 contributi. I tre volumi sono stati analizzati integralmente e utilizzati come mappa iniziale del campo.

- Espansione attraverso reti citazionali (Google Scholar, Web of Science). A partire dagli anchor texts, sono stati individuati contributi ricorrentemente citati tramite analisi delle reti di citazione (Google Scholar, Web of Science), limitando la ricerca al periodo 2016–2025 e a riviste internazionali di valutazione.
- Selezione finale attraverso criteri di rilevanza epistemica. I testi sono stati selezionati in base a tre criteri: (a) rilevanza per il rapporto tra IA ed evidenza valutativa; (b) diversità di prospettive (strumentali, critiche, orientate alla governance); (c) originalità concettuale (framework, riflessioni teoriche), escludendo contributi orientati a meri casi applicativi.

Il corpus finale comprende diciassette contributi principali (Tabella 1), analizzati attraverso un processo iterativo in quattro stadi: (a) codifica aperta; (b) aggregazione in cluster tematici; (c) analisi comparativa per identificare convergenze, divergenze e lacune; (d) sintesi concettuale che ha generato il framework dei cinque stadi proposto nel paragrafo 4.

1.2. Limitazioni

Tre limitazioni meritano di essere esplicitate. In primo luogo, il focus sulla letteratura in lingua inglese introduce un bias linguistico e geografico che tende a privilegiare le prospettive provenienti dal Global North, con il rischio di sotto-rappresentare approcci e contributi elaborati in altri contesti. In secondo luogo, il focus sui tre *anchor texts* comporta un bias legato alla rete accademica di riferimento: questa impostazione favorisce una specifica comunità epistemica, prevalentemente AEA-centrica, e può escludere voci critiche alternative che circolano in altri ambiti disciplinari. Infine, va considerato il posizionamento dell'autore, che ha contribuito direttamente al dibattito analizzato. Ciò comporta il rischio di una selezione non neutrale delle fonti. Questo coinvolgimento viene tuttavia reso esplicito e assunto come parte integrante del metodo: se da un lato consente una comprensione profonda delle dinamiche interne al

campo, dall'altro richiede una costante vigilanza critica per evitare possibili bias confermativi. Queste limitazioni non invalidano l'approccio ma ne definiscono il perimetro di validità. Una critical narrative review non ambisce a rappresentatività statistica ma a profondità analitica e sviluppo concettuale (Greenhalgh *et al.*, 2018).

Tab. 1 - Corpus dei principali contributi analizzati

ANCHOR TEXTS

Autori	Anno	Focus tematico	Concettualizzazione IA
Mason & Montrosse-Moorhead (eds.)	2023	Special Issue <i>NDE</i> , Vol. 177-178: Evaluation and AI	Anchor text
Nielsen, Mazzeo Rinaldi & Petersson (eds.)	2025	Volume Routledge: IA and Evaluation	Anchor text
Potluka, Harten, Kocks & Dvorak (eds.)	2025	Special Issue <i>Evaluation: Digitalization in evaluations and evaluations of digitalization</i>	Anchor text

PRECURSORI

Autori	Anno	Focus tematico	Concettualizzazione IA
Bamberger	2016	Big Data e M&E	Precursore
Petersson & Breul	2017	Big Data e valutazione	Precursore
Mazzeo Rinaldi	2018	Big Data e valutazione (Italia)	Precursore
Picciotto	2020	Big Data: sfide per la valutazione	Precursore
York & Bamberger	2020	Big Data e valutazione	Precursore

CONTRIBUTI VARI (Special Issues/Routledge/...)

Autori	Anno	Focus tematico	Concettualizzazione IA
Nielsen	2023	Disruption valutativa – S.I.	Framework
Montrosse-Moorhead	2023	Criteri valutativi per IA – S.I.	Framework
Azzam	2023	Validità – S.I.	Rischio epistemico
York & Bamberger	2025	Applications of Big Data (Routledge)	Estensione strumentale
Greenstein & Cho	2025	Etica ed equità (Routledge)	Rischio, epistemico
Mazzeo Rinaldi & Nielsen	2025	Identità professionale (Routledge)	Trasformativo
Nielsen	2025	Industria valutativa e tecnologie emergenti (Routledge)	Governance + Estensione
Mazzeo Rinaldi & Occhipinti	2023	Big Data, IA e valutazione in Italia	Estensione strumentale + Rischio
Rompczyk	2025	Standard professionali	Rischio e opportunità
Shapiro & Lam	2025	Applicazioni pratiche	Estensione strumentale
Pudney <i>et al.</i>	2025	Valutazione sistemi IA	IA come oggetto
Wencker <i>et al.</i>	2025	Text as data e NLP	Estensione strumentale

Nota: I contributi di Bamberger (2016), Petersson & Breul (2017), Mazzeo Rinaldi (2018), Picciotto (2020) e York & Bamberger (2020) sono classificati come 'precursori' in quanto precedenti al rilascio di ChatGPT e rilevanti per contestualizzare il dibattito su Big Data, IA e valutazione.

2. Modi di concettualizzare l'IA nella letteratura valutativa

La letteratura recente sull'IA applicata alla valutazione non costituisce un campo unitario, ma si articola attorno a modalità ricorrenti di concettualizzazione che riflettono assunzioni differenti sul rapporto tra analisi, evidenza e giudizio. Tali modalità non sono semplici “posizioni” teoriche, bensì configurazioni epistemiche che orientano il modo in cui l'IA viene integrata nei processi valutativi e, soprattutto, il modo in cui gli output algoritmici vengono trattati come evidenza.

La sintesi critica del corpus consente di individuare tre configurazioni prevalenti: (a) l'IA come estensione strumentale; (b) l'IA come rischio epistemico ed etico; (c) l'IA come oggetto di valutazione e governance. Queste prospettive non sono mutuamente esclusive: spesso coesistono nello stesso contributo. Tuttavia, ciascuna tende a enfatizzare specifici stadi del processo che conduce dall'output al giudizio, lasciandone altri in ombra.

2.1 *L'IA come estensione strumentale della capacità analitica*

Un primo insieme di contributi interpreta l'IA prevalentemente come una estensione strumentale della capacità analitica della valutazione. In questa prospettiva, l'attenzione si concentra sulle opportunità offerte da machine learning, natural language processing e modelli generativi per trattare grandi volumi di dati, analizzare testi non strutturati, automatizzare compiti routinari e produrre sintesi rapide. Questa linea affonda le proprie radici nei lavori sui Big Data che precedono l'attuale ondata di IA generativa. Bamberger (2016) e Petersson e Breul (2017) già sottolineavano come la disponibilità di fonti digitali massive potesse ampliare scala, granularità e tempestività dei processi di monitoraggio e valutazione. York e Bamberger (2020) hanno sistematizzato tale prospettiva mostrando come analytics e tecnologie digitali possano rafforzare la misurazione dei risultati in contesti complessi. Nel contesto italiano, Mazzeo Rinaldi (2018) e Mazzeo Rinaldi e Occhipinti (2023) hanno evidenziato il ritardo della valutazione nell'ingaggiare questo dibattito, ma anche il potenziale trasformativo di tali strumenti.

Nel corpus più recente, questa configurazione assume una forma esplicitamente orientata all'IA. Wencker *et al.*, (2025) mostrano come NLP e large language models consentano di trattare corpus testuali di dimensioni prima inaccessibili, generando nuove forme di “text as data” per l'analisi valutativa. Shapiro e Lam (2025) illustrano applicazioni pratiche dell'IA in diverse fasi del ciclo valutativo, dalla raccolta dei dati alla sintesi dei risultati. Nielsen

(2023) interpreta le tecnologie emergenti come fattori di disruption dell'industria valutativa, capaci di automatizzare una quota crescente di compiti analitici e redazionali. In questi contributi, l'IA è prevalentemente concettualizzata come infrastruttura abilitante: uno strumento che estende ciò che il valutatore può fare in termini di velocità, scala e copertura empirica. Il lessico ricorrente è quello dell'efficienza, dell'ampliamento del raggio osservativo e della riduzione dei costi cognitivi e operativi.

Dal punto di vista analitico, questa configurazione tende a concentrare l'attenzione sugli stadi 1 e 2 del framework: produzione dell'output e sua validazione tecnica. Gli autori discutono in modo approfondito le potenzialità analitiche degli strumenti e, in misura crescente, i problemi di accuratezza e bias. Rimangono invece in larga parte impliciti i passaggi successivi – interpretazione theory-informed, triangolazione, giudizio.

In York e Bamberger (2020, 2025), ad esempio, l'IA e gli analytics sono presentati come mezzi per rafforzare la capacità della valutazione di osservare fenomeni complessi e dinamici. Tuttavia, il modo in cui tali output vengono trasformati in evidenza valutativa è dato per scontato: l'enfasi è sulla produzione di informazione più ricca, non sul governo del passaggio dall'informazione al giudizio. Analogamente, in Wencker *et al.* (2025) l'attenzione è rivolta alle tecniche di estrazione di pattern semantici, mentre resta sullo sfondo il problema di come tali pattern debbano essere interpretati alla luce di teorie di programma o criteri valutativi.

Nell'articolo introduttivo dello special issue di Evaluation dedicato alla digitalizzazione delle valutazioni Potluka *et al.* (2025), identificano tra le sfide principali la necessità di sviluppare conoscenze metodologiche adeguate per evitare approcci “black box”. Questa preoccupazione si colloca proprio nello spazio coperto dal filone strumentale: l'enfasi sulla potenza analitica richiede una corrispondente attenzione ai processi che trasformano tale potenza in comprensione valutativa. Tuttavia, come negli altri contributi

di questo filone, i passaggi attraverso cui tale trasformazione avviene restano in larga parte da esplicitare.

In questa prospettiva, l'evidenza tende implicitamente a coincidere con l'informazione prodotta dall'algoritmo. La sofisticazione tecnica diventa una garanzia tacita di qualità dell'evidenza. Il rischio principale non è l'errore puntuale, ma la naturalizzazione dell'output come “fatto”. Il passaggio inferenziale che trasforma un pattern computazionale in un'affermazione valutativa rimane opaco.

Questa configurazione illumina con forza il potenziale dell'IA per ampliare le capacità analitiche della valutazione, ma oscura il lavoro interpretativo che rende tali capacità rilevanti per il giudizio. In termini di framework, essa

rafforza i primi due stadi e lascia in ombra i successivi. È in questa compressione che si annida il punto cieco comune alla letteratura: l'assunzione che una migliore analisi produca automaticamente una migliore evidenza.

2.2 L'IA come rischio epistemico ed etico

Un secondo filone della letteratura si sviluppa in chiave critica e interpreta l'IA principalmente come una fonte di rischio per la produzione dell'evidenza valutativa. In questa prospettiva, l'attenzione si concentra sui limiti strutturali dei sistemi algoritmici: bias nei dati di addestramento, opacità dei modelli, automation bias, difficoltà di spiegazione, asimmetrie di potere e potenziali effetti di tecnocratizzazione.

Questo orientamento affonda le proprie radici nei lavori che, già nel contesto dei Big Data, mettevano in guardia contro una lettura ingenuamente positivista dell'innovazione digitale. Picciotto (2020) ha sottolineato come la disponibilità di grandi volumi di dati non risolva problemi fondamentali di veracity, volatility e validity, evidenziando il rischio che l'apparente potenza analitica sostituisca una riflessione sui costrutti valutativi. In modo analogo, Bamberger (2016) osservava che la validazione tecnica degli output non equivale a una validazione valutativa dell'evidenza.

Nel dibattito più recente sull'IA, queste preoccupazioni si intensificano. Azzam (2023) discute in modo sistematico le implicazioni dell'IA per la validità, mostrando come modelli opachi possano produrre risultati formalmente coerenti ma epistemicamente fragili. Greenstein e Cho (2025) collocano l'uso dell'IA entro un quadro di etica ed equità, evidenziando come bias algoritmici e automation bias possano rafforzare disuguaglianze esistenti e ridurre la trasparenza dei processi valutativi. Rompczyk (2025) interroga la compatibilità tra l'uso dell'IA e gli standard professionali della valutazione, mettendo in luce il rischio di un allentamento dei principi di indipendenza, trasparenza e responsabilità. Potluka *et al.* (2025) evidenziano l'urgenza di criteri normativi che governino l'uso dell'IA, riconoscendo implicitamente che i rischi epistemici ed etici non possono essere risolti solo sul piano tecnico. Anche qui, tuttavia, il focus resta prevalentemente sui primi stadi del framework: identificare i rischi più che governare l'intero percorso inferenziale.

In questa configurazione, l'IA si presenta come un elemento ambiguo per la pratica valutativa: introduce opacità dove la valutazione rivendica trasparenza, automatismo dove è richiesto giudizio ed efficienza dove servono tempi di riflessione. Il baricentro dell'argomentazione si colloca sugli stadi 1 e 2 del

framework, ma in senso rovesciato rispetto al filone strumentale: non per esaltarne le potenzialità, bensì per metterne in luce le fragilità.

Questi contributi svolgono una funzione cruciale di problematizzazione. Rendono visibili i rischi epistemici ed etici che accompagnano l'utilizzo dell'IA e contrastano la naturalizzazione dell'output algoritmico come evidenza neutrale. Tuttavia, essi tendono a collocare l'IA come elemento esterno rispetto alla pratica valutativa, rispetto al quale adottare una postura prevalentemente difensiva.

Dal punto di vista del framework, questa letteratura illumina con forza le tensioni degli stadi iniziali, opacità dell'output, confusione tra performance e validità, ma lascia relativamente indeterminato il modo in cui tali rischi possano essere governati all'interno di disegni valutativi concreti. L'attenzione è rivolta più ai pericoli che ai processi attraverso cui l'evidenza viene costruita.

Ne risulta una prospettiva che, pur offrendo diagnosi convincenti, fatica a tradursi in una grammatica operativa per la valutazione. Il passaggio dall'output al giudizio resta un problema evocato ma non articolato. L'IA è qualcosa da contenere, limitare o sorvegliare, più che un oggetto da integrare entro un percorso interpretativo. In termini di prospettiva analitica, questa configurazione rende visibili le fragilità dei primi due stadi, ma lascia in ombra quelli successivi: interpretazione, triangolazione e giudizio. La critica resta concentrata sull'algoritmo, mentre il lavoro interpretativo della valutazione rimane sullo sfondo. È proprio questo scarto che apre lo spazio per una concettualizzazione più integrata del governo dell'evidenza.

2.3 L'IA come oggetto di valutazione e governance

Un terzo insieme di contributi propone un cambio di prospettiva più radicale, concettualizzando l'IA come oggetto diretto di valutazione. In questo filone, la comunità valutativa rivendica un ruolo specifico: non si limita a utilizzare l'IA come strumento, ma mette a disposizione criteri, framework e approcci metodologici per analizzarne funzionamento, impatti e implicazioni sociali. Questa prospettiva è particolarmente visibile nei contributi che applicano approcci standards e criteria based all'analisi dei sistemi di IA, enfatizzando dimensioni quali spiegabilità, equità, affidabilità e coerenza con obiettivi di policy (Montrosse-Moorhead, 2023).

In modo analogo, Pudney *et al.* (2025) sviluppano un quadro per la valutazione dei sistemi di IA in ambito pubblico, insistendo sulla necessità di integrare criteri tecnici con criteri normativi e sociali. Nielsen (2025) colloca

esplicitamente la valutazione come attore chiave nella governance dell'IA, capace di connettere performance tecnica, impatti sociali e decisioni di policy.

In questa prospettiva, la valutazione recupera con forza il proprio mandato normativo. Gli stadi 4 e 5 del framework – triangolazione e giudizio – emergono come centrali: l'IA viene sottoposta a confronto con altre fonti di evidenza, con valori pubblici e con aspettative degli stakeholder, e diventa oggetto di decisioni esplicite. Rispetto ai due filoni precedenti, questa configurazione sposta il baricentro dall'uso dell'IA alla sua regolazione. L'attenzione non è più rivolta solo a ciò che l'IA consente di fare, né esclusivamente ai rischi che introduce, ma alla possibilità di governarla attraverso pratiche valutative.

Tuttavia, anche in questo filone il passaggio dagli output algoritmici all'evidenza tende a rimanere implicito. I sistemi di IA vengono trattati come oggetti complessi da giudicare, ma raramente si tematizza il modo in cui, nelle valutazioni ordinarie, gli output prodotti da tali sistemi attraversino l'intera catena inferenziale. In altre parole, la prospettiva della governance rafforza la dimensione del giudizio, ma non rende necessariamente visibile il percorso che conduce dall'analisi computazionale a tale giudizio. Il rischio è che la valutazione si collochi a valle dei sistemi, pronunciandosi sui loro impatti, senza interrogare sistematicamente il modo in cui l'evidenza algoritmica viene costruita e utilizzata nei processi valutativi quotidiani.

In termini di framework, questa configurazione illumina gli stadi finali – triangolazione e giudizio – ma lascia relativamente in ombra quelli intermedi: interpretazione theory-informed e costruzione progressiva dell'evidenza. La governance si esercita sull'oggetto "IA", mentre la catena inferenziale che trasforma output in evidenza resta in larga parte tacita. È proprio questa asimmetria che completa il quadro: ciascun filone rende visibili alcune dimensioni e ne oscura altre. Nessuno, preso singolarmente, tematizza l'intero percorso che conduce dall'output algoritmico al giudizio valutativo.

2.4 Frammentazione concettuale e punto cieco comune

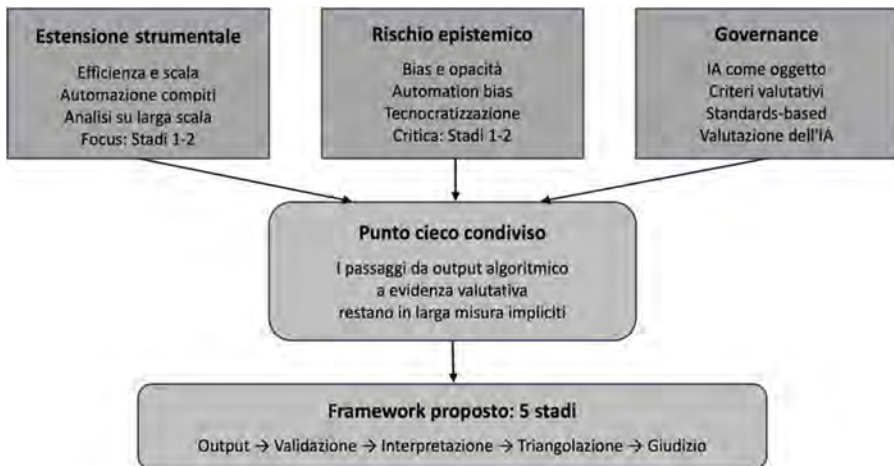
Considerate nel loro insieme, queste tre modalità di concettualizzare l'IA mostrano come la letteratura valutativa sia caratterizzata da una significativa frammentazione concettuale. I contributi differiscono per obiettivi, linguaggi disciplinari e livelli di astrazione, ma condividono un limite ricorrente: la scarsa attenzione ai passaggi attraverso cui gli output algoritmici vengono trasformati in evidenza valutativa. Ogni prospettiva illumina alcuni stadi del framework e ne oscura altri:

- l'estensione strumentale privilegia output e validazione tecnica;

- la prospettiva critica problematizza opacità e bias ma resta esterna ai processi;
- l'IA come oggetto di valutazione enfatizza giudizio e governance senza esplicitare la catena inferenziale.

Il risultato è che il percorso che conduce dall'analisi computazionale al giudizio rimane spesso in larga misura implicito, come rappresentato nella schema sottostante (fig. 1). In questo spazio ancora poco tematizzato si colloca il contributo di questo articolo, che si propone di rendere più espliciti e discutibili i passaggi attraverso cui l'informazione algoritmica viene tradotta in evidenza valutativa.

Fig. 1 - Tre modalità di concettualizzare l'IA nella valutazione



La frammentazione appena descritta suggerisce che il problema non risieda tanto nella presenza dell'IA in sé, quanto nell'assenza di una concettualizzazione condivisa dei passaggi che trasformano l'analisi in evidenza. Per rendere questo nodo maggiormente osservabile e discutibile, appare quindi necessario spostare l'attenzione dagli strumenti ai processi inferenziali che li attraversano.

3. Dagli output algoritmici all'evidenza: il problema

L'evidenza valutativa non coincide con la mera disponibilità di dati o risultati analitici, ma è il prodotto di un processo che integra dati empirici, quadri interpretativi, criteri espliciti e valori normativi. Già Bamberger (2016) e

Picciotto (2020) hanno sottolineato come la qualità dell'evidenza non sia riducibile alla qualità tecnica delle misure, ma dipenda dalla capacità di collegare osservazioni, teoria e giudizio. L'introduzione dell'IA non elimina tali passaggi, ma tende piuttosto a renderli meno visibili, aumentando il rischio di una sovrapposizione impropria tra output algoritmico ed evidenza (Mazzeo Rinaldi & Nielsen, 2025). Classificazioni, cluster, predizioni e testi generati possono essere utilizzati direttamente per informare conclusioni valutative, producendo una compressione epistemica in cui i passaggi interpretativi, inferenziali e valutativi tendono a essere implicitamente fusi: l'output algoritmico viene così tacitamente assimilato all'evidenza. In questo slittamento non risiede semplicemente un rischio tecnico, ma una trasformazione del modo in cui l'inferenza valutativa viene costruita e legittimata.

Le preoccupazioni espresse da Picciotto (2020) rispetto ai Big Data trovano qui una nuova intensità: la potenza analitica può talvolta ridurre lo spazio dedicato alla riflessione sui costrutti e sui criteri di valore. La distinzione tra "Evaluation needs Big Data" e "Big Data needs evaluation" richiama la necessità di un governo metodologico bidirezionale dei passaggi che conducono dall'output all'evidenza. Azzam (2023) e Greenstein e Cho (2025) mostrano come l'opacità algoritmica e l'automation bias rafforzino questa dinamica, rendendo meno trasparenti i passaggi interpretativi che tradizionalmente fondano la credibilità dell'evidenza.

Il problema centrale non è dunque se l'IA produca risultati accurati, ma come tali risultati vengano trasformati in affermazioni valutative. Quando l'output computazionale viene trattato come evidenza autosufficiente, la valutazione rischia di perdere la propria specificità interpretativa, riducendosi a un'estensione dell'analisi dei dati. Per discutere questo passaggio, può essere utile distinguere analiticamente cinque stadi attraverso cui un risultato computazionale può essere trasformato in evidenza valutativa:

- Produzione dell'output algoritmico
- Validazione tecnica
- Interpretazione theory-informed
- Triangolazione
- Giudizio valutativo

Questi stadi non descrivono una sequenza operativa rigida, ma individuano dimensioni epistemiche che possono essere rese più visibili e maggiormente tematizzate. In questo senso, essi ricompongono in forma esplicita aspetti che nella tradizione valutativa risultano spesso distribuiti tra disegno, analisi e uso dell'evidenza (Bamberger, 2016; York & Bamberger, 2020; Picciotto, 2020).

La loro mancata esplicitazione non costituisce soltanto un limite procedurale, ma può incidere sulla trasparenza e sulla legittimazione dell'evidenza prodotta. Rendere esplicita la catena che conduce dall'output al giudizio significa, in questa prospettiva, riaffermare che anche nell'era dell'IA, la valutazione resta una pratica di interpretazione e attribuzione di valore, non una mera funzione di calcolo.

4. Il framework: cinque stadi tra funzione e tensione epistemica

Il framework proposto non descrive una successione standard da applicare meccanicamente a ogni valutazione. Piuttosto, prova a rendere osservabili cinque dimensioni che, nella pratica, possono essere fuse, compresse o solo parzialmente esplicitate. La sua funzione non è prescrittiva, ma analitica: offrire una possibile prospettiva per interrogare il passaggio dall'output algoritmico al giudizio valutativo. Ciascuno stadio individua una funzione che può contribuire alla trasformazione dell'informazione in evidenza, ma richiama anche alcune tensioni che, se non adeguatamente considerate, possono incidere sulla trasparenza e legittimità del processo valutativo.

Tab. 2 - Cinque stadi della trasformazione dell'output algoritmico in evidenza valutativa

<i>Stadio</i>	<i>Funzione analitica</i>	<i>Possibile compressione epistemica</i>
<i>Output algoritmico</i>	Produce pattern, stime, classificazioni o testi	Naturalizzazione dell'output come dato
<i>Validazione tecnica</i>	Verifica accuratezza, robustezza e bias	Confusione tra performance tecnica e validità valutativa
<i>Interpretazione theory-informed</i>	Collega l'output a teoria di programma, contesto e meccanismi	Riduzione del significato a correlazione o pattern
<i>Triangolazione</i>	Confronta l'output con fonti, metodi e prospettive diverse	Autosufficienza dell'algoritmo
<i>Giudizio valutativo</i>	Formula conclusioni di valore sulla base di criteri espliciti	Delega implicita del giudizio

Un esempio illustrativo può aiutare a chiarire come i diversi stadi del framework possano intrecciarsi nella pratica valutativa e dove si producano possibili forme di compressione epistemica.

Box 1 – Output algoritmico ed evidenza valutativa

Si consideri una valutazione di un programma di formazione professionale che utilizza strumenti di natural language processing per analizzare migliaia di commenti aperti raccolti attraverso questionari online. Il sistema algoritmico produce cluster tematici e indicatori di sentiment, identificando automaticamente pattern ricorrenti nelle esperienze dei partecipanti.

In assenza di ulteriori passaggi interpretativi, tali output possono essere assunti direttamente come evidenza della qualità del programma. Tuttavia, il fatto che un commento venga classificato come “positivo” o associato a un determinato tema non chiarisce necessariamente il significato valutativo di tali risultati, né il modo in cui essi si colleghino agli obiettivi del programma, ai meccanismi di cambiamento attesi o alle differenze tra gruppi di partecipanti. È proprio in questo spazio tra output computazionale e giudizio valutativo che si collocano i passaggi di interpretazione, triangolazione e attribuzione di valore che il framework proposto mira a rendere maggiormente visibili.

4.1 Output algoritmico: potenza analitica e opacità incorporata

Il primo stadio riguarda la produzione dell’output da parte del sistema di IA. Che si tratti di classificazioni, cluster, stime predittive o testi generati, tali output sono il risultato di una serie di scelte: selezione delle variabili, definizione delle categorie, metriche di ottimizzazione, architettura del modello, dati di addestramento.

Dal punto di vista valutativo, il punto cruciale non è tanto cosa produce l’algoritmo, quanto come tali risultati incorporino assunzioni che restano spesso invisibili. Nei modelli di machine learning tradizionali, queste assunzioni riguardano il modo in cui il fenomeno viene reso computabile (Wencker *et al.*, 2025); nei modelli generativi, esse si manifestano nella capacità di produrre output semanticamente plausibili anche in assenza di un ancoraggio empirico verificabile.

La letteratura sull’IA ha ampiamente mostrato che l’output algoritmico non è mai neutro: esso riflette distribuzioni di dati, scelte di design e trade-off tra accuratezza, semplicità e generalizzabilità. Sul piano valutativo, ciò implica che l’output non può essere trattato come “dato grezzo”. Esso è già una forma di interpretazione automatizzata.

La tensione strutturale di questo stadio risiede nell’opacità: quanto più l’output appare sofisticato e fluente, tanto più è difficile ricostruirne le premesse. In assenza di un lavoro di esplicitazione, l’output tende a essere assunto come evidenza in sé, innescando un processo di trasformazione che salta i passaggi intermedi.

4.2 Validazione tecnica: necessaria ma non sufficiente

Il secondo stadio riguarda la validazione tecnica dell'output algoritmico. In questa fase vengono valutate accuratezza, stabilità, robustezza e presenza di bias sistematici. Nella letteratura sulla data science, questo livello di analisi è ampiamente sviluppato e rappresenta un prerequisito fondamentale per qualsiasi utilizzo responsabile degli algoritmi (Schwartz *et al.*, 2022). Tuttavia, dal punto di vista valutativo, esso rappresenta una condizione necessaria, ma non sufficiente.

Un modello può risultare altamente performante rispetto alle metriche adottate e, al tempo stesso, produrre informazioni poco rilevanti o fuorvianti rispetto alle domande che la valutazione è chiamata ad affrontare. La tensione strutturale di questo stadio risiede proprio nella possibile sovrapposizione tra performance tecnica e validità valutativa.

Questa distinzione era già stata evidenziata da Bamberger (2016) nel contesto dei Big Data, quando osservava che la validazione tecnica degli output non equivale a una validazione valutativa dell'evidenza prodotta. L'IA tende a rendere questa distinzione meno visibile: l'eccellenza tecnica può essere scambiata per solidità dell'evidenza. Il framework consente di localizzare questo slittamento: l'evidenza diventa fragile non perché l'algoritmo “funziona male”, ma perché la validazione si arresta al livello tecnico.

4.3 Interpretazione contestuale e theory-informed

Il terzo stadio introduce un passaggio tipicamente valutativo: l'interpretazione degli output algoritmici alla luce del contesto e della teoria di programma. In questa fase, i risultati computazionali cessano di essere meri pattern e vengono collocati all'interno di un quadro esplicativo più ampio.

La letteratura valutativa sulla theory-based ha da tempo sottolineato che la comprensione dei meccanismi causali e delle condizioni contestuali è essenziale per attribuire significato ai risultati empirici (Chen, 1990; Pawson & Tilley, 1997). In assenza di questo lavoro interpretativo, l'IA rischia di rafforzare un uso meramente correlazionale dell'evidenza, privilegiando pattern predittivi rispetto alla comprensione dei processi sottostanti poiché i modelli tradizionali di machine learning eccellono nell'identificazione di correlazioni ma non distinguono intrinsecamente se queste rappresentino relazioni causali (Pearl & Mackenzie, 2018). La tensione strutturale di questo stadio è dunque la riduzione correlazionale del significato. Un secondo esempio illustrativo

consente di osservare questa tensione nel caso dell'analisi qualitativa assistita da IA generativa.

Box 2 – IA generativa e interpretazione qualitativa

Si consideri una valutazione qualitativa basata su interviste in profondità a beneficiari di un programma sociale. Un large language model viene utilizzato per codificare automaticamente le trascrizioni, identificare temi ricorrenti e produrre sintesi interpretative.

In questo caso, il sistema non si limita a organizzare informazioni, ma partecipa direttamente alla costruzione del significato analitico. Le categorie generate possono apparire coerenti e plausibili, inducendo il valutatore ad assumerle implicitamente come interpretazioni già valide.

Tuttavia, il passaggio tra similarità linguistica e significato valutativo rimane teoricamente e contestualmente mediato. Le categorie prodotte dall'IA possono riflettere regolarità semantiche senza cogliere meccanismi causali, relazioni di potere o dimensioni normative rilevanti per il giudizio.

In questo senso, il framework proposto aiuta a distinguere il momento della produzione automatizzata di pattern dal lavoro interpretativo attraverso cui tali pattern vengono trasformati in evidenza valutativa.

Quando l'interpretazione theory-informed è assente o debole, l'output viene letto come risposta autosufficiente. In questo modo, la potenza predittiva sostituisce la comprensione dei processi. Il framework rende visibile questo slittamento: ciò che manca non è informazione, ma significato valutativo.

4.4 Triangolazione e confronto tra fonti

Il quarto stadio riguarda la triangolazione con altre fonti, metodi e prospettive. In una prospettiva valutativa, nessuna fonte informativa, algoritmica o meno, può essere considerata autosufficiente, la credibilità emerge dal confronto (Picciotto, 2020). La letteratura valutativa mostra come la triangolazione rappresenti un meccanismo chiave per ridurre il rischio di interpretazioni distorte e per rafforzare la credibilità dell'evidenza (Patton, 1999). L'IA introduce una tensione specifica: la pressione all'autosufficienza dell'algoritmo. La scala, la velocità e la granularità degli output possono indurre una fiducia sproporzionata, rafforzata da fenomeni di automation bias (Goddard *et al.*, 2012; Parasuraman & Manzey, 2010). La triangolazione diventa allora non solo una buona pratica, ma un presidio critico contro l'autosufficienza dell'output algoritmico. Come mostrato anche negli esempi illustrativi precedenti, la disponibilità di output altamente granulari o semanticamente plausibili può rafforzare la tendenza a trattare l'analisi algoritmica come fonte autosufficiente di

evidenza. Il framework consente di analizzare quando e come l'output algoritmico viene sottratto al confronto, diventando una fonte privilegiata o esclusiva. In tali casi, la fragilità dell'evidenza non dipende dall'assenza di dati, ma dalla riduzione del pluralismo informativo.

4.5 Giudizio valutativo: il nodo della responsabilità

Il quinto stadio riguarda la formulazione del giudizio valutativo, in cui l'evidenza viene messa in relazione con criteri espliciti, valori e standard condivisi. Questo passaggio implica una responsabilità normativa e istituzionale che non può essere delegata ai sistemi di IA senza snaturare la funzione stessa della valutazione (Montrosse-Moorhead, 2023; Nielsen, 2025). La tensione strutturale risiede nella delegabilità implicita del giudizio. Quando le conclusioni sono fortemente mediate da sistemi opachi, il confine tra supporto decisionale e sostituzione del giudizio tende a sfumare. Il rischio non è che l'IA "decida", ma che il giudizio venga presentato come tecnicamente necessario. Il framework rende visibile questo punto: il giudizio non è un esito automatico dell'analisi. È un atto che implica responsabilità. Rendere esplicito lo stadio finale significa riaffermare che, anche nell'era dell'IA, la valutazione resta una pratica di attribuzione di valore.

5. Implicazioni metodologiche per la valutazione

Il framework proposto non mira soltanto a chiarire un nodo epistemico astratto, ma suggerisce anche alcune possibili implicazioni per il modo in cui la valutazione viene progettata, condotta e utilizzata in contesti mediati da sistemi algoritmici. Tali implicazioni non riguardano tanto l'adozione di specifici strumenti, quanto l'attenzione ai processi attraverso cui si passa dall'output al giudizio valutativo. In questa prospettiva, le principali implicazioni possono essere ricondotte a quattro ambiti: il disegno valutativo; la costruzione dell'evidenza; l'uso dell'evidenza e le relazioni con gli stakeholders; il ruolo professionale del valutatore.

5.1 Implicazioni per il disegno valutativo

L'integrazione dell'IA nei processi valutativi incide, anzitutto, sul disegno della valutazione. L'uso di sistemi algoritmici tende infatti a spostare

l'attenzione verso fasi analitiche avanzate, rischiando di comprimere o dare per scontate le scelte di scoping, di formulazione delle domande valutative e di esplicitazione delle assunzioni teoriche. In contesti caratterizzati da abbondanza di dati e potenza computazionale, il disegno rischia di diventare reattivo agli strumenti disponibili: ciò che è misurabile o modellizzabile orienta ciò che viene considerato rilevante. In questo senso, il framework può offrire una chiave interpretativa utile a rendere più osservabile tale slittamento (York & Bamberger, 2020). Quando gli stadi relativi all'interpretazione, alla triangolazione e al giudizio risultano deboli o poco esplicitati, il processo valutativo rischia infatti di essere orientato prevalentemente dall'analisi algoritmica, con domande valutative che emergono ex post a partire da ciò che l'algoritmo riesce a identificare. In questa prospettiva, la funzione del disegno valutativo si rafforza: esso diventa un vincolo epistemico che precede e orienta l'uso dell'IA. Progettare una valutazione in contesti algoritmici può quindi implicare la necessità di chiarire:

- quali domande di valore devono essere affrontate;
- quali dimensioni del programma risultano osservabili tramite IA e quali tendono invece a rimanere fuori campo;
- in che modo gli output computazionali vengano collegati a una teoria di programma esplicita.

L'IA, in questo senso, non sembra ridurre l'importanza del disegno valutativo; al contrario, può rendere ancora più rilevante la necessità di esplicitarne presupposti, criteri e finalità. Il framework può contribuire a individuare situazioni in cui il disegno appare fortemente orientato dagli strumenti disponibili e a riportare l'attenzione sulle finalità valutative che guidano il processo.

Sul piano pratico, questo comporta anche una dimensione relazionale: la definizione di domande valutative e teoria di programma non è solo un esercizio interno al valutatore, ma un confronto esplicito con committenti e stakeholder, tanto più necessario quando l'IA rende disponibili nuove fonti o granularità di dati capaci di orientare surrettiziamente il fuoco dell'indagine.

5.2 Costruzione dell'evidenza: validità oltre la performance

Un secondo insieme di implicazioni riguarda la validità dell'evidenza valutativa prodotta attraverso sistemi di IA. Come discusso nella sezione precedente, la validazione tecnica degli output rappresenta solo uno dei livelli necessari per garantire la qualità dell'evidenza. Dal punto di vista valutativo, la qualità dell'evidenza dipende dall'allineamento tra ciò che viene modellizzato

e i costrutti teorici e normativi rilevanti per il giudizio (Patton, 2008; Whittemore *et al.*, 2023). In questo senso, l'uso dell'IA tende a rendere più visibili alcune tensioni classiche della valutazione, in particolare quella tra precisione analitica e rilevanza per il giudizio. Output altamente performanti possono infatti risultare epistemicamente fragili quando i passaggi di interpretazione e triangolazione rimangono deboli o poco esplicitati. Nei casi illustrativi discussi, tale fragilità non deriva necessariamente da errori tecnici dell'algoritmo, ma dalla parziale esplicitazione dei passaggi che collegano pattern computazionali e giudizio valutativo. In questa prospettiva, il framework può aiutare a individuare dove si collocano tali fragilità: non necessariamente nel funzionamento dell'algoritmo, quanto piuttosto nella parziale esplicitazione dei processi che collegano l'output al giudizio valutativo.

Sul piano operativo, ciò suggerisce l'importanza che la documentazione valutativa renda maggiormente espliciti:

- le assunzioni incorporate nei modelli;
- i criteri attraverso cui l'output viene interpretato;
- le fonti con cui viene triangolato;
- il modo in cui tali elementi informano il giudizio.

In assenza di questa esplicitazione, l'uso dell'IA potrebbe favorire forme dinamiche di tecnocratizzazione del processo valutativo, riducendo lo spazio per il confronto critico e l'uso riflessivo dell'evidenza. Sul piano pratico, ciò richiede di riconsiderare le modalità di triangolazione, affiancando sistematicamente agli output algoritmici metodi qualitativi e approcci interpretativi – interviste, focus group, osservazione – non come verifica accessoria, ma come componente costitutiva della costruzione dell'evidenza, da concordare con i committenti fin dal disegno.

5.3 Uso dell'evidenza e relazioni con gli stakeholder

Un ulteriore insieme di implicazioni riguarda l'uso dell'evidenza e le relazioni con gli stakeholder. La produzione algoritmica di output ad alta granularità può infatti rafforzare forme di razionalizzazione tecnica, riducendo lo spazio per il confronto interpretativo e per la discussione sui criteri di valore sottesi ai risultati.

In questa prospettiva, il framework suggerisce l'opportunità di considerare l'uso dell'evidenza come parte integrante della catena inferenziale. Rendere maggiormente visibili i diversi stadi può contribuire a trasformare l'output algoritmico in un oggetto di discussione e interpretazione, piuttosto che in

“autorità silenziosa”. In particolare, interpretazione e triangolazione possono essere lette come spazi deliberativi in cui diversi attori possono interrogare significato, limiti e implicazioni dei risultati.

Da questo punto di vista, l’IA non conduce inevitabilmente alla chiusura del processo valutativo; può, se governata, aprire nuove forme di dialogo basate su evidenze più ricche ma anche più controverse.

Questo implica, sul piano pratico, di ricondurre al confronto con gli stakeholder anche momenti che la prassi valutativa tende a dare per acquisiti a monte, come la definizione della teoria di programma e dei criteri di giudizio, tanto più quando questi vengono messi in tensione da output algoritmici ad alta granularità.

5.4 Ruolo professionale del valutatore

Il quarto ambito riguarda il ruolo professionale del valutatore e la distribuzione delle responsabilità nei processi valutativi mediati dall’IA. L’introduzione di sistemi algoritmici tende infatti a ridefinire, almeno in parte, la distribuzione del lavoro cognitivo, spostando alcune attività analitiche verso strumenti automatizzati e richiedendo al tempo stesso competenze di supervisione, interpretazione e governance.

In questa prospettiva, il ruolo del valutatore non sembra necessariamente ridursi, quanto piuttosto trasformarsi. La letteratura più recente sottolinea, in particolare, come tale evoluzione implichi anche nuove competenze: alfabetizzazione digitale, capacità di interrogare criticamente gli output algoritmici e collaborazione interdisciplinare, con attenzione a privacy, bias e responsabilità (Montrosse-Moorhead, 2023; Mazzeo Rinaldi & Nielsen, 2025; Nielsen, 2025).

In questo scenario, il contributo specifico del valutatore potrebbe risiedere non tanto nella produzione di analisi sempre più sofisticate, quanto nella capacità di rendere espliciti e governare i passaggi che collegano i diversi stadi del processo valutativo. In questo senso, il valutatore può essere interpretato come una figura di mediazione interpretativa, chiamata a esplicitare assunzioni, limiti, criteri interpretativi e responsabilità.

Il framework, da questo punto di vista, può aiutare a individuare situazioni in cui tale funzione tende a indebolirsi: ad esempio quando l’output algoritmico viene fatto convergere direttamente verso conclusioni valutative, riducendo la visibilità dei passaggi interpretativi intermedi. La questione, allora, non riguarda tanto l’uso o meno dell’IA, quanto le modalità attraverso cui viene governato il passaggio dall’informazione al giudizio.

Le competenze richieste non si esauriscono quindi nell'alfabetizzazione digitale individuale, ma includono la capacità di gestire il rapporto con gli stakeholder attorno a questi passaggi – teoria di programma, criteri, triangolazione – ridefinendo il ruolo del valutatore anche come funzione di mediazione tra analisi algoritmica e comunità degli interlocutori della valutazione.

Le implicazioni metodologiche discusse mostrano come la prospettiva analitica proposta possa contribuire a riorientare il disegno, la costruzione e l'uso dell'evidenza. Resta tuttavia importante interrogare criticamente anche il framework stesso, le sue condizioni di applicabilità e i limiti che lo attraversano. La sezione seguente affronta queste questioni.

6. Discussione critica del framework

Il framework proposto non va inteso come una sequenza operativa da applicare integralmente a ogni valutazione. La sua finalità è prevalentemente analitica: offrire una chiave di lettura per rendere più visibili alcune dimensioni epistemiche che tendono a rimanere implicite quando l'evidenza è mediata da sistemi algoritmici. Uno dei rischi principali è interpretare il framework come una definizione di “buona pratica”. I cinque stadi non descrivono necessariamente ciò che dovrebbe accadere in ogni valutazione, ma ciò che può essere distinto quando si analizza il passaggio dall'output al giudizio. Nella pratica, tali stadi risultano spesso intrecciati: l'interpretazione può avvenire durante la validazione tecnica; la triangolazione può essere incorporata in procedure automatizzate; il giudizio può essere parzialmente inscritto nelle metriche utilizzate. Il framework non intende negare questa realtà operativa, quanto piuttosto offrire strumenti per osservarla e discuterla. In questo senso, può aiutare a individuare dove avvengano forme di compressione epistémica e quali effetti possano avere sulla trasparenza e sulla legittimazione dell'evidenza. La piena articolazione dei cinque stadi presuppone inoltre condizioni che non sempre risultano disponibili: tempo, competenze interdisciplinari, spazi deliberativi e committenze disposte a confrontarsi con margini di incertezza. In molte valutazioni tali condizioni sono solo parzialmente presenti. Il framework non propone quindi una soluzione generale, ma uno strumento analitico per rendere più espliciti i trade-off che attraversano il processo valutativo: aiuta a riconoscere che fermarsi alla sola validazione tecnica può riflettere non solo vincoli operativi, ma anche specifiche scelte istituzionali. Allo stesso tempo, il framework non elimina il problema dell'opacità algoritmica né garantisce che il giudizio rimanga pienamente umano. Esso può, contribuire a rendere maggiormente visibili le modalità attraverso cui alcune funzioni interpretative vengono

delegate ai sistemi algoritmici. La sua utilità dipende quindi anche dalla disponibilità istituzionale a utilizzare tale visibilità come occasione di discussione, riflessività e assunzione di responsabilità. Il contributo teorico del framework può essere letto come un tentativo di spostare l'attenzione dal solo controllo tecnico dell'IA alle modalità di costruzione e governo epistemico dell'evidenza, in linea con le più recenti proposte di collocare la valutazione come attore chiave nella governance dei sistemi algoritmici (Pudney *et al.*, 2025; Nielsen, 2025). Una parte della letteratura si interroga soprattutto sull'accuratezza, l'equità o la spiegabilità degli algoritmi. Il framework prova ad aggiungere una domanda complementare: attraverso quali passaggi l'output viene trasformato in evidenza valutativa, e chi governa tale trasformazione. Da questo punto di vista, il framework non si pone in alternativa agli approcci regolativi o tecnici, ma tenta piuttosto di integrarli entro una prospettiva valutativa. In questa prospettiva, la legittimità dell'evidenza non dipende esclusivamente dalle proprietà dell'algoritmo, ma anche dai processi inferenziali e istituzionali attraverso cui l'output viene interpretato e utilizzato.

Conclusioni

L'introduzione dell'IA nei processi valutativi non sembra poter essere interpretata soltanto come un'innovazione tecnica o come una semplice estensione degli strumenti di analisi disponibili. Come emerge dalla sintesi critica della letteratura, il dibattito sull'IA nella valutazione è caratterizzato da una pluralità di approcci e concettualizzazioni che, pur evidenziando potenzialità rilevanti, tendono spesso a concentrarsi sugli strumenti e sugli output prodotti, lasciando meno esplorati i passaggi attraverso cui tali output vengono trasformati in evidenza valutativa. L'analisi proposta ha evidenziato una certa frammentazione epistemica del campo, articolabile, almeno in forma analitica, in tre configurazioni principali: l'IA come estensione strumentale, l'IA come rischio interpretativo ed etico, e l'IA come oggetto di valutazione e governance. Ciascun filone mette in luce aspetti importanti: il primo richiama le potenzialità offerte dalla scala e dalla capacità analitica; il secondo rende visibili questioni legate a bias, opacità e possibili dinamiche di tecnicizzazione; il terzo riporta l'attenzione sulla dimensione normativa e sul ruolo della valutazione nella governance dei sistemi algoritmici. La review suggerisce tuttavia l'esistenza di un elemento ancora relativamente poco tematizzato. I passaggi attraverso cui l'output algoritmico viene trasformato in evidenza valutativa rimangono spesso impliciti. L'IA tende così a essere alternativamente valorizzata, problematizzata o regolata, ma più raramente collocata entro un processo

esplicito che colleghi analisi, interpretazione e giudizio. In assenza di tale esplicitazione, l'output rischia di essere assunto come dato autosufficiente oppure trattato come un oggetto esterno al processo valutativo, mentre il lavoro interpretativo che sostiene la costruzione dell'evidenza rimane in secondo piano. Il framework dei cinque stadi proposto in questo articolo nasce dal tentativo di rendere maggiormente osservabili i passaggi attraverso cui l'informazione algoritmica viene trasformata in evidenza valutativa. Distinguere output, validazione, interpretazione, triangolazione e giudizio significa riportare l'attenzione su dimensioni che l'uso dell'IA può comprimere o rendere meno esplicite. Tale esplicitazione non riguarda soltanto la pratica cognitiva del singolo valutatore, ma investe anche il modo in cui teoria di programma, criteri e triangolazione vengono negoziati con gli stakeholder, ridefinendo in parte il management stesso del processo valutativo.

In questa prospettiva, il contributo dell'articolo non riguarda soltanto l'analisi delle implicazioni tecniche dell'IA, ma la possibilità di interrogare criticamente i processi attraverso cui l'evidenza viene costruita, interpretata e utilizzata nei contesti valutativi. Il framework proposto non intende offrire una nuova prescrizione metodologica, ma una possibile chiave di lettura per analizzare dove e come tali processi vengano compressi, delegati o opacizzati. Ulteriori sviluppi empirici potranno contribuire a verificare come le diverse forme di compressione epistemica si manifestino concretamente in specifici contesti valutativi e organizzativi. Nell'era dell'IA, la questione centrale per la valutazione potrebbe non riguardare soltanto l'adozione di nuovi strumenti, ma la capacità di mantenere visibili, discutibili e governabili i passaggi attraverso cui l'informazione viene trasformata in evidenza e l'evidenza in giudizio.

Dichiarazioni

L'autore dichiara che il Software ChatGPT (GPT-5.6 Sol, OpenAI) è stato utilizzato esclusivamente per limitate attività di revisione linguistica e sintesi. Non sono stati inseriti dati riservati. L'autore si assume la piena responsabilità del manoscritto finale.

Riferimenti bibliografici

Athey, S., Berg, A. (2023). AI and the transformation of social science research: Careful bias management and data fidelity are key. *Science*, 381(6663), 1319–1322. <https://doi.org/10.1126/science.adi1778>

- Azzam, T. (2023). Artificial intelligence and validity. *New Directions for Evaluation*, 178–179, 85–95. <https://doi.org/10.1002/ev.20565>
- Bamberger, M. (2016). *Integrating big data into the monitoring and evaluation of development programmes*. UN Global Pulse.
- Goddard, K., Roudsari, A., Wyatt, J.C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121–127. <https://doi.org/10.1136/amiajnl-2011-000089>
- Grant, M.J., Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Greenhalgh, T., Thorne, S., Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, 48, e12931. <https://doi.org/10.1111/eci.12931>
- Greenstein, N., Cho, S.W. (2025). Ethics & equity in data science for evaluators. In S.B. Nielsen, F. Mazzeo Rinaldi, G. J. Petersson (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation* (pp. 56–77). Routledge. <https://doi.org/10.4324/9781003512493>
- Leeuw, F.L., Bamberger, M. (2025). *Artificial intelligence and big data: Lessons from evaluations of rule of law and development*. Edward Elgar.
- Mason, S., Montrosse-Moorhead, B. (Eds.). (2023). Artificial intelligence and evaluation (Special issue). *New Directions for Evaluation*, 178–179, 1–134. <https://doi.org/10.1002/ev.20554>
- Mazzeo Rinaldi, F. (2018). Big Data e valutazione: Una relazione ancora da costruire. *RIV Rassegna Italiana di Valutazione*, 68, 83–100.
- Mazzeo Rinaldi, F., Nielsen, S.B. (2025). Artificial intelligence: Challenges for evaluators. In S. B. Nielsen, F. Mazzeo Rinaldi, G.J. Petersson (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation* (pp. 287–306). Routledge. <https://doi.org/10.4324/9781003512493>
- Mazzeo Rinaldi, F., Occhipinti, O. (2023). Big Data, intelligenza artificiale e valutazione: Cosa accade in Italia. *RIV Rassegna Italiana di Valutazione*, 85–86, 185–207. <https://doi.org/10.3280/RIV2023-085010>
- Montrosse-Moorhead, B. (2023). Evaluation criteria for artificial intelligence. *New Directions for Evaluation*, 178–179, 123–134. <https://doi.org/10.1002/ev.20566>
- Nielsen, S.B. (2023). Disrupting evaluation? Emerging technologies and their implications for the evaluation industry. *New Directions for Evaluation*, 178–179, 47–57. <https://doi.org/10.1002/ev.20558>
- Nielsen, S.B. (2025). The evaluation industry and emerging technologies. In S. B. Nielsen, F. Mazzeo Rinaldi, & G. J. Petersson (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation* (pp. 266–286). Routledge. <https://doi.org/10.4324/9781003512493>
- Nielsen, S.B., Mazzeo Rinaldi, F., Petersson, G.J. (Eds.). (2025). *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation*. Routledge. <https://doi.org/10.4324/9781003512493>
- NIST - National Institute of Standards and Technology. (2022). AI test, evaluation, validation and verification (TEVV). <https://www.nist.gov/ai-test-evaluation-validation-and-verification-tevv>
- Parasuraman, R., Manzey, D.H. (2010). Complacency and automation bias in the use of imperfect automation. *Human Factors*, 52(3), 381–410. <https://doi.org/10.1177/0018720810376055>
- Paterson, B.L., Thorne, S.E., Canam, C., Jillings, C. (2001). *Meta-study of qualitative health research: A practical guide to meta-analysis and meta-synthesis*. Sage.

- Patton, M.Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34(5 Pt 2), 1189–1208.
- Patton, M.Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage Publications.
- Pawson, R., Tilley, N. (1997). *Realistic evaluation*. Sage Publications.
- Pearl, J., Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic Books.
- Petersson, G.J., Breul, J. (Eds.). (2017). *Cyber society, big data, and evaluation*. Routledge.
- Picciozzo, R. (2020). Evaluation and the Big Data challenge. *American Journal of Evaluation*, 41(2), 166–181. <https://doi.org/10.1177/1098214019850334>
- Potluka, O., Harten, S., Kocks, A., Dvorak, J. (2025). Digitalization in evaluations and evaluations of digitalization: The changing landscape of evaluations. *Evaluation*, 31(3), 289–302. <https://doi.org/10.1177/13563890251357650>
- Pudney, S., Mills, D., Alaci, A.R., Sellers, S., Dvorak, J., Potluka, O. (2025). Evaluation of artificial intelligence-enhanced critical infrastructure systems: A conceptual framework. *Evaluation*, 31(3), 412–443. <https://doi.org/10.1177/13563890251350677>
- Rompczyk, K. (2025). Technological revolution in evaluation: Artificial intelligence and the adherence to evaluation standards. *Evaluation*, 31. <https://doi.org/10.1177/13563890251331066>
- Saunders, M.N.K., Rojon, C. (2011). On the attributes of a critical literature review. *Coaching: An International Journal of Theory, Research and Practice*, 4(2), 158–162. <https://doi.org/10.1080/17521882.2011.596485>
- Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence* (NIST Special Publication 1270). National Institute of Standards and Technology.
- Shapiro, S., Lam, V. (2025). Artificial intelligence in program evaluation: Insights and applications. *Canadian Journal of Program Evaluation*, 39(2), 382–391. <https://doi.org/10.3138/cjpe-2024-0027>
- Sharma, V.D. (2025). Artificial intelligence as a research tool in social sciences and emerging challenges. *International Journal for Multidisciplinary Research*, 7(3). <https://doi.org/10.36948/ijfmr.2025.v07i03.49410>
- Wencker, T., Borst-Graetz, J., Niekler, A. (2025). Text as data for evaluation: Natural language processing and large language models to generate novel insights from unstructured text data. *Evaluation*, 31(3), 369–393. <https://doi.org/10.1177/13563890251330911>
- Whitemore, R., Vilar-Compte, M., De La Cerda, S., Delvy, R., Jaser, S., Stowell, C., Breen, C., Gobeil-Lavoie, A.P., Naranjo, D., Crimston, C., Weinger, K., Pérez-Escamilla, R. (2023). Getting rigor right: A framework for methodological choice in adaptive monitoring and evaluation. *Global Health: Science and Practice*, 11(Suppl. 2). <https://doi.org/10.9745/GHSP-D-22-00243>
- York, P., Bamberger, M. (2020). *Measuring results and impact in the age of big data: The nexus of evaluation, analytics, and digital technology*. Rockefeller Foundation.
- York, P., Bamberger, M. (2025). The applications of Big Data to strengthen evaluation. In S.B. Nielsen, F. Mazzeo Rinaldi, G.J. Petersson (Eds.), *Artificial intelligence and evaluation: Emerging technologies and their implications for evaluation* (pp. 37–55). Routledge.

La facilitazione valutativa nell'approccio partecipativo

*Walter Antonio Canu**

Group facilitation is increasingly used within participatory approaches to social research and program evaluation, particularly in qualitative and mixed-methods designs. However, its methodological role across the different stages of the evaluation process remains under-theorized. Drawing on the literature on utilization-focused evaluation and participatory approaches, this paper distinguishes between ethical and utilitarian-methodological perspectives on participation and proposes an original typology based on the intersection of four evaluation stages and five facilitative purposes, with particular attention, in pragmatic terms, to the functions of meaning-making and sense-making. The proposed framework identifies group facilitation as a cross-cutting methodological resource capable of strengthening both stakeholder participation and the use of evaluation findings.

Keywords: participatory evaluation; group facilitation; utilization-focused evaluation; stakeholder participation; sense-making; pragmatics.

1. La facilitazione valutativa

La facilitazione di gruppi, ossia l'applicazione di tecniche di gruppo in sessioni facilitate, è una pratica ormai comunemente adottata nell'ambito delle politiche pubbliche, sia nel public engagement, sia nella ricerca sociale valutativa di tipo partecipativo. Wilkinson (2012) definisce la facilitazione di gruppi a partire dal concetto di sessione facilitata, ossia «un incontro altamente strutturato in cui il leader della riunione (il facilitatore) guida i

* Walter Antonio CANU, Sapienza Università di Roma (walterantonio.canu@uniroma1.it) e Centrale Valutativa Srl (walter.canu@centralevalutativa.it)
Invio Proposta. 30-04-2026. Accettazione: 17-07-2026.

partecipanti attraverso una serie di passaggi predefiniti per arrivare a un risultato che viene creato, compreso e accettato da tutti i partecipanti». Coinvolgere in modo significativo i gruppi nel processo di ricerca è tipico degli approcci partecipativi propriamente intesi, ossia quelli in cui il valore della partecipazione non è solo simbolico, poiché porta concreti benefici sia dal punto di vista dell'*empowerment* dei portatori di interesse sia dal punto di vista metodologico (Bezzi, 2010a). Sono un esempio di ciò pratiche come la co-programmazione e la co-progettazione, che direttamente ineriscono il lavoro valutativo (Torrighiani, 2025), così come – in termini più generali - associazioni quali la *International Association for Public Participation* (IAP). Bisogna ricordare, tuttavia, che applicare tecniche di gruppo non è mai sufficiente a garantire che gli attori sociali coinvolti in una ricerca valutativa abbiano effettivamente un ruolo significativo nel suo processo (Springett e Wallerstein, 2003), anche perché – mette in guardia Bezzi (2010a) – basta condurre un *workshop* valutativo in modo troppo direttivo, intrusivo o forzatamente guidato verso il consenso, per rendere la partecipazione mera retorica. La vera partecipazione, invece, ha un'utilità concreta in quanto potente stimolo per l'utilizzazione dei risultati e del processo di valutazione (Cousins, 2003). Su questa scia, Patton (2017), l'ideatore della *Utilization-Focused Evaluation*, la valutazione focalizzata sull'utilizzo, ha sistematizzato il ruolo della facilitazione nella valutazione. Nel suo libro, *Facilitating Evaluation*, egli concettualizza (p. 21) la facilitazione valutativa come:

(...) una nicchia specializzata all'interno del più ampio mondo della facilitazione. La facilitazione della valutazione applica e adatta le conoscenze e le tecniche generali di facilitazione alle sfide specifiche del lavoro con i gruppi di stakeholder coinvolti nella valutazione del programma. Lo scopo di tale facilitazione è quello di migliorare la pertinenza, la credibilità, la significatività e l'utilità delle valutazioni.

Per l'autore, la facilitazione valutativa è contraddistinta da proattività, reattività, interattività e adattabilità, e segue cinque principi operativi:

- “Sii guidato/a dal fattore personale”, che suggerisce di acquisire informazioni relativamente a conoscenze, interessi, motivazioni e preoccupazioni di coloro che saranno facilitati, nell'ottica di creare relazioni di fiducia e mutuo rispetto;
- “Coinvolgi tramite le opzioni”, che suggerisce di accompagnare i partecipanti nella generazione e considerazioni di varie opzioni, identificando tutte le alternative, comparandole con i relativi pro e contro, ed

evidenziando le implicazioni di ogni scelta, nell'ottica di aumentare il senso di responsabilità dell'intero processo;

- “Osserva, interpreta, adatta”, che suggerisce di monitorare continuamente le dinamiche di gruppo e lo stato di avanzamento verso i risultati desiderati, trattando tutto ciò che accade come dati di cui tenere conto in modo da rimanere sempre aperti ad eventuali necessità di modifica e riuscire a adattarsi adeguatamente;
- “Incorpora il pensiero valutativo ovunque”, che suggerisce di considerare ogni facilitazione valutativa come un'opportunità di sensibilizzazione dei partecipanti e dei portatori di interesse del programma valutato rispetto al pensiero valutativo come chance per offrire *capacity-building* rispetto alla pratica della valutazione;
- “Rinvigorisci con input all'avanguardia”, che suggerisce di disseminare di spunti il processo di facilitazione valutativa indicando nuove direzioni, idee emergenti, possibilità di evoluzione, opportunità creative, e innovazioni ispiranti, al fine di massimizzare la rilevanza e la potenziale efficacia del processo di valutazione.

In continuità con Patton, e riferendosi in particolare all'aspetto interattivo della sua visione della facilitazione applicata all'ambito valutativo, King e Stevahn (2012) hanno approfondito la questione delle dinamiche interpersonali, mettendo sotto i riflettori il “fattore interpersonale” in aggiunta a quello che Patton chiama “fattore personale”. La loro impostazione si sviluppa a partire dalla teoria dell'interdipendenza sociale (Johnson e Johnson, 2005), che identifica le modalità tramite cui gli individui possono essere in relazione tra loro (cooperativa o competitiva), il tipo di interazione caratteristica di ciascuna relazione (promozionale o oppositiva) e il probabile risultato successivo (reciprocamente o non reciprocamente benefico/dannoso), in termini di soluzioni *win-win*, *win-lose*, *lose-lose*. Per i due autori, che considerano l'interdipendenza sia tra *stakeholder* sia tra *stakeholder* e valutatore, la pratica valutativa interattiva implica che quest'ultimo assuma tre ruoli: *decision-maker*, attore (agente), praticante riflessivo. Infatti, stante la responsabilità deontologica sullo studio valutativo, il ruolo può essere interpretato in vari modi a seconda che si tratti di una valutazione diretta dal valutatore, di una valutazione collaborativa o di una valutazione diretta dagli *stakeholder* partecipanti. La differenza tra i tre tipi deriva da chi ha l'ultima parola sulle decisioni importanti: nel primo caso è il valutatore ad averla (dopo essersi consultato con gli *stakeholder*), nell'ultimo caso sono gli *stakeholder* ad averla (dopo essersi consultati con il valutatore). Vista da questo punto di vista, la facilitazione valutativa di cui parla Patton, non sembra essere affatto

qualcosa di opzionale per chi fa valutazione. Bezzi (2010b) è molto netto su questo punto, quando fa presente che considerare la valutazione come un processo partecipativo sia una necessità metodologica e che non vi siano altre strade percorribili se si punta ad una ricerca che abbia senso e che sia utile.

Eppure, nella letteratura scientifica gli approcci partecipati alla ricerca sociale applicata sono certamente distinti dagli altri. A parte la ricerca-azione proposta da Lewin, chiaramente distinguibile per via del ruolo attivo assunto dai partecipanti, oltre che per lo scopo trasformativo (Barbier, 2008), la ricerca valutativa partecipata è generalmente intesa come un tipo specifico di valutazione, anche perché epistemologicamente legata ad una visione necessariamente sensibile alle suggestioni del costruttivismo sociale (Cousins e Earl, 1992) e della teoria delle rappresentazioni sociali di Moscovici (Howarth, Andreouli, e Kessi, 2014). In questo quadro, è possibile distinguere tra valutazione partecipata nell'accezione "utilitaristico-metodologica" (in cui il coinvolgimento attivo degli attori è un mezzo, più che un fine) e valutazione partecipata nell'accezione "etica", rispetto alla quale il fine del coinvolgimento attivo dei vari portatori di interesse è legato alla volontà di dare voce (e peso) ai molteplici punti di vista e alle relative istanze di coloro che sono toccati dal programma oggetto di valutazione, compresi quelli più svantaggiati o in minoranza, in una prospettiva di giustizia sociale. Nei confronti di queste due accezioni di ricerca valutativa partecipata, la facilitazione è il comune denominatore ma assume una valenza differente a seconda che si tratti dell'una e dell'altra.

2. Valutazione partecipata nell'accezione "etica"

La letteratura mostra come i processi partecipativi siano considerati tali in virtù di due macro-finalità, una definibile come 'deliberativa' ossia finalizzata alla presa di decisioni (Abelson, *et al.*, 2003), l'altra definibile come relativa all'*empowerment*, ovvero connessa all'idea di trasferire potere (decisionale, simbolico, di rappresentanza, etc.) in ottica trasformativa (Yeich e Levine, 1992). A questo proposito, Palumbo e Congiu (2009) ripartiscono i gradi della partecipazione in tre modalità: a) informare e istruire; b) ascoltare e apprendere (consultazione); c) dialogare e deliberare. Sono tuttavia molti i modi per concepire i processi di ricerca partecipativi. Nel suo contributo, Foresti (2003) riporta il modello proposto da Riquier (1997), in cui il grado di partecipazione viene incrociato con il grado di estensione della partecipazione, dando vita a quattro quadranti: il primo, con partecipazione molto intensa (co-produzione) e molto estesa, proprio della valutazione partecipata

(per come Guba e Lincoln l'hanno intesa); il secondo, con partecipazione poco intensa ma molto estesa, in cui i valutatori hanno pressoché totale responsabilità della produzione dei risultati e delle scelte metodologiche; il terzo, con partecipazione molto intensa ma poco estesa, in cui la partecipazione è limitata solo ad alcuni degli *stakeholder* del programma (solitamente lo staff operativo); il quarto, in cui la partecipazione è sia poco intensa che poco estesa. Mclean Petras e Porpora (1993), invece, hanno delineato tre modelli di ricerca partecipativa: il modello di processo parallelo, in cui il ricercatore e la comunità studiata perseguono i loro obiettivi in modo indipendente, seppur collaborando in uno scambio reciprocamente vantaggioso; il modello di coinvolgimento reciproco, che pone il ricercatore in un ruolo attivista, lavorando fianco a fianco con i membri di un gruppo verso obiettivi concreti; il modello dell'Università dell'America Centrale (UCA), che prevede il coinvolgimento diretto dei membri della comunità, in collaborazione con i ricercatori, per la progettazione delle domande di ricerca e per la rilevazione di dati e informazioni che possono poi essere utilizzate dal gruppo. Concettualmente, è la celebre scala della partecipazione di Arnstein (1969) a fornire una delle suggestioni più efficaci per inquadrare il tema rispetto alla valutazione, pur originariamente sviluppata in relazione ai processi di *public engagement*. La scala parte dai livelli non partecipativi, legati alla manipolazione o alla persuasione, fino ad arrivare ai livelli del controllo dei cittadini (o portatori di interesse, nel caso di una ricerca), dove vi è una vera e propria condivisione di responsabilità decisionale (Ciucci, 2010) e, talvolta, una delega da parte degli amministratori (o ricercatori). I livelli intermedi sono perlopiù di partecipazione simbolica, seppur considerabili partecipativi, perché comunque basati su informazione e consultazione, seppur senza un significativo trasferimento di potere decisionale (Arnstein, 1969). Sulla stessa falsariga, Turnbull (1999) ha riconosciuto come l'uso della valutazione partecipativa possa essere anche solo strumentale o simbolico, sostenendo che l'efficacia partecipativa derivi anche dal livello di partecipazione pregresso all'interno del contesto in cui si sviluppa il programma nonché dalla percezione degli individui coinvolti rispetto a quanto un effettivo comportamento partecipativo possa essere, nei fatti, controverso, a dispetto delle intenzioni dichiarate (ad esempio a causa delle dinamiche di potere nascoste).

3. Valutazione partecipata in senso “utilitaristico-metodologico”

Dal lato metodologico, l'argomento della partecipazione è altrettanto complesso. Nonostante si possano concepire vari gradi di “partecipatività”,

delineare dei requisiti minimi tramite cui distinguere la natura partecipativa di un processo richiede comunque una traduzione metodologica, specialmente nel caso della ricerca partecipativa e nella valutazione partecipativa.

Sul piano dei metodi, infatti, vari autori, tra cui Palumbo e Torrigiani (2009) fanno notare una certa congruenza tra 'qualitativo' e 'partecipativo', pur avvertendo a riguardo della necessità di non sovrapporre i piani. In effetti, i due approcci hanno in comune, oltre alla base epistemologica e alla prospettiva dal basso (*bottom-up*), anche l'affinità verso l'idiografia, l'inferenza induttiva, l'orientamento alla comprensione (*Verstehen*) e, non da ultimo, la dimensione testuale del dato (certamente verbale, ma non esclusivamente), che viene co-prodotto. Di questo avviso è anche Lennie (2006), quando sostiene che la valutazione partecipativa è sostenuta da una filosofia interpretativa e da un quadro costruttivista. In quest'ottica, l'uso di tecniche di gruppo nella ricerca valutativa partecipata (e dunque della facilitazione intesa complessivamente come processo) ha ricadute importanti sul lato operativo. Quest'ultime, infatti, oltre ad essere generalmente efficienti (Vindrola-Padros e Johnson, 2020), specialmente perché consentono di raggruppare in una volta più stakeholder), incorporano spesso aspetti quantitativi e qualitativi (categorizzazioni, gerarchizzazioni, votazioni, etc.), risultando per questo motivo particolarmente adatte ai disegni misti, soprattutto se eseguiti con tecniche meno tradizionali del Focus Group (Cyr, 2017), e più innovative, come il *brainstorming* valutativo, la *Nominal Group Technique*, e la *SWOT* relazionale (Bezzi 2013). Quest'ultima, ad esempio, è stata usata nella valutazione del Programma di Sviluppo Rurale 2014-2020 della Regione Sardegna (Buscemi, Licciardo e Solca, 2015) e della Regione Friuli-Venezia Giulia (Cisilino, 2014), per approfondire le interrelazioni tra le caratteristiche del contesto d'intervento e le priorità programmatiche. Vi sono poi le numerose tecniche di gioco di ruolo appartenenti al filone delle Liberating Structures (Lipmanowicz, *et al.*, 2015), che – insieme a quelle provenienti dalla psicologia sociale e dal management, come la Field Force Analysis di Lewin (Swanson e Creed, 2014) o il Fishbone Workshop (Ipsen, *et al.*, 2018) – possono essere affiancate a veri e propri metodi di facilitazione, come il Consensus Workshop (Stanfield, 2002), lo European Awareness Scenario Workshop (Amodio, 1999), il LEGO Serious play® (Beltrami, 2017), il World Café (Celements, Sharples, e Bishop, 2024), l'Open Space Technology (Owen, 2008). Metodi valutativi veri e propri (intesi come combinazione codificata di tecniche), come l'Outcome Mapping (Earl, Carden e Smutylo, 2001), la Most Significant Change technique (Dart e Davies, 2003) e l'Outcome Harvesting (Wilson-Grau, 2018), invece, sono intrinsecamente fondati sull'uso di tecniche miste, e per questo motivo è utile confrontarli

sinteticamente (tab. 1), così da mettere in luce i vari ruoli che le tecniche di gruppo possono avere in un disegno valutativo complessivo.

Tab. 1 - Confronto tra metodi valutativi (OM, MSC e OH) rispetto alla facilitazione dei gruppi

<i>Metodo</i>	<i>Descrizione sintetica del metodo</i>	<i>Ruolo della facilitazione di gruppi</i>
<i>Outcome Mapping (OM)</i>	Ricostruzione partecipata della strategia di un intervento, identificazione dei soggetti strategici (<i>boundary partner</i>) nella sfera di influenza dell'intervento, individuazione e monitoraggio di specifici indicatori di progresso (<i>progress marker</i>) collegati alla strategia e ai singoli soggetti strategici tramite diari di bordo.	Nell'OM l'uso di tecniche di gruppo è previsto esplicitamente nella fase di inquadramento e disegno (per la ricostruzione della strategia, l'identificazione dei soggetti strategici e la definizione degli indicatori progresso). Nelle altre fasi, la facilitazione dei gruppi può avere un ruolo rilevante grazie a workshop dedicati alla riflessione condivisa sull'andamento dei <i>progress marker</i> (fase di costruzione della base empirica) o al commento in itinere dei diari di bordo (fase di analisi ed interpretazione).
<i>Most Significant Change technique (MSC)</i>	Identificazione e selezione partecipata e stratificata di storie indicative di cambiamenti significativi (positivi e negativi) derivanti da un progetto/servizio/prodotto, individuazione delle relative motivazioni ed argomentazioni.	Nella MSC l'uso di tecniche di gruppo è previsto esplicitamente nella fase di inquadramento e disegno (per la definizione delle regole di identificazione e selezione delle storie, tra cui i domini di cambiamento) nonché nella fase di costruzione della base empirica e nella fase di analisi ed interpretazione, rispettivamente per l'identificazione e per la selezione delle storie. Nella fase di divulgazione ed uso, la facilitazione dei gruppi può avere un ruolo rilevante grazie a workshop dedicati alla restituzione condivisa degli esiti della selezione.
<i>Outcome Harvesting (OH)</i>	Identificazione partecipata e verifica indipendente tramite testimoni qualificati ("sostanziazione") degli effetti attesi e inattesi di un intervento, della loro significatività contestuale e del contributo causale degli agenti di cambiamento.	Nell'OH l'uso di tecniche di gruppo è previsto esplicitamente nella fase di costruzione della base empirica (per l'ingaggio degli agenti di cambiamento nell'identificazione e formulazione collaborativa degli effetti da verificare) e nella fase di divulgazione e uso (per l'accompagnamento all'utilizzo dei risultati valutativi). Nelle altre fasi, la facilitazione dei gruppi può avere un ruolo rilevante grazie a workshop dedicati al contributo causale ipotizzato degli agenti di cambiamento (fase di inquadramento e disegno) o alla disamina della significatività contestuale degli esiti verificati (fase di analisi ed interpretazione).

4. Facilitare nel processo valutativo

Per identificare il ruolo della facilitazione nel processo valutativo, si può usare una tipologia derivabile dall'incrocio di due dimensioni concettuali. La

prima riguarda le quattro macro-fasi comunemente attribuite ai processi valutativi:

- Fase di inquadramento e disegno
- Fase di generazione dei dati e delle informazioni (costruzione della base empirica)
- Fase di analisi ed interpretazione dei dati e delle informazioni
- Fase di divulgazione e uso.

La seconda dimensione, invece, attiene agli obiettivi tipici della facilitazione dei gruppi, che sono sintetizzabili essenzialmente nei seguenti cinque punti:

- il coinvolgimento (emotivo-cognitivo) o *engagement* dei partecipanti;
- l'elicitazione di informazioni e l'esplorazione e costruzione dei significati (*meaning-making*);
- la costruzione di senso (*sense-making*) e la formulazione di giudizi (*assessment*);
- la presa di decisioni (*decision making*);
- la creazione di soluzioni.

L'obiettivo del coinvolgimento consiste nella stimolazione di un atteggiamento collaborativo, proattivo, e produttivo nei confronti del gruppo e del processo in generale, ed è considerabile come uno scopo trasversale proprio di qualunque settore, da quello educativo (Eller, 2004) a quello della cittadinanza attiva (Alabdulkarim e Macaulay, 2007). Si tratta di una sorta di pre-condizione, naturalmente imprescindibile per poter aspirare agli altri obiettivi e, non a caso, intuitivamente concepibile come un aspetto definitorio di base della nozione di facilitazione di gruppi. In modo analogo, la presa di decisioni è comunemente annoverata tra gli obiettivi principali – se non il principale – dei processi facilitativi (Kaner, 2014). Questo obiettivo riguarda la creazione condivisa di qualcosa, come la soluzione ad un problema (Vick, 2015), che sia un piano di azione, un modello di intervento, una rappresentazione grafica o addirittura una scultura simbolica. Per le peculiarità della facilitazione valutativa, ho preferito separare l'obiettivo della creazione di soluzioni (Priest e Gass, 1997) da quello della presa di decisioni. Rimangono da chiarire due obiettivi: uno relativo all'elicitazione delle informazioni e all'esplorazione e comprensione dei significati, l'altro relativo alla costruzione di senso condiviso e alla formulazione di giudizi. A differenza degli altri tre, si tratta di obiettivi compositi, ossia formati da più di una funzione.

La scelta di raggrupparli in due gruppi è dovuta ai seguenti motivi (oltre alla parsimonia):

- nella pratica, queste funzioni trovano spesso concretizzazione negli stessi momenti (nella stessa sessione facilitata o addirittura nel corso della stessa attività, come illustrerò successivamente);
- esse rappresentano – a mio avviso – il cuore della facilitazione valutativa, in quanto ruotano attorno a due concetti fondamentali per la ricerca sociale, significato e senso, la cui complessità e rilevanza impone una trattazione a parte nel prossimo paragrafo.

Infine, con ‘elicitazione delle informazioni’ si intende l’azione dello stimolare direttamente o indirettamente la condivisione, da parte dei partecipanti, di informazioni di qualunque natura, relative ad esperienze, emozioni, impressioni, opinioni, conoscenze, spiegazioni, etc. (Barton, 2015) anche facenti parte della conoscenza tacita (Marradi, 2003). La ‘formulazione di giudizi’, invece, consiste nel far esprimere ai partecipanti delle valutazioni assolute o relative (comparative) rispetto ad uno o più referenti, seguendo dei criteri esplicitati; tale valutazione può essere o meno espressione di un consenso condiviso e può avere una forma quantitativa o qualitativa.

5. Una necessaria digressione: significati e senso nella facilitazione valutativa

Per intendersi sui concetti di *meaning making* e di *sense making*, centrali nella facilitazione valutativa, è necessario affrontare i concetti di significato e di senso. Come distinguerli?

Secondo Bezzi (2019) – che considera la costruzione di senso condiviso la funzione principale della ricerca sociale – la soluzione è linguistica, e deriva dalla differenza tra semantica e pragmatica (Bezzi, 2022). Solitamente, si considera la prima come inerente allo studio dei significati intrinseci del linguaggio, e la seconda come inerente allo studio dei significati estrinseci derivanti dall’uso del linguaggio in contesti specifici. Tuttavia, usare questo *fundamentum divisionis* come criterio per distinguere significato e senso è discutibile, perché anche alcuni significati connotativi sono tali in quanto propri di un determinato contesto d’uso e di una situazione specifica, in virtù dell’indessicalità evidenziata da Putnam (1975).

L’alternativa, allora, consiste nel considerare il concetto di senso come ‘pragmatico’ non dal punto di vista linguistico, bensì filosofico. In questa prospettiva il senso di un concetto è legato alla sua utilità, alla sua capacità

di produrre effetti, sulla scia di Peirce, James, Dewey e i vari epigoni del pragmatismo (Campelli, 2018). Così facendo, al *meaning-making*, ossia alla costruzione di significati, si lascia il compito di esplorare come i partecipanti concepiscono i referenti, identificando gli aspetti dell'intensione sovrapposti (e non) nei vari schemi mentali più o meno cristallizzati (Neisser, 1989), così da giungere eventualmente – magari proprio attraverso tecniche di gruppo facilitate o tramite l'analisi e la restituzione di interviste in profondità (Krauss, 2005) – a dei significati condivisi. Ad esempio, relativamente al concetto di 'presa in carico integrata', si potrebbe arrivare a definire cosa voglia dire, secondo quegli operatori sociali lì, in quell'ambito spazio-temporale lì, prendere in carico gli utenti in modo integrato. Questo è ciò di cui parla Bezzi (2003) riferendosi all'esplorazione del campo semantico nella ricerca valutativa, ovvero chiedersi "cosa intendiamo quando parliamo di...?".

Alla (co-)costruzione di senso (*sense-making*), invece, si affida un compito diverso, che va oltre la semantica, nella pragmatica. Bezzi (2006) concettualizza la valutazione pragmatica a partire dalla centralità del contesto e degli scopi degli attori, mettendo in evidenza come entrambi gli elementi vadano studiati ermeneuticamente, favorendo altresì la partecipazione dei portatori di interesse. Ciò è particolarmente rilevante per la presente riflessione, anche perché la facilitazione di gruppi racchiude in sé un importante elemento ermeneutico, come messo in luce da Walsh e Andersen (2013). Tuttavia, intendere la costruzione di senso da una prospettiva filosoficamente pragmatica potrebbe voler dire qualcosa di ancora diverso, oltre la *frame analysis* di Goffman (1974), lo studio dei significati locali di Geertz (1983) e l'interpretazione fenomenologica dell'intersoggettività di Schütz (1971).

Per Weick (1995), che ha affrontato il tema proprio in riferimento ai contesti organizzativi, il *sense-making* è definibile come un processo attraverso il quale le persone, attingendo alle loro identità e ai contesti sociali, interpretano retrospettivamente, in modo attivo, le esperienze fatte e le situazioni incontrate, al fine di giungere ad una comprensione plausibile che permetta di agire. Questa concezione della costruzione di senso, indubitabilmente pragmatica (visto l'accento sull'azione), è probabilmente la più condivisa. In quest'ottica, nell'ambito della facilitazione valutativa, al *sense-making* è affidato il compito di valorizzare l'interazione sociale di gruppo con l'obiettivo di trasformare i concetti "semantici" (frutto del processo di significazione condivisa) in concetti "pragmatici", ossia agibili e agiti poiché utilizzabili e utili, seguendo l'approccio valutativo *Utilization-Focused* di Patton (2011).

In altre parole: non più solo "la presa in carico integrata, in questo programma, vuol dire X", bensì "la presa in carico integrata, in questo

programma, è stata fatta / si può fare / si farà in questo modo X... con questi effetti (supposti o verificati) Y, T, Z.”. Il focus è dunque sull’azione e sulle conseguenze pratiche di tale azione, e ciò vale anche per concetti che non esprimono di per sé azioni, bensì atteggiamenti (ad esempio, il pregiudizio etnico degli assistenti sociali nei confronti di beneficiari di etnia rom in un programma di integrazione sociale), capacità (ad esempio, la capacità di gestione dello stress degli operatori sociosanitari in un servizio pubblico-privato di prevenzione del gioco d’azzardo), conoscenze (ad esempio, la conoscenza degli agricoltori relativamente alle varie forme di impollinazione in un progetto di sensibilizzazione sull’agricoltura biologica). In tutti questi casi, il contesto, composto anche dagli stessi attori sociali, rende possibile passare dal significato al senso (in particolare, nel caso della valutazione, al senso condiviso, più che individuale), valorizzando l’uso potenziale ed effettivo dei significati e le relative conseguenze. La facilitazione di gruppi può essere un modo particolarmente efficace per operare questa trasformazione.

6. Scopi facilitativi e fasi della valutazione: una tipologia

La tipologia proposta dà luogo a venti tipi rappresentativi di altrettante modalità di applicazione della facilitazione valutativa. Si noti che, mentre le fasi valutative sono considerabili piuttosto lineari (nonostante vi sia comunque una certa circolarità del percorso, nel caso di disegni qualitativi e misti), gli scopi facilitativi sono da ritenersi privi di una sequenzialità precisa, fatta (parzialmente) eccezione per la costruzione dei significati e la costruzione del senso condiviso.

Nella tabella seguente - sviluppata sulla base dell’esperienza professionale dell’autore - per ognuna delle modalità di applicazione facilitativa (indicata dall’incrocio tra riga e colonna) è stato inserito un giudizio sul livello di utilità che la facilitazione può avere rispetto alla fase valutativa in questione. La domanda che guida la lettura dello schema è: “Stante la possibilità teorica di perseguire gli scopi facilitativi in ogni fase valutativa, quand’è che la facilitazione di gruppi è relativamente più utile?”

La colonna relativa al coinvolgimento è naturalmente quella con un livello di utilità sempre alto, per le ragioni già precedentemente esposte relative alla partecipazione degli *stakeholder*, sia essa perseguita prevalentemente per fini metodologici o etici. Rispetto all’elicitazione di informazioni e all’esplorazione e costruzione dei significati, invece, il livello di utilità è alto nella fase di inquadramento e disegno così come in quella di costruzione della base empirica, e più basso nelle altre fasi. Infatti, le tecniche di gruppo

finalizzate a rilevare informazioni e a processarle semanticamente possono essere usate, oltre che per l'esplorazione del campo semantico dell'evaluando (Bezzi, 2003), anche per l'analisi situazionale (Patton, 2011), per l'analisi delle condizioni di valutabilità (Wholey, 1987), per la (ri)co-costruzione della teoria del programma, nonché per la costruzione condivisa di strumenti d'indagine, come nel caso di un brainstorming valutativo (Bezzi, 2011) con operatori sociali finalizzato ad enucleare i significati attribuiti alla riabilitazione degli alcolisti ed usarli poi induttivamente per costruire possibili indicatori da operativizzare all'interno di un questionario. Anche rispetto alla rilevazione di dati e informazioni, le tecniche di gruppo sono particolarmente efficaci in termini di *meaning-making* e forse questa modalità di applicazione è quella più comune, come dimostra ad esempio il grande successo del Focus Group, nonostante i suoi limiti (Bezzi, 2005).

Tab. 2 - Tipologia delle modalità di applicazione della facilitazione valutativa con il livello di utilità della facilitazione di gruppi per ogni tipo

Scopo facilitativo Fase valutativa	Coinvolgimento emotivo-cognitivo dei partecipanti	Elicitazione di informazioni e meaning-making	Sense-making e formulazione di giudizi (assessment)	Presa di decisioni (decision making)	Creazione di soluzioni
Inquadramento e disegno	alto	alto	medio	alto	alto
Costruzione della base empirica	alto	alto	basso	medio	basso
Analisi ed interpretazione di dati e informazioni	alto	medio	alto	medio	basso
Divulgazione ed uso	alto	basso	alto	alto	alto

Diversamente, per quanto concerne l'analisi e l'interpretazione di dati e informazioni, le tecniche di gruppo risultano meno utili se orientate alla costruzione di significati (condivisi o meno), mentre risultano estremamente utili per la costruzione di senso e per la formulazione di giudizi. Per comprendere le ragioni di questo differente livello di utilità, conviene innanzitutto non considerare l'elicitazione delle informazioni e la formulazione di giudizi, ossia le funzioni di cui si compongono i due scopi confrontati; la prima non va considerata per via del fatto che non è tipicamente considerata una funzione centrale rispetto all'analisi e all'interpretazione della base empirica (anche se non è raro elicitare informazioni anche in questa fase, ad

esempio tramite un'intervista di *follow-up*, oppure nei disegni di ricerca circolari, in cui rilevazione e analisi procedono parallelamente); la seconda non va considerata per via del fatto che l'*assessment*, pur essendo spesso abbinato alla costruzione di senso intesa pragmaticamente (quindi in riferimento agli usi pratici dei concetti e ai relativi effetti) grazie alla sua natura sia analitica (e dunque propedeutica alla costruzione di senso) che sintetica (e dunque consequenziale alla costruzione di senso), non è né condizione necessaria né effetto inevitabile del *sense-making* (diversamente dall'elicitazione delle informazioni, che è invece essenziale per il *meaning-making*).

Infatti, le ragioni per cui, in fase di analisi ed interpretazione, la facilitazione tramite tecniche di gruppo è meno utile per la costruzione di significati rispetto alla costruzione di senso condiviso riguardano proprio le già trattate differenze tra significato e senso. Del resto, se rispetto al primo (il significato), il ricercatore, una volta chiarito cosa si intende con il concetto X, può anche analizzare ed interpretare autonomamente i risultati del processo di significazione collettiva e fornire ai partecipanti una restituzione potenzialmente autoconclusiva (almeno relativamente a quella fase di ricerca), rispetto al secondo (il senso condiviso) non è possibile immaginare che il ricercatore eviti di coinvolgere direttamente i partecipanti nell'analisi e nell'interpretazione, perché la domanda a cui risponde non è “Cosa intendiamo con il concetto X?”, bensì “Cosa ci facciamo con il concetto X?”.

L'ultimo quesito, infatti, implica necessariamente un confronto partecipato, proprio perché lo scopo è creare un senso condiviso (non individuale), e le tecniche di gruppo sono, come visto, il mezzo più adeguato a giungervi. Certamente, visto che la costruzione di senso è qualcosa che riguarda anche l'uso e gli effetti pregressi del concetto (quindi anche quelli del passato e non solo quelli del presente o del futuro), la domanda potrebbe anche essere “Cosa ci abbiamo fatto con il concetto X?” e, teoricamente, in quel caso, si potrebbe anche prescindere dal ricorso a tecniche di gruppo in sede di analisi ed interpretazione, per costruire un senso condiviso. Il ricercatore, in effetti, potrebbe condurre un'analisi fenomenologica delle interviste (Eatough e Smith, 2017) per identificare gli aspetti primari e secondari dell'intensione del fenomeno (concetto) studiato per come descritto dalle esperienze soggettive dei partecipanti, senza ricorrere ad un confronto di gruppo e semmai sottoponendo agli intervistati le sue risultanze solo dopo e individualmente, massimizzando l'intersoggettività tramite il *member checking* (Candela, 2019). Tuttavia, in questo caso, la prospettiva pragmatica verrebbe valorizzata solo retrospettivamente, rispetto all'uso dei concetti e ai loro effetti, e non prospettivamente, cioè in relazione all'usabilità e utilità futura di tali concetti. Per questo motivo, anche considerato che la ricerca applicata e,

nello specifico, quella valutativa, differisce dalla ricerca di base proprio per l'orientamento verso l'uso potenziale dei risultati, è ragionevole considerare le tecniche di gruppo particolarmente utili per il *sense-making* in fase di analisi ed interpretazione.

Si può dare una lettura del tutto simile anche per quanto concerne la differenza nell'utilità delle tecniche di gruppo rispetto al *meaning-making* e al *sense-making* nella fase di divulgazione ed uso dei risultati della valutazione. In questa fase, la facilitazione assume una notevole utilità nel secondo caso (contrariamente al primo) in considerazione del ruolo assumibile dalle tecniche di gruppo nei confronti di questioni quali l'*assessment* di opzioni alternative suggerite dalle raccomandazioni valutative o la socializzazione dei risultati della valutazione con un pubblico allargato di *stakeholder*, finalizzata ad una costruzione più allargata del senso condiviso rispetto al programma.

A riguardo della presa di decisioni e della creazione di soluzioni, le differenze in termini di utilità della facilitazione dei gruppi sono meno marcate. Per entrambi questi due scopi, le tecniche di gruppo possono considerarsi molto utili, sia relativamente alla fase di inquadramento e disegno che alla fase di divulgazione ed uso. Rispetto a quest'ultima, le decisioni da prendere (ad esempio, relativamente a come dar seguito a raccomandazioni valutative specifiche o, più in generale, all'eventuale prosecuzione o interruzione del programma) possono essere facilitate tramite tecniche di gruppo come, ad esempio, la *Qualitative Weight and Sum*, suggerita da Scriven e Davidson (2005), che consente di eliminare progressivamente le alternative tramite una procedura iterativa di confronto sistematico dei criteri. Nella fase di inquadramento e disegno, invece, la facilitazione è indispensabile qualora, ad esempio, si vogliano costruire criteri di giudizio condivisi e relativi livelli di *performance* tramite una rubrica valutativa avente descrittori qualitativi e/o parametri quantitativi in grado di discriminare tra livello molto carente, carente, accettabile, buono e ottimo. In fase di analisi ed interpretazione, la stessa rubrica può essere oggetto di un ulteriore workshop facilitato al fine di rivedere / ricalibrare i criteri usati.

Nei confronti della creazione di soluzioni, invece, le tecniche di gruppo non mantengono la stessa utilità nelle fasi di costruzione della base empirica e di analisi ed interpretazione rispetto alle altre due. Infatti, nell'inquadramento e nel disegno di una *Developmental Evaluation*, finalizzata a facilitare lo sviluppo di un programma innovativo in un contesto emergente (Patton, 1994), il coinvolgimento del gruppo può rendere possibile la creazione condivisa della teoria del cambiamento, del piano di erogazione della prestazione (Rossi, Lipsey, e Henry, 2018), o di una batteria di affermazioni da inserire in un questionario, così come nella fase di divulgazione ed uso esso

può facilitare l'ideazione di una strategia di *outreach* per comunicare gli apprendimenti principali della valutazione o l'elaborazione di modifiche al programma sulla base dei risultati della ricerca valutativa. Per contro, nella costruzione della base empirica e nell'analisi ed interpretazione della stessa, le tecniche di gruppo hanno un ruolo più limitato in termini di creazione di soluzioni, poiché queste due fasi della valutazione non sono normalmente deputate alla creazione intesa come produzione di un *output* intellettuale considerabile come soluzione condivisa ad un problema (a meno che di non includere l'eventuale creazione condivisa dei trascritti delle interviste e il relativo piano di codifica).

Conclusioni

La facilitazione dei gruppi è una pratica ormai comunemente adottata anche nella ricerca sociale valutativa, soprattutto nell'ambito di valutazioni orientate da approcci partecipativi e realizzate tramite metodi qualitativi e misti. Ciononostante, la letteratura che tratta l'argomento raramente entra nel merito degli aspetti strettamente metodologici implicati dall'uso della facilitazione e delle relative tecniche di gruppo nell'architettura complessiva del disegno valutativo, soffermandosi, invece, o sulle singole tecniche, o sui benefici che l'approccio partecipativo porta in termini etici, ossia di coinvolgimento ed *empowerment* degli stakeholder (sorvolando spesso sul come, concretamente, la facilitazione dei gruppi possa rendere partecipativi tali approcci). Per colmare questa lacuna, è stata sviluppata una tipologia attorno al concetto di 'facilitazione valutativa', derivante dall'incrocio tra le fasi della ricerca valutativa e le funzioni della facilitazione dei gruppi, tramite cui è possibile concepire la facilitazione come un'opzione praticabile trasversalmente, non solo nel quadro degli approcci partecipativi¹, poiché essa è particolarmente adatta all'integrazione metodologica. Inoltre, con riferimento agli approcci partecipativi, le tecniche di gruppo, se concepite coerentemente all'interno di un metodo valutativo a tutto tondo, possono giocare un ruolo essenziale per rendere la partecipazione effettiva (andando oltre la partecipazione simbolica), perché consentono di valorizzare i contributi individuali e collettivi lungo tutto il processo di valutazione, in un'ottica che privilegia la costruzione di senso condiviso (a livello pragmatico), piuttosto che la

¹ Si pensi, ad esempio, al ruolo della facilitazione dei gruppi nel coinvolgere esperti per identificare le variabili più importanti su cui basare il matching statistico nel caso di un disegno quasi-sperimentale.

costruzione di significati (a livello semantico). In prospettiva, i prossimi approfondimenti dovrebbero riguardare essenzialmente come fare in modo che la facilitazione valutativa alimenti reciprocamente tanto i benefici metodologici quanto quelli etici.

Dichiarazioni

Si dichiara l'assenza di conflitti di interesse e di finanziamenti, totali o parziali, della ricerca da cui deriva l'articolo. L'autore dichiara, inoltre, di non aver utilizzato strumenti di IA nella redazione del contributo.

Riferimenti bibliografici

- Abelson, J., Forest, P.G., Eyles, J., Smith, P., Martin, E., Gauvin, F.P. (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social science e medicine*, 57(2), 239-251.
- Alabdulkarim, A.A., Macaulay, L.A. (2007). Facilitation patterns and citizen engagement. *International Journal of Technology, Policy and Management*, 7(2), 122-133.
- Amodio, L. (1999). *Atelier del futuro: la metodologia European awareness scenario workshop per promuovere la partecipazione nei processi di innovazione e sviluppo sostenibile*. CUEN.
- Arnstein, S.R. (1969). A ladder of citizen participation. *Journal of the American Institute of planners*, 35(4), 216-224.
- Barbier, R. (2008). *La ricerca-azione*. Armando Editore.
- Barton, K.C. (2015). Elicitation techniques: Getting people to talk about ideas they don't usually talk about. *Theory e Research in Social Education*, 43(2), 179-205.
- Beltrami, G. (2017). *Leg® Serious Play® pensare con le mani: Valore per le persone, valore per le organizzazioni*. FrancoAngeli.
- Bezzi, C. (2003). Cosa valutiamo, esattamente?. *Rassegna Italiana di Valutazione*, 27, 1000-1025.
- Bezzi, C. (2005). Il focus group: consideriamone i limiti. *Sociologia e ricerca sociale*, 76/77, 1000-1008.
- Bezzi, C. (2006). Evaluation pragmatics. *Evaluation*, 12(1), 56-76.
- Bezzi, C. (2010a). La partecipazione valutativa dalla retorica al metodo. *Rassegna Italiana di Valutazione*, 47, 119-130.
- Bezzi, C. (2010b). *Il nuovo disegno della ricerca valutativa*. FrancoAngeli.
- Bezzi, C. (2011). Evaluational brainstorming. *Sociology Mind*, 1(04), 151.
- Bezzi, C. (2013). Fare ricerca con i gruppi. Guida all'utilizzo di focus group, brainstorming, Delphi e altre tecniche. FrancoAngeli.
- Bezzi, C. (2019). La ricerca sociale e valutativa come costruzione di senso condiviso. *Quaderni di Sociologia Clinica*, (18/2019).
- Bezzi, C. (2022). *La costruzione del dato nelle scienze sociali*. FrancoAngeli.
- Buscemi, V., Licciardo, F., e Solca, F. (2015, 17-18 aprile). *La valutazione del contesto di riferimento del PSR della Regione Sardegna 2014-2020 attraverso la SWOT relazionale*. XVIII Congresso Annuale dell'Associazione Nazionale di Valutazione, Genova.

- Campelli, E. (2018). *Da un luogo comune: Introduzione alla metodologia delle scienze sociali* (4a ed.). Carrocci.
- Candela, A.G. (2019). Exploring the function of member checking. *The qualitative report*, 24(3), 619-628.
- Cisilino, F. (a cura di). (2014). *Il valore aggiunto del partenariato nella costruzione del Programma di sviluppo rurale 2014–2020: Friuli-Venezia Giulia*. INEA.
- Ciucci, F. (2010). Valutazione e Pubblica Amministrazione: il contributo della partecipazione alla decisione. *Rivista trimestrale di scienza dell'amministrazione*, 1, 2010.
- Clements, A. J., Sharples, A., Bishop, J. (2024). The World Café method for engaging groups in conversation: Practical considerations and an agenda for critical evaluation. *Occupational Psychology Outlook*, 3(1), 6-18.
- Cousins, J.B. (2003). Utilization effects of participatory evaluation. In *International handbook of educational evaluation* (pp. 245-265). Dordrecht: Springer Netherlands.
- Cousins, J.B., Earl, L.M. (1992). The case for participatory evaluation. *Educational evaluation and policy analysis*, 14(4), 397-418.
- Cyr, J. (2017). The unique utility of focus groups for mixed-methods research. *PS: Political Science e Politics*, 50(4), 1038-1042.
- Dart, J., Davies, R. (2003). A dialogical, story-based evaluation tool: The most significant change technique. *The American Journal of Evaluation*, 24(2), 137-155.
- Davidson, E.J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Sage Publications.
- Eatough, V., Smith, J.A. (2017). Interpretative phenomenological analysis. In Willig, C., e Rogers, W.S. (Eds). *The SAGE handbook of qualitative research in psychology*. Sage Publications.
- Earl, S., Carden, F., e Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs*. IDRC.
- Eller, J. (2004). *Effective group facilitation in education: How to energize meetings and manage difficult groups*. Corwin Press.
- Foresti, M. (2003). La partecipazione in ambito valutativo: mito o realtà? Teoria e pratica degli approcci partecipati in valutazione. *Rassegna Italiana di Valutazione*, 25, 1000-1026.
- Geertz, C. (1983) Local Knowledge: Further Essays in Interpretative Anthropology. *Basic Books*.
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.
- Howarth, C., Andreouli, E., Kessi, S. (2014). Social representations and the politics of participation. In *The Palgrave handbook of global political psychology*. Palgrave Macmillan UK.
- Ipsen, C., Sørensen, O.H., Poulsen, S., Gish, L. (2018). Using High-Involvement Fishbone Workshops to Transform Problem Identification into Tailor-Made Organizational Interventions. In *Organizational Interventions for Health and Well-being* (pp. 25-41). Routledge.
- Johnson, D.W., Johnson, R.T. (2005). New developments in social interdependence theory. *Genetic, social, and general psychology monographs*, 131(4), 285-358.
- Kaner, S. (2014). *Facilitator's guide to participatory decision-making*. John Wiley e Sons.
- King, J. A., Stevahn, L. (2012). *Interactive evaluation practice: Mastering the interpersonal dynamics of program evaluation*. Sage Publications.
- Krauss, S.E. (2005). Research paradigms and meaning making: A primer. *The qualitative report*, 10(4), 758-770.
- Lennie, J. (2006). Increasing the rigour and trustworthiness of participatory evaluations: learnings from the field. *Evaluation Journal of Australasia*, 6(1), 27-35.
- Lipmanowicz, H., Singhal, A., McCandless, K., Wang, H. (2015). Liberating structures: Engaging everyone to build a good life together. *Communication and "the good life"*, 233-246.

- Marradi, A. (2003). Il ruolo della conoscenza tacita nella vita quotidiana e nella scienza. In Lazzari, F., e Merler, A. (Eds.) *La sociologia della solidarietà*. FrancoAngeli.
- Neisser, U. (1989). *Concepts and conceptual development: Ecological and intellectual factors in categorization*. CUP Archive.
- Owen, H. (2008). *Open space technology: A user's guide*. Berrett-Koehler Publishers
- Palumbo, M., Congiu, D. (2009). Partecipazione e valutazione per il Piano Regolatore Sociale. In *La partecipazione tra ricerca e valutazione* (pp. 37-72). FrancoAngeli.
- Palumbo, M., Torrigiani, C. (2009). *La partecipazione fra ricerca e valutazione*. FrancoAngeli.
- Patton, M. Q. (1994). Developmental evaluation. *Evaluation practice*, 15(3), 311-319.
- Patton, M. Q. (2011). *Essentials of utilization-focused evaluation*. Sage Publications.
- Patton, M. Q. (2017). *Facilitating Evaluation: Principles in Practice*. Sage Publications.
- Petras, E. M., e Porpora, D. V. (1993). Participatory research: Three models and an analysis. *The American Sociologist*, 24(1), 107-126.
- Priest, S., Gass, M. (1997). An Examination of "Problem-Solving" versus "Solution-Focused" Facilitation Styles in a Corporate Setting. *Journal of Experiential Education*, 20(1), 34-39.
- Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131-193.
- Riquier, G. (1997). *Stakeholders' participation in public policy evaluations: Impact on organizational learning*. European Evaluation Society, Stockholm, Sweden.
- Schütz, A. (1971) *Collected Papers*. Martinus Nijhoff.
- Springett, J., Wallerstein, N. (2003). Issues in participatory evaluation. *Community-based participatory research for health*, 2, 263-288.
- Stanfield, R.B. (2002). *The workshop book: From individual creativity to group action*. New Society Publishers.
- Swanson, D. J., Creed, A.S. (2014). Sharpening the focus of force field analysis. *Journal of change management*, 14(1), 28-47.
- Torrigiani, C. (2025). Co-programmazione, co-progettazione e valutazione delle politiche sociali. *RIV: rassegna italiana di valutazione*: 91, 1, 2025, 30-47.
- Turnbull, B. (1999). The mediating effect of participation efficacy on evaluation use. *Evaluation and program Planning*, 22(2), 131-140.
- Vick, W. (2015). *Process-Based Facilitation. Facilitation for meeting leaders, consultants and group facilitators*. iUniverse.
- Vindrola-Padros, C., Johnson, G. A. (2020). Rapid techniques in qualitative research: a critical review of the literature. *Qualitative health research*, 30(10), 1596-1604.
- Walsh, K.D., Andersen, P. (2013). Group facilitation as hermeneutic practice. *Group Facilitation*, 12, 15.
- Weick, K.E., Weick, K.E. (1995). *Sensemaking in organizations*. Sage Publications.
- Wholey, J.S. (1987). Evaluability Assessment: Developing Program Theory. *New directions for program evaluation*, 33, 77-92.
- Wilkinson, M. (2012). *The secrets of facilitation: The SMART guide to getting results with groups*. John Wiley e Sons.
- Wilson-Grau, R. (2018). *Outcome harvesting: Principles, steps, and evaluation applications*. IAP.
- Yeich, S., Levine, R. (1992). Participatory Research's Contribution to a Conceptualization of Empowerment. *Journal of Applied Social Psychology*, 22(24), 1894-1908.

Contesti di attuazione, mediazioni operative e traiettorie di cambiamento. Un'analisi valutativa del progetto Ri-Belli

*Elvira Celardi**

This article examines, through the case of the Ri-Belli project, an intervention aimed at addressing educational poverty, the implementation conditions and operational mediations that shape trajectories of change. Adopting a possibilist perspective and a theory-based reading of implementation, it interprets change as a situated outcome produced through the interaction between project resources, implementation contexts, and family arrangements. The analysis reconstructs recurrent territorial and family configurations and shows, through the case of an affective education initiative, how the same intervention can generate different trajectories: convergence, misalignment, or transformative adaptation. The article argues that evaluating these interventions requires shifting attention away from mere outcome verification and towards understanding the conditions that make change possible, hinder its consolidation, or redirect its course.

Keywords: educational poverty; implementation contexts; operational mediations; trajectories of change; evaluative possibilism; theory-based evaluation.

Introduzione

Il contributo muove dallo studio valutativo del progetto Ri-Belli, un intervento di contrasto alla povertà educativa minorile realizzato in alcune aree del territorio ragusano segnate da forte vulnerabilità sociale: il centro storico di Ragusa e la cosiddetta fascia trasformata, zona rurale caratterizzata da precarietà abitativa e lavorativa, fragilità delle reti di supporto e rilevante

* Elvira CELARDI, Università di Catania (elvira.celardi@unict.it)

Invio Proposta: 24-03-2026. Accettazione: 17-07-2026.

presenza migratoria (Petino, Wilson e Torre, 2022; Celardi, 2025). In questi contesti, la povertà educativa si intreccia con forme di deprivazione materiale, marginalità sociale, traiettorie familiari e migratorie complesse, traducendosi in percorsi educativi discontinui e in relazioni di supporto ridotte e instabili. La ricerca valutativa analizza le condizioni entro cui il cambiamento ha preso (o meno) forma grazie all'attivazione di Ri-Belli. Il progetto non è stato assunto come un oggetto chiuso da verificare *ex post*, bensì come un processo da osservare nella sua interazione con contesti differenti, vincoli situati e margini di mediazione.

L'analisi del cambiamento è stata orientata da una postura possibilista (Stame, 2022), che considera l'implementazione come un processo esplorativo e adattivo, in cui deviazioni, resistenze ed esiti inattesi diventano elementi conoscitivi decisivi (Hirschman, 1967; Tandler, 1992). In questa chiave, la ricerca si colloca nel filone della *Theory-Based Evaluation* (Weiss, 1997) e utilizza la *Realist Evaluation* (Pawson e Tilley, 1997) come lente interpretativa. Entro questa cornice, l'articolo ricostruisce alcune configurazioni ricorrenti emerse nel percorso valutativo e si sofferma, come caso esemplificativo, sul laboratorio di affettività per adolescenti dell'azione "*Ri-Belli per... Volersi bene*". Il caso mostra come il cambiamento non dipenda soltanto dalla coerenza del disegno progettuale, ma anche dalle mediazioni attivabili nei diversi contesti di attuazione, mettendo in luce traiettorie non uniformi e nuove domande valutative.

1. Il progetto Ri-Belli e il suo percorso valutativo

Coerentemente con l'Obiettivo 4 dell'Agenda 2030, Ri-Belli mirava a ridurre il divario che colpisce bambini e adolescenti in situazione di povertà educativa e materiale, ampliandone le opportunità educative e intervenendo in modo integrato sulle condizioni di vita che ne ostacolano i percorsi. Il progetto si ispirava all'idea della bellezza come leva di rigenerazione sociale e simbolica e a quella della ribellione come tensione etica al cambiamento (Camus, 1951), un nesso che già trovava sintesi nel nome stesso di "Ri-Belli". L'intervento era articolato lungo tre assi strategici: il rafforzamento delle competenze emotive e relazionali e il miglioramento delle condizioni di vita nell'ambiente domestico e familiare; lo sviluppo della comunità educante e il potenziamento dei servizi territoriali; il sostegno all'inclusione educativa, al successo scolastico e alla continuità formativa. Le azioni progettuali si traducevano, a loro volta, in una pluralità di attività concrete –

laboratori, consulenze, accompagnamento personalizzato, e servizi di prossimità – distribuite tra famiglia, territorio e scuola.

Il finanziatore del progetto e committente della valutazione, Impresa sociale Con i Bambini, ha predisposto, per gli interventi finanziati nell'ambito del medesimo bando, un set articolato di indicatori, rispetto al quale Ri-Belli è stato chiamato a documentare in modo trasparente la propria efficacia. Tale impianto consente di osservare in modo comparabile sia i risultati di realizzazione – platee raggiunte e intensità del coinvolgimento – sia i cambiamenti attesi tra i destinatari diretti e nel territorio. Alla luce di questi indicatori, Ri-Belli restituisce un quadro complessivamente positivo, con diversi scostamenti favorevoli rispetto ai target iniziali. Sul piano della realizzazione, i destinatari coinvolti sono stati 659 minori, 216 genitori e 162 tra insegnanti ed educatori: rispettivamente 409, 76 e 62 unità in più rispetto a quanto atteso. Anche rispetto ai cambiamenti osservati, i risultati appaiono in larga misura coerenti con le previsioni e, in alcuni casi, superiori: 318 minori, ossia 118 in più rispetto ai 200 previsti, hanno mostrato un rafforzamento delle competenze non cognitive, mentre il target di 75 genitori con un miglioramento delle competenze educative è stato pienamente raggiunto. Sul versante scolastico, 37 minori, a fronte dei 31 attesi, hanno registrato un incremento significativo della frequenza. Più contenuti, pur restando prossimi ai valori previsti, risultano invece gli esiti relativi alla ripresa degli studi, con 13 casi su 15 attesi, e alla conclusione positiva dell'anno scolastico, con 11 casi su 15. Anche per quanto riguarda il potenziamento della comunità educante, i risultati appaiono nel complesso coerenti con le attese: sono stati attivati 14 protocolli con enti esterni su 15 previsti, sono stati implementati tutti e 12 i servizi integrativi leggeri attesi e sono state avviate 7 reti dopo l'inizio del progetto, a fronte delle 3 inizialmente previste.

Questi dati restituiscono un quadro incoraggiante, ma non bastano a ricostruire i processi attraverso cui tali risultati sono stati prodotti, né a comprendere le dinamiche che ne hanno favorito o ostacolato il dispiegamento. Una lettura fondata esclusivamente sui target tende infatti ad appiattire la variabilità interna del progetto, come se le azioni si svolgessero in contesti neutri e producessero effetti sostanzialmente omogenei. Ciò che resta in ombra è proprio il modo in cui la logica d'intervento prende forma nell'operatività quotidiana, attraverso il lavoro di coordinamento tra partner, le mediazioni costruite dagli operatori, le risposte differenziate dei destinatari e i vincoli posti dai contesti di attuazione.

Da questo scarto tra la performance osservata e l'intelligibilità dei processi prende avvio il percorso di approfondimento sviluppato nelle pagine che seguono. Accanto alla domanda su che cosa sia stato realizzato e con

quali esiti misurabili, la ricerca si è interrogata, in una prospettiva di possibilismo valutativo (Stame, 2019; 2022), su come, per chi e a quali condizioni tali esiti si siano generati, quali possibilità di cambiamento abbiano effettivamente aperto e perché, in altri casi, non si siano consolidate. I risultati osservati non sono stati intesi come semplice conferma o smentita della teoria iniziale del progetto, bensì come base per riformulare le domande valutative alla luce delle differenze emerse nei contesti di implementazione e delle possibilità di cambiamento effettivamente aperte.

1.1. L'implementazione come spazio di produzione del cambiamento

In contesti sociali e istituzionali ad alta variabilità, come quelli in cui si è inserito Ri-Belli, la teoria del programma diventa intelligibile solo se osservata nel suo farsi concreto, cioè nel modo in cui viene interpretata, negoziata e tradotta nelle pratiche dagli attori coinvolti (Weiss, 1997; Koleros et al., 2020). Per questa ragione, la ricerca valutativa ha assunto l'implementazione come spazio privilegiato di osservazione, considerando il cambiamento non come esito automatico del disegno progettuale, ma come processo situato che prende forma nell'interazione tra risorse del progetto, risposte dei destinatari, vincoli contestuali e mediazioni operative. In questa prospettiva, il contesto è stato trattato come un insieme di condizioni attive che modulano l'attivazione dei processi di cambiamento. È qui che il possibilismo assume il suo significato più operativo: la valutazione non si limita a verificare se gli esiti attesi si siano prodotti, ma tratta scarti, resistenze e aggiustamenti come materiale conoscitivo decisivo per comprendere quali possibilità di cambiamento il progetto abbia effettivamente aperto nei diversi contesti e a quali condizioni esse abbiano potuto consolidarsi, indebolirsi o non attivarsi affatto. L'inferenza causale è stata pertanto trattata in termini di contribuzione più che di attribuzione. L'obiettivo non è stato isolare un effetto netto imputabile al progetto, ma ricostruire se e come l'intervento abbia contribuito a produrre determinati esiti attraverso meccanismi generativi operanti in condizioni contestuali specifiche. In questa direzione, la ricerca valutativa ha assunto come riferimento principale la *Realist Evaluation*, utile a leggere l'interazione tra contesti, meccanismi ed esiti (Pawson e Tilley, 1997; Pawson, 2006). Operativamente, lo schema CME è stato applicato dapprima alle singole azioni progettuali, per ricostruire le relazioni tra contesti, meccanismi ed esiti emerse nel corso dell'implementazione, individuando le condizioni rilevanti, i meccanismi che si attivavano, si bloccavano o richiedevano mediazioni e gli esiti osservati nei diversi gruppi di destinatari. In un secondo

momento, le evidenze raccolte sono state ricomposte in una lettura integrata, consentendo di ricostruire configurazioni interpretative ricorrenti dei processi di implementazione. La ricerca si è avvalsa, inoltre, della *Street-Level Evaluation* per mettere a fuoco lo sforzo quotidiano degli operatori, le strategie adattive elaborate in risposta agli imprevisti e le mediazioni attraverso cui l'intervento veniva tradotto nella realtà (Brodtkin, 2008; Lipsky, 2010). In chiave più generale, il percorso ha mantenuto anche un'attenzione di tipo *developmental* (Patton, 2011), trattando l'implementazione come un processo in evoluzione e gli scarti come un'occasione per riformulare le domande valutative alla luce di quanto emergeva dal campo.

La base empirica della ricerca, prevalentemente qualitativa, ha combinato due livelli tra loro complementari. Da un lato, i materiali prodotti dal monitoraggio del progetto – fogli firma, registri delle attività, diari di bordo, questionari di autovalutazione, osservazioni partecipative, relazioni tecniche e report raccolti nei diversi momenti del monitoraggio – hanno fornito la base descrittiva dell'attuazione, consentendo di seguire l'andamento formale delle azioni, i livelli di partecipazione e gli esiti rilevati nel corso del progetto. Dall'altro, la ricerca valutativa in profondità ha utilizzato focus group, interviste e osservazione per interrogare i significati di quei dati, ricostruire le difficoltà incontrate, i successi inattesi, le strategie di adattamento e le condizioni che avevano favorito o ostacolato il cambiamento.

Il percorso si è sviluppato in modo iterativo nel corso delle diverse annualità del progetto. Dopo una prima ricognizione dei contesti e della documentazione disponibile, al termine del primo anno di attuazione sono stati condotti tre focus group con referenti e operatori, dedicati alle differenze tra i contesti, alle criticità incontrate, alle strategie di superamento e ai successi inattesi. A distanza di alcuni mesi, un secondo ingresso sul campo ha previsto interviste in profondità con la referente dell'associazione capofila e con alcune operatrici, finalizzate ad approfondire temi emersi nei focus group. Tale ricostruzione è stata integrata da colloqui con testimoni privilegiati del territorio, utili a far emergere le configurazioni locali del cambiamento e i fattori di innesco o di resistenza. L'analisi è stata condotta attraverso la triangolazione delle fonti e la progressiva stabilizzazione delle configurazioni interpretative.

2. Contesti di attuazione e condizioni del cambiamento

Seguendo il progetto nel suo percorso di attuazione, la ricerca ha ricostruito tre configurazioni territoriali – *integrazione limitata*, *integrazione*

frammentata e isolamento estremo – che restituiscono articolazioni interne rilevanti delle aree di intervento. A differenziarle non è tanto il profilo territoriale in senso stretto, quanto la diversa combinazione di disponibilità e qualità degli spazi, accessibilità ai servizi, condizioni abitative, presenza di luoghi di aggregazione, grado di infrastrutturazione sociale e tenuta delle reti territoriali. È da tali combinazioni che, di volta in volta, dipende la possibilità del progetto di intercettare bisogni, costruire relazioni e sostenere percorsi di cambiamento.

2.1. Il progetto nel contesto di integrazione limitata

L'*integrazione limitata* coincide prevalentemente con il centro storico di Ragusa, un'area urbana segnata dallo spopolamento e dall'isolamento sociale, ma inserita in un contesto cittadino dotato di una rete di servizi relativamente strutturata. In questa configurazione, il problema principale non risiede tanto nell'assenza di risorse quanto nella loro accessibilità selettiva: i servizi esistono, ma restano poco praticabili per i nuclei più fragili. La scuola rappresenta uno dei pochi presidi stabili del territorio, ma il raccordo tra famiglie, servizi e opportunità educative non è automatico e richiede mediazioni capaci di ridurre le distanze materiali, organizzative e simboliche (Cecardi, Mazzeo Rinaldi e Occhipinti, 2026).

Qui, il progetto Ri-Belli ha intercettato soprattutto famiglie monoparentali, composte da madri sole con figli a carico, esposte alla precarietà economica, alle difficoltà di conciliazione tra lavoro e cura e alla debolezza delle reti di supporto.

Sebbene la dispersione scolastica risulti relativamente contenuta, la continuità educativa dei minori risente dell'intreccio di vincoli materiali, relazionali e organizzativi che caratterizza la loro vita quotidiana. In questo contesto, il progetto ha svolto una funzione di intermediazione, favorendo l'emersione di bisogni difficilmente esprimibili e accompagnando i nuclei familiari nell'accesso alle risorse territoriali disponibili.

L'azione *Ri-Belli per... Abitare*, orientata al miglioramento delle condizioni abitative e al sostegno della quotidianità familiare attraverso interventi domiciliari, ha reso particolarmente evidente questo meccanismo. La presenza continuativa dell'équipe nelle abitazioni ha favorito la costruzione di una relazione di fiducia e l'emersione progressiva di bisogni, fragilità e risorse che difficilmente sarebbero stati rilevabili in contesti di intervento diversi da quello domestico. Su questa base è stato possibile orientare e

accompagnare le famiglie verso i servizi più pertinenti e promuoverne una partecipazione più attiva alle opportunità educative offerte dal progetto.

I cambiamenti osservati si sono collocati su due piani complementari. Da un lato, le abitazioni hanno progressivamente cessato di apparire come spazi anonimi, spogli o provvisori, assumendo i tratti di luoghi più curati, vissuti e relazionalmente significativi. Dall'altro, si è registrato un rafforzamento delle dimensioni educativa e relazionale, soprattutto nel rapporto tra genitori e figli e nell'attenzione al percorso scolastico dei minori. Tra gli effetti osservati, l'inserimento dei figli nelle attività pomeridiane e nei laboratori educativi ha alleggerito il carico di cura delle madri e, in alcuni casi, ha reso più praticabili percorsi formativi o lavorativi.

Il laboratorio di affettività per adulti, dell'azione *Ri-Belli per... Volersi bene* si è collocato nella medesima logica. L'accessibilità del servizio, la flessibilità degli orari e l'approccio paritario delle operatrici hanno favorito l'emersione di bisogni inespressi e la costruzione di spazi di ascolto e di confronto condivisi.

Nel contesto di integrazione limitata, dunque, il cambiamento non deriva dalla creazione di nuove opportunità, ma dalla capacità del progetto di rendere accessibili risorse e servizi già presenti sul territorio attraverso relazioni di prossimità e percorsi di accompagnamento rivolti ai nuclei più fragili.

2.2. Il progetto nel contesto di integrazione frammentata

L'integrazione frammentata caratterizza alcune aree della provincia di Ragusa, tra cui Santa Croce Camerina, Acate, Vittoria e Scoglitti. Qui la fragilità sociale si manifesta nella limitata disponibilità di servizi di supporto alle famiglie, nella debolezza dei presidi educativi, ricreativi e aggregativi e nelle limitate opportunità di partecipazione e di socializzazione dei minori.

A Santa Croce Camerina, dove questa configurazione si manifesta in modo particolarmente evidente, molti minori coinvolti nel progetto appartenevano a famiglie segnate dalla precarietà economica e da una presenza adulta intermittente, condizionata dai tempi e dai vincoli del lavoro agricolo nelle serre. In questo quadro, la principale criticità riguardava la discontinuità dei percorsi educativi e relazionali al di fuori della scuola. La debolezza dei presidi educativi e delle reti comunitarie favoriva, soprattutto tra gli adolescenti, dinamiche di gruppo conflittuali, episodi di bullismo e comportamenti oppositivi.

In risposta a queste criticità, grazie alla collaborazione con la Parrocchia San Giovanni Battista e in sinergia con l'azione *Ri-Belli per... Imparare*, il

progetto ha coinvolto operatori di comunità, volontari e mediatori culturali, attivando uno spazio educativo e di incontro rivolto a minori e madri. Per i ventisei ragazzi partecipanti, questo spazio è divenuto un presidio educativo di prossimità in cui studiare, ricevere supporto scolastico e partecipare ad attività ludiche e cooperative. Anche i momenti conviviali, come le merende condivise e le feste di quartiere, hanno assunto una funzione educativa, contribuendo a rafforzare le relazioni tra pari e a promuovere modalità più costruttive di gestione dei conflitti.

Il lavoro con le quattordici madri coinvolte ha svolto una funzione complementare nel rafforzare i legami educativi. I momenti di incontro e i percorsi linguistici avviati in corso d'opera hanno offerto occasioni di confronto e contribuito a ridurre le difficoltà nella comprensione delle comunicazioni scolastiche e amministrative. In questo modo, il progetto ha rafforzato il rapporto tra famiglie e istituzioni educative, rendendo più accessibili informazioni, servizi e opportunità che, in assenza di mediazione, sarebbero stati difficilmente fruibili.

2.3. *Il progetto nel contesto di isolamento estremo*

L'isolamento estremo emerge nelle aree più marginali della fascia trasformata, in particolare nella zona di Marina di Acate, dove si concentra una consistente presenza di famiglie migranti, spesso residenti in abitazioni precarie e inadeguate, prive delle condizioni abitative minime. Qui il problema non riguarda soltanto le opportunità educative, ma le stesse condizioni di accesso ai servizi, alla mobilità e alla vita sociale.

L'area è segnata da una forte deprivazione infrastrutturale: i servizi sono assenti o molto distanti, il trasporto pubblico manca del tutto e le abitazioni sono disperse sul territorio, rendendo difficoltosi gli spostamenti. L'isolamento geografico, sociale e istituzionale rende inoltre la popolazione residente poco visibile, difficilmente raggiungibile e debolmente connessa ai circuiti ordinari di tutela, di istruzione e di partecipazione.

L'assenza di servizi per la prima infanzia e la normalizzazione del lavoro minorile limitano profondamente i percorsi educativi dei minori (Save the Children, 2023). In molti casi, la cura dei bambini più piccoli ricade sui fratelli e sulle sorelle maggiori; in altri, i minori accompagnano i genitori al lavoro o sono coinvolti direttamente nelle attività agricole, con ripercussioni che arrivano fino alla non frequenza e, nei casi più gravi, all'abbandono scolastico. Per molti bambini e adolescenti che vivono a Marina di Acate, il mondo conosciuto coincide quasi interamente con l'ambiente familiare, il

lavoro nelle serre e i codici relazionali della comunità di appartenenza. Le possibilità di futuro tendono così a presentarsi come rigidamente predeterminate: per i ragazzi, il lavoro agricolo; per molte ragazze, ruoli familiari e di cura fortemente tradizionali. L'isolamento, dunque, limita anche la possibilità stessa di immaginare percorsi di vita differenti.

In questo contesto, il progetto ha agito soprattutto aprendo spazi di esperienza che il territorio tendeva a comprimere. In particolare, l'azione *Ri-Belli per... Stupire*, pensata per stimolare nei minori la capacità di meravigliarsi e di riconoscere la bellezza nei propri contesti di vita quotidiana, ha trovato nel laboratorio fotografico uno strumento privilegiato.

L'efficacia del laboratorio non risiede tanto nell'acquisizione di una competenza tecnica, quanto nell'opportunità offerta ai partecipanti di uscire dai contesti abituali, esplorare il territorio e sperimentare luoghi e relazioni fino ad allora estranei alla loro quotidianità. Le uscite, le escursioni, il contatto con operatori ed educatori e l'uso della fotografia come strumento di osservazione hanno ampliato l'orizzonte di esperienza dei minori, favorendo una rilettura più consapevole del contesto di appartenenza e aprendo alla possibilità di immaginare percorsi di vita diversi da quelli considerati fino ad allora scontati.

L'intervento ha inoltre permesso a molti bambini e adolescenti di sperimentare capacità e ruoli fino a quel momento impensati, rafforzando l'autostima e sostenendo la costruzione di aspettative scolastiche e professionali diverse da quelle considerate abituali nel proprio ambiente di vita.

Più fragili e discontinui sono risultati, invece, i cambiamenti nelle figure adulte di riferimento. Gli uomini sono rimasti in larga misura ai margini del progetto, pur continuando spesso a orientare in modo decisivo le scelte familiari. Anche per molte donne l'attivazione è stata selettiva: alcune azioni pensate per le madri non sono riuscite ad avviarsi o non hanno incontrato domande effettive. Su questo hanno inciso la limitata autonomia di movimento, il sovraccarico di cura e di lavoro, la dipendenza dalle decisioni maschili, la difficoltà di raggiungere i servizi e, più in generale, la distanza culturale rispetto a interventi percepiti come non pienamente compatibili con i confini della sfera personale e familiare. Per questo motivo, le aperture generate nei minori hanno richiesto un costante lavoro di mediazione con le famiglie, per evitare che nuove aspirazioni o possibilità di scelta entrassero in conflitto con aspettative familiari fortemente strutturate.

La configurazione di isolamento estremo mostra così, al tempo stesso, le maggiori potenzialità e il limite più netto dell'azione progettuale. Proprio nei contesti in cui il progetto ha ampliato gli orizzonti di esperienza dei minori, è emersa con maggiore evidenza la difficoltà di consolidare tali cambiamenti

nella vita familiare, anche perché le condizioni di isolamento e deprivazione che caratterizzano il territorio continuano a ostacolare il consolidamento dei cambiamenti avviati dal progetto. In assenza di trasporti, presidi educativi stabili e servizi accessibili, infatti, anche le trasformazioni più promettenti rischiano di restare fragili.

2.4. Assetti familiari e praticabilità del cambiamento

Come si è visto, l'accesso alle attività, la continuità della partecipazione e, più in generale, la praticabilità del cambiamento non sono riconducibili soltanto a caratteristiche socio-territoriali, ma anche agli assetti domestici di provenienza dei minori, che differiscono per le risorse disponibili, gli stili educativi e gli orientamenti culturali.

A partire dal lavoro sul campo, sono state ricostruite tre configurazioni ricorrenti – *famiglie ponte*, *sopraffatte* e *ancorate* – utili a comprendere in che modo il contesto familiare abbia ridefinito il tipo di mediazione necessario all'attivazione del progetto. Le configurazioni familiari proposte hanno valore euristico e sono state ricostruite a partire dall'incrocio di tre dimensioni analitiche: il *capitale familiare*, inteso come insieme delle risorse culturali, sociali ed economiche disponibili nel nucleo (Bourdieu, 1986); gli *stili educativi*, riferiti a orientamenti più autorevoli o autoritari, ma letti alla luce di contesti segnati da forti vincoli di sopravvivenza quotidiana (Baumrind, 1967); l'*orientamento interculturale*, inteso come grado di apertura, oscillazione o chiusura nei confronti di norme, modelli educativi e riferimenti culturali diversi da quelli prevalenti nel contesto familiare (Banks, 2004; Berry, 1997).

Tali configurazioni si collegano a tre modi ricorrenti di rapporto tra bisogni, risorse progettuali e possibilità di cambiamento: la *convergenza*, quando il progetto incontra bisogni già riconosciuti; la *mediazione*, quando deve costruire le condizioni dell'incontro; la *mancata connessione*, quando non riesce ad attivare un aggancio significativo (Celardi, 2025, pp. 230-243).

Le *famiglie ponte* sono nuclei che, pur disponendo di risorse economiche limitate e di reti di sostegno fragili, riconoscono il valore dell'istruzione e mostrano una certa attenzione al percorso educativo dei figli. Il vincolo principale non è l'assenza di investimento educativo, ma la difficoltà a tradurlo in un effettivo accesso alle opportunità formative, extrascolastiche o di socializzazione. Qui prevale una dinamica di *convergenza*: il progetto intercetta bisogni già almeno in parte riconosciuti e rende concretamente accessibili opportunità che il contesto familiare o territoriale non consentirebbe di

sostenere autonomamente. Ciò accade, per esempio, con l'assegnazione di *doti educative*, utilizzate per sostenere le spese scolastiche, acquistare materiali didattici e rendere accessibile la partecipazione dei minori ad attività educative e ricreative. In questi casi, il progetto trasforma una disponibilità educativa esistente in una concreta possibilità di accesso.

Le *famiglie sopraffatte* sono nuclei in cui la vita quotidiana è assorbita dalla pressione dei bisogni urgenti: casa, lavoro, lingua, pratiche amministrative, gestione dei figli, problemi di mobilità o di documentazione. In questi casi, il problema non è l'assenza di interesse per la dimensione educativa, ma la difficoltà a investire energie quando il presente è dominato dall'emergenza. Il cambiamento non si produce quindi per semplice adesione a un'offerta già definita, ma richiede un lavoro di prossimità capace di agganciare i genitori ai problemi più immediati. Solo quando questa relazione di fiducia si consolida, si apre uno spazio ulteriore in cui anche la relazione genitoriale e i bisogni educativi dei figli possono essere nominati e affrontati. Qui la traiettoria prevalente è quella della *mediazione*.

Le famiglie *ancorate*, infine, sono nuclei in cui le fragilità materiali si intrecciano con cornici culturali e normative più rigide, spesso segnate da modelli educativi autoritari, ruoli di genere gerarchizzati e orientamenti interculturali più chiusi. In questi casi non è in gioco soltanto la scarsità di risorse, ma anche la presenza di confini normativi che limitano la legittimità di alcune scelte educative e delle aspirazioni dei minori. Il lavoro di mediazione è perciò più lento e faticoso, perché non basta ridurre gli ostacoli pratici: occorre costruire fiducia, lavorare sui codici culturali e negoziare gradualmente la possibilità stessa del cambiamento. Quando questa mediazione si attiva, il progetto riesce ad aprire varchi; quando non si attiva, le possibilità rese visibili dall'intervento tendono a restare senza consolidamento, ad esempio nei casi in cui la partecipazione dei minori viene interrotta o non autorizzata dalle figure adulte di riferimento. È qui che la traiettoria può slittare dalla mediazione alla *mancata connessione*.

3. Un caso esemplificativo delle dinamiche di implementazione: il laboratorio di affettività

L'esperienza del laboratorio di affettività, rivolto a ragazze e ragazzi tra gli 11 e i 18 anni e promossa nell'ambito dell'azione *Ri-Belli per... Volersi bene* mostra come una stessa proposta educativa possa generare traiettorie differenti in relazione ai contesti di attuazione, alle soglie di accettabilità dei destinatari e alle mediazioni introdotte in corso d'opera. Il laboratorio è stato

progettato secondo un approccio educativo attivo e partecipativo, ricorrendo a *role playing*, *brainstorming*, lavori di gruppo e materiali visivi, con l'obiettivo di stimolare il coinvolgimento e favorire l'elaborazione di argomenti poco trattati in famiglia. I temi sono stati sviluppati attraverso il metodo *Teen STAR*,¹ che propone un percorso di conoscenza progressiva della corporeità e delle sue implicazioni relazionali, affettive e valoriali, in una prospettiva olistica dello sviluppo personale.

La ricerca sul campo ha tuttavia evidenziato differenze significative tra i livelli di partecipazione e i tipi di coinvolgimento riscontrati nei diversi contesti di attuazione. Presso il Consultorio Familiare Diocesano di Vittoria, i partecipanti hanno mostrato continuità nella frequenza, disponibilità al confronto e un progressivo coinvolgimento attivo. In questo contesto, la regolarità della partecipazione e il clima relazionale cooperativo hanno reso via via possibile affrontare non solo i cambiamenti fisiologici, le emozioni e i sentimenti, ma anche temi più sensibili, come l'omosessualità, l'omofobia, l'innamoramento e l'amore.

Nel laboratorio svolto presso il Presidio della Caritas di Marina di Acate, invece, la stessa proposta educativa ha incontrato maggiori difficoltà, soprattutto quando l'attenzione si è concentrata su temi più delicati, come la salute riproduttiva e la sessualità. In alcuni casi tali difficoltà si sono tradotte in rifiuti espliciti e in comportamenti di boicottaggio, rendendo visibili soglie di accettabilità più restrittive e la necessità di un progressivo riadattamento dell'intervento.

3.1. Tra resistenza e coinvolgimento: il caso di Marina di Acate

A Marina di Acate, fin dall'avvio, il laboratorio ha incontrato sfide complesse che ne hanno condizionato il funzionamento. Nei gruppi di età compresa tra gli 11 e i 14 anni, le difficoltà hanno riguardato soprattutto il mantenimento dell'attenzione e l'elaborazione dei contenuti proposti. Tali difficoltà, probabilmente riconducibili a una limitata esposizione dei minori a stimoli culturali e relazionali, sono state presto contenute dall'équipe attraverso

¹ Il metodo Teen STAR (*Sexuality Teaching in the Context of Adult Responsibility*), sviluppato da Hanna Klaus, propone un approccio educativo fondato sulla crescita integrale della persona e sull'esplorazione delle dimensioni fisica, emotiva,itoriale, intellettuale e sociale dello sviluppo umano. L'obiettivo è fornire agli adolescenti strumenti per comprendere i cambiamenti del corpo e delle emozioni, promuovendo una visione responsabile della sessualità (Klaus, 1998).

l'introduzione di strategie più dinamiche e coinvolgenti, come l'animazione di strada, che hanno reso gli incontri più interattivi e stimolanti.

Le criticità più marcate si sono invece riscontrate nel gruppo dei più grandi (15-17 anni), in particolare durante le sessioni dedicate alla conoscenza degli organi riproduttivi, che hanno incontrato una forte opposizione e, in alcuni casi, un vero e proprio boicottaggio da parte dei partecipanti maschi. Alcune operatrici, tuttavia, hanno riferito che, dopo episodi di aperta ribellione, alcuni ragazzi abbandonavano la stanza, salvo poi osservare da dietro la porta, spinti da una curiosità latente per contenuti taciuti nel contesto domestico. Secondo l'équipe, questo comportamento non riflette un rifiuto lineare dei temi proposti, bensì un conflitto tra ciò che era socialmente consentito nel gruppo di appartenenza e ciò che attraeva sul piano individuale. Nel racconto delle operatrici, tale dinamica appare chiaramente differenziata per genere. Le ragazze, infatti, sono rimaste in aula, mostrando fin da subito una maggiore apertura e curiosità per le tematiche proposte, anche per il desiderio di acquisire informazioni difficilmente reperibili nel contesto familiare. Questa predisposizione iniziale, tuttavia, non si è mantenuta invariata. La presenza di fratelli e cugini maschi nei gruppi ha progressivamente generato un clima di pressione sociale, limitando la libertà di espressione delle ragazze e rendendo più difficile la condivisione aperta delle esperienze. Per sbloccare la situazione ed evitare la chiusura del laboratorio è stato necessario l'intervento di una mediatrice culturale, che ha favorito una migliore comprensione delle dinamiche culturali e sociali in gioco e, di conseguenza, un adattamento più consapevole delle strategie educative. L'équipe ha così ricalibrato l'approccio e modificato le tematiche in programma, orientandosi verso un'impostazione più attenta al contesto di provenienza dei minori e alla costruzione di un dialogo non impositivo. Negli incontri successivi si è registrata una trasformazione graduale nel coinvolgimento dei partecipanti: le ragazze hanno recuperato fiducia e partecipazione attiva, mentre, verso la fine del percorso, si è osservata anche una maggiore apertura da parte di alcuni ragazzi, con modalità di relazione più autentiche sia con le operatrici sia con i coetanei.

3.2. *Dalla divergenza empirica alle configurazioni di implementazione*

Dall'analisi dell'attuazione del laboratorio nei due contesti è stato possibile ricostruire tre configurazioni di implementazione, lette attraverso lo schema CME della valutazione realista: una situazione di *convergenza* tra la proposta educativa e le domande dei destinatari; una situazione di *parziale*

disallineamento; e una situazione di *adattamento trasformativo*, in cui l'intervento è stato progressivamente rimodellato in base alle esigenze emergenti. Tali configurazioni mostrano come le caratteristiche locali e la strategia progettuale abbiano interagito, dando luogo a traiettorie di cambiamento diverse. A Vittoria (Tab. 1), la mancanza di luoghi di incontro ha alimentato il bisogno di relazione e di appartenenza e il limitato dialogo familiare su temi affettivi e sessuali ha generato un vuoto informativo che il laboratorio ha intercettato, favorendo la partecipazione attiva e il progressivo consolidamento del senso di appartenenza al gruppo.

Tab.1 - Configurazione CME – Convergenza tra richieste e intervento

<i>Contesto</i>		<i>Meccanismo</i>	<i>Esito</i>
<i>Dove</i>	<i>Chi</i>	<i>Perché + Come</i>	<i>Che cosa accade</i>
<i>Territorio</i> : urbano, con scarsità di spazi aggregativi per adolescenti.	Minori (11–17 anni), italiani e stranieri, provenienti da famiglie con retrotterra culturale medio-alto.	Socialità compensativa : il laboratorio intercetta un bisogno di relazione e appartenenza.	Risultati di attuazione : alta partecipazione, forte coinvolgimento emotivo, consolidamento del senso di appartenenza al gruppo.
<i>Ambiente domestico</i> : dialogo limitato su temi affettivi e relazionali.		Curiosità emancipativa : il limitato dialogo familiare stimola interesse verso contenuti nuovi.	Cambiamenti osservati : aumento della consapevolezza emotiva e relazionale; rafforzamento dell'autoefficacia e della capacità di espressione personale.
<i>Ambiente di implementazione</i> : Consultorio Familiare Diocesano, riconosciuto come luogo neutro e sicuro.	Operatrici	L'approccio empatico delle operatrici favorisce fiducia e partecipazione.	

A Marina di Acate, le operatrici hanno inizialmente adottato la medesima strategia educativa e relazionale utilizzata con successo a Vittoria, fondata su un'impostazione predefinita elaborata in fase di progettazione. Anche in questo caso, l'assenza di spazi aggregativi ha alimentato il bisogno di socialità tra i minori, incentivando la partecipazione al laboratorio. Nella fase iniziale dell'attuazione (t₁), tuttavia, la proposta educativa si è scontrata con resistenze culturali più marcate, determinando un parziale *disallineamento* tra l'intervento e le richieste dei partecipanti. Mentre per molte ragazze i contenuti del laboratorio hanno rappresentato un'opportunità inedita di conoscenza e riflessione su aspetti affettivi e corporei raramente discussi in ambito familiare, per i ragazzi la curiosità iniziale si è trasformata in un conflitto identitario. Il laboratorio è stato percepito come uno spazio potenzialmente minaccioso per il ruolo sociale ricoperto nel proprio contesto di riferimento.

Tab.2 - Configurazione CME – Parziale disallineamento tra richieste e intervento

<i>Contesto</i>	<i>Meccanismo</i>		<i>Esito</i>
<i>Dove</i>	<i>Chi</i>	<i>Perché + Come</i>	<i>Che cosa accade</i>
<i>Territorio:</i> rurale, con presenza di comunità straniere chiuse	Minori (11–17 anni) in isolamento sociale	Socialità controllata. Il bisogno di relazione, alimentato dalla scarsità di spazi di incontro, si confronta con un ambiente fortemente regolato da vincoli familiari e sociali.	Partecipazione formalmente elevata, ma coinvolgimento disomogeneo e condizionato dalle dinamiche di controllo interno al gruppo.
<i>Ambiente domestico:</i> migranti con basso capitale culturale, rigide norme tradizionali e marcate asimmetrie di genere	Operatrici del progetto	La presenza di parenti e coetanei appartenenti allo stesso nucleo o gruppo culturale limita la libera espressione individuale. Le operatrici, ancora ancorate a una strategia predefinita, faticano inizialmente a ricalibrare il setting educativo rispetto alle dinamiche emergenti.	
	Ragazze (15–17 anni) con scarsa alfabetizzazione affettiva	Curiosità repressa. Le ragazze mostrano inizialmente interesse e apertura verso i contenuti del laboratorio, percepito come uno spazio protetto per accedere a conoscenze altrimenti difficilmente disponibili. Tuttavia, la pressione sociale esercitata dalla presenza di fratelli e cugini maschi inibisce progressivamente la libertà di espressione.	Coinvolgimento inizialmente attivo, seguito da ritiro progressivo
<i>Ambiente di implementazione:</i> presidio della Caritas di Marina di Acate, frequentato da minori inseriti in reti parentali e familiari ristrette.	Operatrici del progetto		
Strategia educativa pre-strutturata	Ragazzi (15–17 anni) con scarsa alfabetizzazione affettiva	Resistenza negoziata. La curiosità iniziale verso i contenuti trattati si trasforma in conflitto identitario: il laboratorio viene percepito come uno spazio potenzialmente minaccioso per il ruolo sociale ricoperto nel gruppo e nel contesto familiare. Il timore di perdere la propria posizione di forza innesca forme di opposizione e sabotaggio.	Reazioni di rifiuto, comportamenti oppositivi e boicottaggio dell'attività.
	Operatrici del progetto		

Il timore di perdere la propria posizione di forza all'interno del gruppo ha innescato forme di opposizione e di sabotaggio, che hanno ostacolato il processo di apprendimento e di partecipazione. Anche per le ragazze, tuttavia, la pressione sociale esercitata dal gruppo ha progressivamente ridotto gli spazi di espressione, indebolendo la partecipazione attiva (Tab. 2).

Tab.3 - Configurazione CME – Adattamento trasformativo

<i>Contesto</i>		<i>Meccanismo</i>	<i>Esito</i>
<i>Dove</i>	<i>Chi</i>	<i>Perché + Come</i>	<i>Che cosa accade</i>
<i>Territorio:</i> Rurale, con presenza di comunità straniere chiuse	Mediatrice culturale	Curiosità facilitata L'intervento della mediatrice consente di comprendere meglio le reazioni dei minori e di rimodulare attività, linguaggi e contenuti in modo più rispettoso delle sensibilità culturali del contesto	Miglioramento del clima relazionale
<i>Ambiente domestico:</i> migranti con basso capitale culturale, rigide norme tradizionali e marcate asimmetrie di genere	Operatrici del progetto		Riduzione dei comportamenti oppositivi; minore vergogna; maggiore disponibilità alla narrazione personale e a forme di partecipazione meno conflittuali.
<i>Ambiente di implementazione:</i> modalità educative progressivamente più flessibili, gradualità e non impositive.	Ragazzi (15-17) con scarsa alfabetizzazione affettiva	Attraverso un accompagnamento graduale, la diffidenza iniziale può trasformarsi in una partecipazione più consapevole.	
	Ragazze (15-17) con scarsa alfabetizzazione affettiva	In un contesto percepito come meno giudicante e più rispettoso, le ragazze possono riattivare il desiderio di esplorare temi legati all'affettività e alla sessualità, precedentemente ostacolato dalla pressione sociale esercitata dal gruppo.	Recupero della partecipazione attiva; maggiore libertà di espressione; sviluppo di nuove consapevolezze rispetto alla propria affettività e ai propri diritti.

Con il progredire delle attività (t_2), le operatrici hanno avviato un progressivo riadattamento dell'intervento, trasformando il laboratorio in uno spazio più accessibile e rispettoso delle specificità culturali dei partecipanti (Tab. 3). In particolare, il percorso educativo ha potuto valorizzare il ruolo delle figure femminili, offrendo alle ragazze un'occasione di espressione personale e di riconoscimento identitario. Parallelamente, per i ragazzi si è aperto un percorso di trasformazione, lento ma significativo, nel quale la diffidenza

iniziale ha lasciato spazio a forme di partecipazione più consapevole. L'attenzione alla relazione educativa, unita alla prossimità e all'ascolto, ha così favorito una progressiva riattivazione della fiducia, della condivisione e dell'apertura.

Osservazioni conclusive

A uno sguardo d'insieme, il progetto ha funzionato: i target risultano in larga misura raggiunti e, in alcuni casi, superati.

Il caso Ri-Belli suggerisce però anche un insegnamento più generale. Una valutazione che assuma il progetto come causa del cambiamento rischia di attribuire al disegno progettuale una coerenza e una potenza che, nell'azione, non si danno mai in forma pura. Nessuna ipotesi di cambiamento, per quanto ben costruita, opera in modo indistinto in contesti diversi. Ogni progetto, nel suo farsi concreto, è piuttosto una configurazione singolare prodotta dall'incontro tra intenzioni progettuali, condizioni locali, mediazioni professionali, risposte dei destinatari, scarti e adattamenti. In questo senso, il cambiamento non va inteso come la semplice esecuzione riuscita di una logica predefinita, bensì come un esito contingente e situato, reso possibile da una combinazione specifica di condizioni che non può essere riprodotta identicamente altrove.

Emerge, inoltre, una tensione cruciale sul piano della sostenibilità. Nei contesti più fragili, il progetto ha aperto, nei minori, nuove possibilità di esperienza, riconoscimento e immaginazione del futuro; tuttavia, quando tali aperture non trovano un contesto capace di sostenerle – servizi stabili, spazi educativi, reti di supporto, condizioni materiali minime – il rischio è che il cambiamento resti esposto a regressioni o si traduca in uno scarto ancora più doloroso tra aspirazioni attivate e opportunità realmente praticabili.

Ciò non implica rinunciare a interrogare la causalità dell'intervento, ma spostarla da una logica di attribuzione lineare a una ricostruzione delle condizioni e delle mediazioni attraverso cui il progetto ha contribuito a rendere possibile il cambiamento. In questa prospettiva, il contributo della valutazione non consiste soltanto nella verifica degli esiti, ma anche nella ricostruzione delle condizioni che li hanno resi possibili. Scarti rispetto alle attese, resistenze, adattamenti e successi inattesi non vanno trattati come residui, bensì come materiali conoscitivi centrali. Più che offrire un modello replicabile in senso stretto, il caso Ri-Belli consente di riconoscere quali condizioni, soglie di accesso e forme di mediazione meritino di essere considerate nella progettazione e nella valutazione di interventi futuri.

Dichiarazioni

L'autrice dichiara di aver utilizzato il Software ChatGPT (OpenAI) esclusivamente nella fase finale di revisione del contributo, come supporto al controllo di refusi e alla revisione formale del testo. L'elaborazione scientifica del contributo, l'analisi e l'interpretazione dei risultati sono state svolte interamente dall'autrice, che ha inoltre verificato integralmente la versione finale del testo e se ne assume la piena responsabilità.

Riferimenti bibliografici

- Banks, J.A. (2004). Multicultural education: Historical development, dimensions, and practice. In: Banks, J.A. and Banks C.A.M., a cura di, *Handbook of Research on Multicultural Education*. 2nd ed. San Francisco: Jossey-Bass, pp. 3-29.
- Baumrind, D. (1967). Childcare practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, 75, 1: 43-88. DOI: 10.1037/h0024919
- Berry, J.W. (1997). *Immigration, acculturation, and adaptation*. *Applied Psychology: An International Review*, 46, 1: 5-34. DOI: 10.1111/j.1464-0597.1997.tb01087.x.
- Bourdieu, P. (1986). The forms of capital. In: Richardson J.G., a cura di, *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood Press, pp. 241-258.
- Brodtkin, E.Z. (2008). Accountability in street-level organizations. *International Journal of Public Administration*, 31, 3: 317-336. DOI: 10.1080/01900690701590587
- Camus, A. (1951). *L'homme révolté*. Paris: Gallimard (trad. it.: L'uomo in rivolta. Milano: Bompiani, 1981).
- Celardi, E. (2025). *Riconoscere e valutare il cambiamento. Diario di un'esperienza di contrasto alla povertà educativa*. Milano: FrancoAngeli.
- Celardi, E., Mazzeo Rinaldi, F., Occhipinti, O. (2026). Il ruolo del contesto negli interventi di contrasto alla povertà educativa. Un'esperienza di ricerca. In: Gozzo S., a cura di, *La povertà educativa da una prospettiva sociologica*. Milano: FrancoAngeli, pp. 73-90.
- Hirschman, A.O. (1967). *Development Projects Observed*. Washington, DC: Brookings Institution.
- Klaus, H. (1998). *The International Teen STAR Program*. Bethesda, MD: Teen STAR Program. ERIC Document ED424510.
- Koleros, A., Mulkerne, S., Oldenbeuving, M., Stein, D. (2020). The actor-based change framework: A pragmatic approach to developing program theory for interventions in complex systems. *American Journal of Evaluation*, 41, 1: 34-53. DOI: 10.1177/1098214018786462
- Lipsky, M. (2010). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. 30th anniversary expanded ed. New York: Russell Sage Foundation.
- Patton, M.Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.
- Pawson, R. (2006). *Evidence-Based Policy: A Realist Perspective*. London: Sage.
- Pawson, R., Tilley N. (1997). *Realistic Evaluation*. London: Sage.
- Petino, G., Wilson, J.S., Torre, S. (2022). La fascia trasformata siciliana: il territorio tra crisi socioecologica e vulnerabilità. In: Amato F., Amato V., de Falco S., La Foresta D. and Simonetti L., a cura di, Catene/Chains. *Memorie geografiche* NS 21. Firenze: Società di Studi Geografici, pp. 489-494.

Save the Children Italia (2023). *Piccoli schiavi invisibili. Dentro lo sfruttamento: un'indagine sui figli dei braccianti a Latina e Ragusa*. Roma: Save the Children Italia - ETS.

Stame, N. (2019). Doubt, Surprise, and the Ethical Evaluator: Lessons from the Work of Judith Tandler. *Evaluation*, 25, 4: 449-468. DOI: 10.1177/1356389019879458

Stame, N. (2022). Tra possibilismo e valutazione. Judith Tandler e Albert Hirschman. Soveria Mannelli: Rubbettino.

Tandler, J. (1992). *Progetti ed effetti. Il mestiere di valutatore*. Napoli: Liguori.

Weiss, C.H. (1997). *Theory-based evaluation: Past, present, and future*. *New Directions for Evaluation*, 76: 41-55. DOI: 10.1002/ev.1086

L'infrastrutturazione sociale come leva di sviluppo: Fondazione con il Sud a Castelbuono

*Alfonsina Mastrolia**, *Marco Palmieri***, *Davide De Valeri****

The article presents the results of an ex-post evaluation of the local development program promoted by Fondazione con il Sud in the territory of Castelbuono. The aim of the work is to analyse the extent to which the intervention contributed to strengthening local social infrastructure, understood as the consolidation and reorganization of collaborative relationships among territorial actors. The evaluation design integrates Social Network Analysis and Nominal Group Technique within a mixed methods perspective, combining structural network analysis and a participatory device for constructing evaluative judgment. The hybridization of quantitative and qualitative tools enables the articulation of objectification of relational configurations with dialogical interpretation of observed changes, in accordance with the participatory evaluation approach. The article proposes a methodological reflection on the integrated use of SNA in evaluation contexts, highlighting its potential in producing shared knowledge about local development processes.

Keywords: participatory evaluation; social network analysis; local development program; social infrastructure; mixed methods.

Il modello d'intervento di Fondazione con il Sud

La Fondazione con il Sud è un ente privato nato nel 2006 nell'ambito di un'alleanza tra le fondazioni bancarie e le reti nazionali del terzo settore,

* Alfonsina MASTROLIA, Sapienza Università di Roma (alfonsinamastrolia@gmail.com)

** Marco PALMERI, Sapienza Università di Roma (marco.palmieri@uniroma1.it)

*** Davide DE VALERI, Fondazione con il Sud (d.devaleri@fondazioneconilsud.it)

Invio Proposta: 20-05-2026. Accettazione: 22-07-2026.

finalizzata a promuovere l'infrastrutturazione sociale e lo sviluppo socioeconomico del Mezzogiorno d'Italia.

La Fondazione con il Sud sostiene progetti promossi da organizzazioni della società civile in cinque ambiti principali: i. l'educazione dei giovani, ii. lo sviluppo del capitale umano di eccellenza, iii. la valorizzazione dei beni comuni, iv. l'innovazione dei servizi sociosanitari, v. l'accoglienza e l'integrazione dei cittadini stranieri. Tra il 2006 e il 2025 la Fondazione ha erogato oltre 335 milioni di euro per sostenere circa 2.000 iniziative che hanno coinvolto oltre 7.000 enti pubblici, privati e del terzo settore.

L'intervento della Fondazione può essere letto alla luce di modelli valutativi basati sulla teoria¹, fondati su uno schema che collega territori, azioni e cambiamenti. La Fondazione opera secondo un approccio strutturale e integrato di analisi dei territori, secondo il quale le fragilità che caratterizzano il Mezzogiorno – povertà, esclusione sociale, degrado ambientale, marginalità culturale, ecc. – non sono riconducibili esclusivamente a bisogni individuali o alla carenza di servizi, ma riflettono una debolezza strutturale degli ecosistemi territoriali. Debolezza che si manifesta nella frammentazione delle reti tra attori pubblici e privati, nella mancata valorizzazione delle risorse locali e nella carenza di capitale sociale, inteso come insieme di relazioni fiduciarie, norme condivise e capacità collettive di azione (Putnam, 1993; Cartocci, 2007; Svimez, 2024). La Fondazione identifica nel terzo settore, o meglio nel terzo settore specializzato, l'attore centrale dei processi di cambiamento promossi attraverso le diverse linee di finanziamento programmatiche. Le organizzazioni del terzo settore, soprattutto quando specializzate in ambiti tematici definiti, sono considerate attori in grado di svolgere un ruolo di attivazione e coordinamento degli altri soggetti comunitari: cittadini, enti pubblici, scuole, imprese. In questa prospettiva, l'intervento della Fondazione persegue una catena causale che va dalla costruzione di partenariati multi-attore, allo sviluppo di infrastrutture sociali in termini di servizi educativi e sanitari, all'attivazione di spazi comunitari e di modelli di governance condivisi, all'innovazione delle modalità di risposta ai bisogni, fino all'attivazione e alla valorizzazione di risorse e competenze locali.

L'ipotesi causale che orienta la missione della Fondazione considera il rafforzamento del terzo settore e la costruzione di reti collaborative territoriali elementi propedeutici alla nascita di infrastrutture sociali permanenti, in grado di aumentare la capacità delle comunità di rispondere collettivamente ai bisogni sociali, culturali, ambientali ed economici. In questa prospettiva, i diversi ambiti di finanziamento non costituiscono iniziative autonome, bensì diverse

¹ *Theory-Based Evaluation* in: Chen, 1990; Weiss, 1995; Stame, 2004.

modalità attraverso le quali la Fondazione persegue un medesimo obiettivo: la costruzione di infrastrutture sociali. I progetti, nel loro insieme di risorse, azioni e relazioni, rappresentano i beni abilitanti intorno ai quali si organizzano reti di collaborazione, servizi di prossimità, percorsi di autonomia e nuove forme di partecipazione civica. Ne deriva un'azione articolata su tre livelli. A livello micro, i progetti promuovono cambiamenti nelle condizioni di vita delle persone, aumentando l'autonomia, le competenze, l'inclusione e la capacità di autodeterminazione. A livello meso, rafforzano le organizzazioni del terzo settore, consolidano partenariati e generano nuove infrastrutture sociali. A livello macro, contribuiscono alla costruzione di ecosistemi territoriali più coesi, collaborativi e resilienti, in cui la cooperazione tra attori pubblici e privati diviene una componente strutturale dello sviluppo locale.

Attraverso il monitoraggio e la valutazione dei progetti finanziati, la Fondazione svolge un esercizio di apprendimento continuo per una riprogettazione costante dei propri strumenti di finanziamento, affinché dalle sperimentazioni progettuali possano nascere modelli di cambiamento efficaci e replicabili.

1. La linea di finanziamento “Sviluppo locale” nel comune di Castelbuono (PA)

Lo “Sviluppo locale” è una linea di finanziamento di Fondazione con il Sud, non ancorata a uno specifico settore tematico; sostiene iniziative di sviluppo economico e sociale promosse da partenariati ampi. È una linea di finanziamento promossa in diverse edizioni che, a partire dal 2008, individua territori di intervento da cui far emergere progettualità di inclusione e di sviluppo. Nel corso del tempo, si è passati da una logica di bando, che metteva in competizione le reti del territorio per la presentazione dei progetti, a una logica di coprogettazione, che ha visto la Fondazione impegnata fianco a fianco con le organizzazioni locali.

Nel 2013, sulla base dei dati con cui sono stati costruiti i Sistemi Locali del Lavoro dell'Istat, in particolare quelli relativi al disagio economico e sociale e alle organizzazioni del terzo settore, i comuni di Castelbuono e Isnello, in provincia di Palermo, sono stati scelti dalla Fondazione per la realizzazione di un programma di sviluppo locale del valore di 5 milioni di euro. La scelta del territorio è stata determinata dalla presenza di risorse locali che consentissero l'avvio di processi di sviluppo, dalla capacità di aggregare soggetti istituzionali e della società civile, da una buona connotazione identitaria. I comuni di Castelbuono e Isnello sono stati scelti anche per le loro

potenzialità turistiche, la posizione geografica, la natura di borghi medievali, un settore agroalimentare dinamico.

L'iniziativa ha previsto la consultazione degli attori locali, la raccolta di idee e l'elaborazione di progettualità condivise, attraverso una fase di innesto, una di consultazione pubblica e una di coprogettazione esecutiva. Il programma ha avuto inizio con una fase di innesto, condotta dalla Fondazione e finalizzata a coinvolgere la cittadinanza in un processo di costruzione condivisa delle linee di sviluppo che avrebbero dato origine ai progetti da finanziare. È stata tenuta un'assemblea presso la sede del consiglio comunale di Castelbuono, alla quale hanno partecipato i soggetti, le imprese, il terzo settore e le istituzioni locali. Al primo momento di confronto sono seguiti numerosi incontri, che hanno circoscritto la progettazione in quattro ambiti tematici: a) agricoltura sociale; b) servizi alla persona; c) raccolta differenziata e riciclo dei rifiuti; d) cultura, enogastronomia e turismo.

Successivamente è stato attivato un percorso di coprogettazione, supportato dal finanziamento di alcune iniziative di animazione comunitaria: installazione di una rete Wi-Fi comunale pubblica, animazione digitale, allestimento di un centro aggregativo, attività di allestimento urbano con creativi e designer. La coprogettazione nell'ambito ambientale, come la raccolta differenziata, è stata interrotta per assenza di interesse da parte del territorio. L'ambito d'intervento sull'agricoltura sociale è stato sviluppato in due direzioni: una collegata alla produzione orticola e all'inclusione sociale di soggetti svantaggiati, e l'altra volta al recupero della produzione e della lavorazione tradizionali della manna da frassino. Gli ambiti del turismo e dei servizi alla persona hanno visto protrarsi una serie di incontri di consultazione interni al territorio.

Il vincolo di partecipazione attiva, imposto per presenziare ai tavoli, assieme alla complessità di una progettazione che tenesse insieme la molteplicità di istanze e punti di vista delle due comunità coinvolte, ha portato al progressivo disimpegno della componente pubblica e del terzo settore del Comune di Isnello, lasciando il territorio di Castelbuono come unico interlocutore della Fondazione. L'esito della coprogettazione è stato l'individuazione di un gruppo di giovani motivati e competenti, che si sono messi a disposizione per le attività progettuali. La Fondazione con il Sud è stata dunque in grado di far emergere la capacità progettuale dei soggetti locali, decidendo di finanziare i progetti che nascevano dal basso e avevano una loro storia e un loro percorso alle spalle. Al termine della coprogettazione, è stato finanziato un intervento di agricoltura sociale denominato "AGRI-ETICA Promozione di un Sistema a Rete di Agricoltura Sociale", nell'ambito del quale è stato assegnato un contributo di €1.230.000 al Consorzio di

cooperative sociali Madonie 33. A novembre 2015 è stato completato il percorso di progettazione esecutiva di un secondo intervento di agricoltura sociale denominato “Manna nuovo benessere – Quando la Manna non cade dal cielo”, nell’ambito del quale è stato assegnato un contributo di €1.170.000 al neocostituito Consorzio sociale Manna Madonita.

Nel 2018 è stata completata l’ultima linea di progettazione, relativa a un percorso di valorizzazione turistica in chiave inclusiva, collegato al festival internazionale Ypsigrock. Al progetto “Tutti Inclusi”, promosso dall’associazione culturale Glenn Gould, è stato assegnato un contributo di €850.000.

Il progetto AGRI-ETICA è stato realizzato grazie alla collaborazione tra soggetti profit e non profit, con l’obiettivo di sperimentare un modello di agricoltura sociale che coniugasse sviluppo rurale ed economico con azioni di riabilitazione, inclusione sociale e inserimento lavorativo di soggetti svantaggiati o a rischio di esclusione sociale. Nell’ambito del progetto erano previsti un caseificio per la realizzazione di prodotti lattiero-caseari, un frantoio per la lavorazione delle olive e l’imbottigliamento dell’olio, e interventi per la lavorazione dei prodotti orticoli e per la trasformazione in conserve.

Il progetto Manna nuovo benessere “Quando la Manna non cade dal cielo” è stato concepito per ripristinare la produzione della manna da frassino (coltura tipica dell’area di Castelbuono) come risorsa da valorizzare per lo sviluppo economico e occupazionale del territorio. Il progetto ha previsto la rimessa in coltura di terreni con presenza di frassineti abbandonati e il trasferimento di competenze e conoscenze da parte dei produttori di manna più anziani alle giovani generazioni. Il progetto ha previsto la costituzione di un consorzio incaricato della lavorazione e della vendita del prodotto.

Il progetto Tutti Inclusi ha inteso valorizzare le potenzialità dell’indotto economico, sociale e turistico, promuovendo la partecipazione inclusiva e l’accessibilità ai luoghi di interesse storico, artistico e religioso di Castelbuono. Tutti gli interventi hanno stimolato la collaborazione tra le organizzazioni territoriali, in particolare del terzo settore, determinando la costituzione di reti ampie (composte sempre da oltre 10 organizzazioni), sia rafforzando le collaborazioni già esistenti sia creandone di nuove.

2. La valutazione a posteriori del programma di sviluppo locale a Castelbuono: le domande valutative, il disegno della ricerca, l’ibridazione delle tecniche

Il disegno di valutazione a posteriori è stato progettato per perseguire la dimensione valutativa dell’efficacia esterna, «che consiste nel confronto

degli effetti ottenuti dal programma o intervento, specialmente i risultati e gli impatti, con il contesto economico-sociale e i suoi bisogni, ovvero con la domanda iniziale rispetto alla quale quel programma o intervento è stato concepito» (Bezzi, 2021, p. 59). La valutazione si è posta i seguenti interrogativi.

- Il programma di sviluppo locale ha contribuito allo sviluppo dell'infrastrutturazione sociale del territorio?
- Ha implementato le reti di collaborazione tra le organizzazioni attive a Castelbuono?
- Ha promosso la propensione delle cooperative sociali del territorio a lavorare in rete?
- Ha favorito l'integrazione tra le organizzazioni del terzo settore castelbuonese, anche con le altre realtà profit e non profit del territorio?

Si è deciso di adottare un approccio partecipato dagli stakeholder per valutare la rilevanza dei cambiamenti generati dagli interventi e la loro capacità di soddisfare i bisogni dei destinatari. Questa strategia mira a cogliere il valore che il programma ha avuto per la comunità locale, a valorizzare il punto di vista dei beneficiari e a stabilire la relazione tra l'intervento e i cambiamenti prodotti, mettendo a confronto punti di vista differenti (anche conflittuali) tra i portatori di interesse. Tra i modelli di valutazione partecipata riportati in letteratura (Foresti, 2003), il nostro può considerarsi un modello consultivo², per l'elevato coinvolgimento degli stakeholder nelle attività di ricerca sul campo, funzionali al ricercatore per porsi in ascolto di ciò che i suoi intervistati hanno da riferire, e per un'interpretazione dei risultati discussa in gruppo con i vari portatori di interesse.

Da un punto di vista metodologico, si è deciso di integrare strumenti di ricerca diversificati per rispondere alle domande valutative stabilite dalla committenza, secondo il principio dell'ibridazione delle tecniche e non della loro triangolazione³. La triangolazione è un'espressione molto utilizzata, anche se impropriamente, nelle scienze sociali e si riferisce alla strategia di indagare lo stesso fenomeno da punti di osservazione indipendenti, ricorrendo a tecniche diverse. «Se un'ipotesi supera un controllo effettuato mediante una serie di tecniche complementari, ottiene un grado di conferma tale che non può mai essere raggiunto dall'applicazione di una singola tecnica»

² Sugli altri modelli di valutazione partecipata, come la partecipazione generica e la valutazione democratica, vedi Foresti (2003).

³ «Mutuata dalla geodetica, la parola "triangolazione" è riferita alla procedura seguita per stabilire l'esatta posizione di un dato oggetto partendo da più punti di osservazione: usare più di un punto di riferimento permetterebbe una maggiore accuratezza nel processo di rilevazione» (Martire, Parra Saiani e Cataldi, 2023, p. 223).

(Webb *et al.*, 1966, p. 174; poi citata da Denzin, 1978, p. 302). La critica principale rivolta alla triangolazione è che mira alla convergenza dei risultati della ricerca, obiettivo sopravvalutato. In realtà, le tecniche qualitative e quantitative non sono in competizione tra loro. Anzi, esse consentono di studiare lo stesso fenomeno fornendo rappresentazioni diverse (Greene, 2007). Invece di puntare alla convergenza dei risultati, è più utile per il ricercatore trasformare queste differenze in un'occasione per valorizzare la divergenza prodotta, così da superare l'idea che combinare tecniche diverse aiuti a confermare l'unica ipotesi sottoposta al vaglio empirico e rafforzare la posizione dell'arricchimento della comprensione attraverso prospettive multiple.

I committenti richiedono risposte certe e univoche, ma la ricerca spesso produce risultati contrastanti, dissonanti e divergenti, che costituiscono il potenziale generativo della scienza valutativa, perché alimentano il confronto tra chi la pensa diversamente. Da ciò nasce l'esigenza di un approccio di ricerca misto, che ibrida le tecniche, una strategia *mixed methods* (Amaturo e Punziano, 2016) «nella quale un ricercatore o un gruppo di ricercatori combina elementi di ricerca qualitativa e quantitativa (incluse visioni del mondo, raccolta analisi e tecniche di inferenza sui dati) al fine di migliorare la ricerca stessa in termini di ampiezza, profondità e validità nei risultati prodotti e nei livelli di comprensione raggiunti, sia in un singolo studio che in più studi tra loro correlati» (Johnson *et al.*, 2007, p. 129). Grazie a un approccio *mixed methods*, questo studio mostra i cambiamenti che l'intervento di Fondazione con il Sud ha prodotto nell'infrastrutturazione sociale di Castelbuono, la complessità del contesto in cui questi cambiamenti hanno preso forma e il meccanismo che li ha attivati: lo sviluppo locale è l'esito dell'attivazione di un capitale sociale comunitario, che si identifica con la qualità delle relazioni tra gli attori che insistono sullo stesso territorio: la capacità di cooperare assieme, la fiducia reciproca, le esperienze passate di collaborazione, la densità delle reti formali che connettono persone, organizzazioni profit e non profit e istituzioni. Lo sviluppo, in questa accezione, è un processo che prende forma nelle singole comunità storicamente situate, perché sono le comunità territoriali gli agenti autentici dei loro percorsi di trasformazione. Le condizioni di partenza contano, ma ciò che fa la differenza è la capacità dei soggetti locali di mobilitarsi intorno a obiettivi condivisi.

Alla luce di queste riflessioni, si è ricorso a due tecniche distinte, secondo il principio che «i difetti di una tecnica sono spesso la forza di un'altra e gli osservatori, combinandole, possono raggiungere il meglio di ognuna» (Parra Saiani, 2004, p. 31). Innanzitutto, si è ricorso alla social network analysis (SNA), che, applicata agli studi di valutazione, focalizza l'analisi sul sistema di interconnessione tra gli attori sociali che operano nel contesto sociale in

cui si svolge il programma. In questo disegno di ricerca, la SNA è stata utile per valutare l'impatto dei finanziamenti della Fondazione sull'infrastrutturazione sociale del territorio. Con il termine "infrastrutturazione sociale" intendiamo il processo attraverso cui le relazioni tra gli attori territoriali si consolidano, si specializzano e assumono carattere stabile, creando una base relazionale che facilita la collaborazione e la condivisione delle risorse (Triglia, 2005). Il concetto di infrastrutturazione sociale, tuttavia, rimanda a una dimensione più ampia, richiamando il capitale sociale, inteso come l'insieme di elementi relazionali – fiducia, aspettative, reciprocità – che valorizzano la relazionalità sociale (Donati e Colozzi, 2006), cui si affiancano la capacità di azione collettiva e la sostenibilità delle relazioni nel tempo. In questo lavoro si è scelto di circoscrivere l'osservazione empirica alle collaborazioni formalizzate e ai relativi indici strutturali di rete (densità, centralità, transitività, *k-core*). Tale scelta deriva dall'ipotesi secondo cui la strutturazione e l'istituzionalizzazione delle collaborazioni costituiscono una dimensione osservabile dell'infrastrutturazione sociale. Una rete densa, caratterizzata da opportunità di coordinamento, descrive le condizioni strutturali che possono favorire l'emergere di capacità collettive di azione. È tuttavia necessario precisare che l'analisi di rete, pur fornendo una descrizione delle configurazioni relazionali, non consente, di per sé, di inferire un rafforzamento dell'infrastrutturazione sociale nella sua accezione più ampia. L'integrazione della SNA con strumenti qualitativi consente di colmare questa lacuna, mettendo in luce i meccanismi alla base dei cambiamenti attivati dal capitale sociale della comunità (Lo Presti, Dentale e Celardi, 2018). Per questo motivo, nel presente lavoro il giudizio valutativo ha richiesto l'uso combinato della SNA con la tecnica del gruppo nominale.

Lo scopo è stato comprendere in che misura il programma di sviluppo locale sia riuscito a incrementare la propensione delle cooperative sociali e del tessuto associativo a collaborare in rete, promuovendo l'integrazione tra le diverse realtà del terzo settore. Con questo obiettivo, sono state condotte 28 interviste tramite un questionario sociometrico, somministrato dal ricercatore ai referenti delle organizzazioni attive sul territorio. Ogni intervistato ha indicato i propri partner collaborativi, consentendo di identificare ulteriori 33 organizzazioni che, pur non essendo state intervistate⁴, risultano strutturalmente rilevanti per la comprensione dell'architettura relazionale territoriale. L'unità di analisi della rete è costituita da organizzazioni attive nel

⁴ Queste organizzazioni sono state contattate per partecipare all'indagine ma i referenti hanno rifiutato l'intervista.

territorio di Castelbuono, appartenenti al terzo settore, alle istituzioni pubbliche locali, ai consorzi, alle cooperative sociali e agli attori economici privati.

Distinguendo tra le collaborazioni promosse dalla Fondazione e quelle attivate in sua assenza, la rilevazione si è concentrata sulla presenza di collaborazioni formalizzate. Agli intervistati è stato chiarito che una “collaborazione formalizzata” è una collaborazione disciplinata da un rapporto formale che prevede la realizzazione di attività operative; porta alla creazione di sistemi interdipendenti tra le organizzazioni coinvolte (ad es. partenariati); le decisioni sono condivise e prese in modo consensuale; i rapporti sono scritti e formalizzati; la comunicazione è gestita in modo professionale; ruoli, tempi e procedure di valutazione sono formalizzati in base a parametri e soglie di impatto.

Successivamente, si è impiegata la tecnica valutativa del gruppo nominale (Bertin, 2011), adottando un disegno di ricerca sequenziale, in cui tecniche diverse sono utilizzate in successione. Il gruppo nominale ha previsto una prima fase direttiva, durante la quale i partecipanti hanno compilato individualmente scale di valutazione relative al raggiungimento degli obiettivi dell'intervento. La seconda fase è stata semidirettiva, poiché gli stessi partecipanti sono stati chiamati a discutere in gruppo gli esiti aggregati delle valutazioni fornite in precedenza individualmente. Il pregio della tecnica si può così riassumere: se i partecipanti valutano l'intervento in modo uniforme, ciò sta a significare che il gruppo condivide la medesima rappresentazione dell'impatto prodotto (in questo caso c'è consenso in merito al fatto che gli esiti sono stati positivi o negativi); se i partecipanti invece valutano gli esiti dell'intervento in modo difforme, allora il gruppo non condivide il medesimo punto di vista (in questo secondo caso non c'è consenso sull'esito) (Bezzi, 2013). «Naturalmente non è affatto detto che la discussione della seconda fase porti a un accordo nel gruppo» (Bezzi, 2021, p. 147). Anzi, «l'interazione fra gli attori non deve essere indirizzata alla costruzione del consenso, ma allo sviluppo di processi di senso che contribuiscono alla costruzione di un linguaggio, dei concetti e di una base informativa e decisionale condivisa fra gli attori». (Bertin, 2011, p. 48). A Castelbuono, la tecnica del gruppo nominale ha avuto lo scopo di attivare processi di riflessività, ricostruendo le criticità dei meccanismi attivati nel corso del programma, in un'ottica di autoapprendimento.

Nel ruolo di esperti, hanno partecipato al gruppo nominale i referenti dei progetti e i loro principali partner, per un totale di nove persone, invitati a valutare individualmente e poi a dialogare tra loro sugli esiti d'impatto del programma di sviluppo. Questo articolo riporta gli esiti della valutazione relativa al conseguimento dell'obiettivo strategico-generale di incrementare

l'infrastrutturazione sociale del territorio, rafforzando la collaborazione tra le organizzazioni del territorio e la cooperazione tra queste, l'amministrazione locale e gli altri enti pubblici di Castelbuono.

3. L'ibridazione metodologica nel processo valutativo

3.1 *Trasformazioni delle configurazioni reticolari: l'analisi strutturale*

L'analisi delle reti di collaborazioni formalizzate è stata sviluppata all'interno di un approccio valutativo finalizzato non soltanto a descrivere le configurazioni reticolari esistenti, ma anche a osservare le condizioni che rendono possibile il cambiamento, in linea con la tradizione valutativa che interpreta i programmi di intervento come processi generativi (Stame, 2016). La SNA è stata impiegata per analizzare in che modo il programma di sviluppo locale (Lo Presti, 2008) abbia inciso sulle strutture relazionali entro cui operano gli attori territoriali, trasformando il dato relazionale in oggetto di riflessione condivisa tra valutatori e la comunità locale (Palumbo, 2011). L'analisi, collocandosi sul versante dell'efficacia esterna dell'intervento, si è concentrata sulle collaborazioni formalizzate e operative tra le organizzazioni attive nel territorio di Castelbuono.

L'indagine ha adottato un'impostazione comparativa rispetto a tre configurazioni reticolari: la rete territoriale preesistente alle iniziative finanziate, la rete delle collaborazioni formalizzate attivate nell'ambito dei progetti sostenuti dalla Fondazione, la rete complessiva risultante dall'integrazione delle due. Questa distinzione consente di osservare le modalità con cui l'intervento si innesta nelle reti già esistenti, modificandole.

La costruzione della base empirica si è svolta in tre fasi: *i*. Fondazione con il Sud ha fornito un primo elenco di enti/organizzazioni tratto dai propri archivi; *ii*. l'elenco è stato verificato, ampliato e aggiornato attraverso il dialogo e il coinvolgimento degli stakeholder locali e dei principali referenti dei progetti; *iii*. sono stati eliminati i duplicati, esclusi gli enti/organizzazioni che non operano a Castelbuono, e sono stati mantenuti solo gli attori attivi sul territorio. La rilevazione dei dati si è articolata in tre fasi temporalmente distinte:

1. durante la prima fase (maggio 2024) sono state condotte interviste individuali faccia a faccia con i referenti delle organizzazioni e delle associazioni locali direttamente coinvolte nel programma di sviluppo locale;

2. nella seconda fase (luglio-novembre 2024) le interviste sono state condotte a distanza, contattando i referenti di altri enti e organizzazioni del territorio che non avevano ricevuto i finanziamenti da Fondazione;
3. la terza fase (novembre 2024) ha previsto interviste ai referenti di quelle organizzazioni che, grazie alle interviste precedentemente svolte, erano risultate essere centrali nella rete di collaborazioni formalizzate.

L'esito finale è una lista di 61 enti/organizzazioni, tra cui figurano istituzioni pubbliche (Comune di Castelbuono), attori culturali (Museo Civico Castello, Museo Naturalistico, Associazione Glenn Gould), cooperative sociali, consorzi, attori economici e altri soggetti del terzo settore.

Il confronto tra le tre configurazioni ha consentito di osservare variazioni quantitative nel numero di collaborazioni e trasformazioni strutturali nella densità della rete, nella configurazione del nucleo centrale e nei ruoli di intermediazione. L'analisi comparativa consente di comprendere se e in che misura il programma (Palumbo, 2001) abbia contribuito a rafforzare l'infrastrutturazione sociale del territorio, qui osservata – coerentemente con quanto precisato in apertura – nella sua componente strutturale e formalizzata, ossia nel consolidamento, nell'integrazione e nella specializzazione delle relazioni collaborative (Long, Hibbert e Braithwaite, 2016).

I risultati della SNA hanno fornito una rappresentazione strutturale delle trasformazioni avvenute nel sistema relazionale territoriale, costituendo la base empirica su cui si è innestato l'approfondimento valutativo successivo. L'analisi della densità della rete ha consentito di osservare in che misura il programma di sviluppo locale abbia inciso sull'intensità delle collaborazioni formalizzate tra le organizzazioni del territorio. Nell'ottica della valutazione partecipata (Palumbo, 2003; Tomei, 2004; Cousins e Whitmore, 1998), la densità è intesa come indicatore della frequenza e della sistematicità delle collaborazioni formalizzate, ossia di una condizione relazionale che può facilitare processi di coordinamento, apprendimento e azione collettiva. La densità costituisce, inoltre, un elemento di trasparenza che consente agli attori territoriali di riconoscere la propria collocazione nel sistema complessivo delle relazioni e di interrogarsi non solo sulle collaborazioni attivate, ma anche su quelle potenzialmente attivabili.

Il primo risultato rilevante riguarda l'aumento della densità della rete complessiva rispetto alla configurazione territoriale preesistente. La rete che integra le relazioni attivate dai progetti finanziati segnala una crescita quantitativa dei legami e una maggiore strutturazione delle relazioni collaborative, con una riduzione delle organizzazioni isolate o marginali e una più

ampia interconnessione tra soggetti del terzo settore e, in alcuni casi, con il mondo del profit e con le istituzioni locali. La tabella 1 riporta i principali indici strutturali delle tre configurazioni di rete messe a confronto.

Tab. 1 - Indici strutturali delle tre reti a confronto

	<i>Rete preesistente (senza FCS)</i>	<i>Rete attivata dal pro- gramma di sviluppo locale</i>	<i>Rete complessiva (con FCS)</i>
n. di nodi	61	31	61
n. di relazioni	260	108	362
grado medio	4,26	3,48	5,93
indice <i>K-core</i>	5	4	6
centralizzazione	0,32	0,37	0,31
densità	0,071	0,116	0,099
componenti	7	1	2
transitività	0,28	0,21	0,34
distanza media	2,66	2,53	2,53

Fonte: elaborazioni degli autori

L'analisi comparativa evidenzia che la rete territoriale preesistente già presentava una base relazionale ampia (61 nodi, organizzazioni attive sul territorio), ma con una densità⁵ contenuta (0,071) e una struttura frammentata in 7 componenti⁶. Questo assetto segnala l'esistenza di un tessuto collaborativo, presente ma caratterizzato da legami selettivi distribuiti in modo diseguale, come si osserva nella figura 1.

Con 260 legami formalizzati e una media di 4,26 collaborazioni per attore, le collaborazioni operative erano centralizzate attorno a un numero limitato di organizzazioni – Museo Civico Castello, Museo Naturalistico, Associazione Ristoratori e Comune, insieme ad alcuni attori economici storici del territorio come Fiasconaro e l'agriturismo Bergi – mentre una parte significativa degli attori occupava posizioni periferiche. L'indice *k-core*⁷

⁵ La densità è il rapporto tra il numero di relazioni presenti rispetto a quelle attivabili: in questo caso la densità rappresenta il valore del numero di collaborazioni formalizzate attivate rispetto al numero massimo di collaborazioni che si sarebbero potute attivare.

⁶ La componente è l'area del grafo in cui ogni nodo può raggiungere tutti gli altri attraverso uno o più percorsi, ma non può raggiungere i nodi di un'altra componente. Il grafo mostra 7 componenti: ci sono 5 nodi isolati e un sottogruppo di 2 attori.

⁷ L'indice *k-core* mostra il grado di struttura centro-periferia del reticolo. In questa rete 13 attori costituiscono il nucleo centrale della rete, perché ciascuno di essi dispone di almeno 5 connessioni dirette con altri attori del reticolo. In generale, un grafo è *k-core* se ha almeno *k*

beneficiari e loro partner diretti), e a 108 legami. Nonostante il numero inferiore di attori e di relazioni, questa configurazione presenta una densità superiore (0,116) rispetto alla rete preesistente. La struttura della rete risulta coesa, organizzata in un'unica componente connessa, con una media di 3,48 collaborazioni formalizzate per attore e un indice *k-core* di 4: il nucleo centrale è composto da 8 organizzazioni (Agry Social, Glenn Gould, Consorzio Manna, Il Sorriso, Consorzio Agrietica, Bergi, Comune, Fiasconaro) che costituiscono il cuore operativo della rete attivata dal programma.

Questo dato indica che i progetti hanno funzionato come dispositivi di attivazione di collaborazioni strutturate, favorendo la costruzione di collaborazioni operative tra organizzazioni che condividono obiettivi comuni. Le reti attivate nell'ambito di "Tutti Inclusi", "Agrietica" e "Manna" si articolano attorno a tematiche distinte, evidenziando una specializzazione settoriale. Da un lato il cluster culturale, centrato su Glenn Gould (in connessione con Museo Naturalistico, Pro Loco, AVY, Aquilone e altri partner operativi), dall'altro il cluster agricolo e cooperativo, imperniato su Agry Social e Consorzio Agrietica (in relazione con Consorzio Manna, Bergi, Il Sorriso e il Comune). La centralizzazione (0,37), superiore a quella della rete preesistente, indica una maggiore concentrazione delle relazioni attorno a specifici hub progettuali. La distanza media di 2,53 indica percorsi più brevi tra gli attori e, di conseguenza, una maggiore efficienza operativa interna. L'indice di transitività, invece, non mostra differenze significative rispetto alla rete preesistente. Come mostra la figura 2, la rete di Fondazione si configura come una rete con 31 attori connessi in un'unica componente. La maggiore densità di questa rete, pur limitata alle organizzazioni coinvolte nei finanziamenti, evidenzia come il programma abbia incentivato modalità di collaborazione formale più intense rispetto a quanto accadeva a Castelbuono negli anni precedenti. La rete complessiva, risultante dall'integrazione tra la rete territoriale preesistente e quella attivata dai progetti finanziati, mostra un incremento della densità relazionale rispetto alla configurazione originaria. Con 61 nodi e 362 legami, la rete complessiva presenta un grado medio di 5,93 e un indice *k-core* di 6, a indicare un rafforzamento sostanziale della coesione strutturale e l'emergere di un nucleo centrale robusto. La rete complessiva registra un incremento del 39% dei legami rispetto alla rete preesistente. La densità (0,099) risulta superiore a quella della rete preesistente (0,071), ma inferiore a quella della rete di Fondazione (0,116), poiché l'integrazione include sia le relazioni strutturate attivate dai progetti sia quelle preesistenti, producendo una densità media che riflette la coesistenza di legami di natura diversa.

un accentramento dei flussi relazionali attorno a pochi attori, bensì di un incremento diffuso delle connessioni all'interno del sistema. L'aumento della chiusura triadica segnala, infatti, una maggiore propensione alla formazione di reti più dense e integrate.

In sintesi, l'analisi mostra che il programma ha intensificato le collaborazioni formalizzate, ridotto la frammentazione e rafforzato il nucleo centrale della rete, senza produrre una concentrazione relazionale. L'aumento della transitività e la riduzione della distanza media indicano una maggiore interconnessione distribuita, che non dipende da pochi nodi dominanti. L'incremento della densità osservato nella rete complessiva non fornisce indicazioni sulla qualità delle collaborazioni attivate, né sulla loro sostenibilità nel tempo al di fuori del contesto di finanziamento⁹. Inoltre, una maggiore densità può riflettere sia un rafforzamento delle capacità collaborative sia una concentrazione delle relazioni all'interno di circuiti progettuali ristretti. Per queste ragioni, i risultati della SNA costituiscono una base informativa necessaria, ma non sufficiente, per formulare un giudizio complessivo sull'efficacia del programma. Coerentemente con l'approccio della valutazione pluralista (Stame, 2016), tale risultato segnala condizioni strutturali potenzialmente favorevoli, che richiedono, tuttavia, di essere attivate attraverso processi riflessivi e progettuali degli attori territoriali.

Con riferimento alla rete complessiva, si osserva la riorganizzazione del nucleo centrale. Alcune organizzazioni che già ricoprivano ruoli centrali nella rete territoriale mantengono la propria posizione, mentre altre rafforzano la propria centralità in relazione alle collaborazioni avviate dai progetti finanziati. L'esito dell'analisi della centralità, mostrato nella figura 3, consente di identificare tali attori in base alle loro posizioni nella rete complessiva.

Il Museo Civico Castello emerge come l'attore con i valori più elevati in tutti gli indici di centralità, con una *betweenness centrality*¹⁰ di 425 e una *closeness centrality*¹¹ di 142, configurandosi come hub territoriale capace di connettere trasversalmente ambiti culturali, turistici e istituzionali. Il Museo

⁹ È bene precisare che il disegno di ricerca non ha previsto una valutazione della sostenibilità di lungo periodo; questo studio ha avuto un respiro di medio periodo, poiché si è svolto quando i finanziamenti erano conclusi da tre/quattro anni.

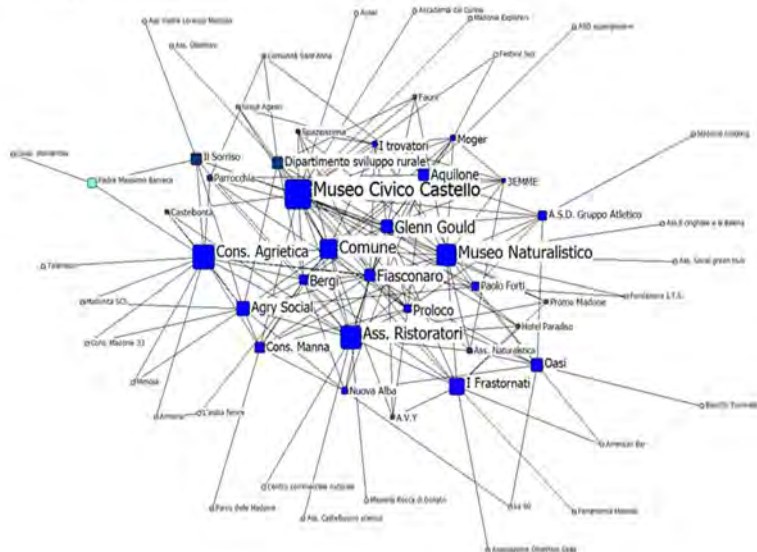
¹⁰ La *betweenness* è la frequenza con cui ogni singolo nodo si trova nel percorso più breve possibile, cioè quello geodetico, che collega ogni altra coppia di nodi della rete. Essa rappresenta la capacità che un attore ha di fungere da intermediario nelle relazioni tra gli attori della rete. Il valore è alto se siamo in presenza di uno snodo importante nella rete.

¹¹ La *closeness* è la somma delle distanze geodetiche che un attore ha dal resto dei nodi della rete. Tanto minore essa è, tanto più il nodo è vicino agli altri nodi del reticolo: valori bassi di *closeness* indicano una distanza ridotta di un attore dagli altri, e quindi maggiore prossimità di *ego* con gli *alter* della rete.

Naturalistico si colloca come secondo polo culturale (*betweenness* 229 e *closeness* 147), svolgendo una funzione complementare al Museo Civico nell'intermediazione culturale e scientifica.

L'Associazione Ristoratori, nonostante disponga di soli 15 legami diretti, registra una *betweenness* pari a 322, segnalando una forte capacità di intermediazione tra sottosistemi diversi, in particolare tra gli ambiti economico, turistico e culturale.

Fig. 3 - Rete complessiva: intermediari secondo la *betweenness* (dimensione del simbolo) e l'indice di degree (dimensione dell'etichetta)



Fonte: elaborazioni degli autori

Il Comune di Castelbuono, pur disponendo di numerose connessioni dirette, registra una *betweenness* contenuta pari a 74, a indicare una funzione di intermediazione meno marcata rispetto ai principali intermediari della rete. Fiasconaro (*betweenness* 113 e *closeness* 151) integra l'ambito imprenditoriale con legami che spaziano dall'ambito culturale a quello produttivo e turistico. Completano il quadro degli attori centrali la cooperativa Aquilone, che collega spazi associativi giovanili e sociali alla rete istituzionale, I Frastornati e l'Associazione Glenn Gould, ciascuno con specializzazioni settoriali complementari nell'ambito culturale. Il Consorzio Agrietica, pur avendo solo 6 legami diretti, registra una *betweenness* elevata, pari a 205. Questa posizione strutturale lo identifica come attore ponte tra il mondo agricolo-produttivo

(Agry Social e Nuova Alba) e gli hub turistico-culturali (Museo Civico, Fiasconaro, Associazione Ristoratori). Altri intermediari rilevanti includono l'Associazione Ristoratori (322), Fiasconaro (113), la cooperativa Aquilone (107) e il Consorzio Manna Madonita, che connettono ambiti settoriali differenti – culturale, economico, sociale e ambientale – contribuendo a ridurre la frammentazione complessiva della rete.

L'analisi reticolare ha reso visibili posizioni di centralità, ruoli di intermediazione e aree di marginalità. Questa base conoscitiva ha prodotto un dispositivo cognitivo in grado di alimentare processi di apprendimento collettivo. La presenza di nuovi hub settoriali solleva interrogativi sulla stabilità e sulla sostenibilità nel tempo di tali posizioni, sull'autonomia delle reti rispetto al sostegno finanziario esterno e sull'effettiva integrazione tra reti nuove e preesistenti. Questi interrogativi richiedono un'integrazione tra l'analisi delle configurazioni reticolari, condotta con la SNA, e l'analisi delle rappresentazioni collettive degli stakeholder del territorio, raccolte mediante la tecnica del gruppo nominale. L'analisi strutturale e l'argomentazione collettiva operano su piani complementari: la prima rende visibili le trasformazioni delle collaborazioni formalizzate; la seconda problematizza il significato e il valore attribuiti a tali trasformazioni, grazie ai giudizi valutativi espressi dai principali stakeholder del territorio, esplorando i significati, i conflitti e le ambivalenze. In questa prospettiva, la partecipazione costituisce una dimensione metodologica capace di integrare le risultanze di ricerca e di interpretazione, contribuendo alla costruzione di una conoscenza condivisa in uno spazio di riflessione dialogica sul significato delle trasformazioni avvenute.

3.2 I risultati del gruppo nominale

Nove referenti di progetto e i loro principali partner sono stati incaricati di valutare il raggiungimento degli obiettivi generali del programma. Gli autori di questo saggio riconoscono che la composizione del gruppo nominale può sollevare dubbi sul rischio di autoreferenzialità del giudizio valutativo. Si potrebbe inoltre pensare che un gruppo nominale omogeneo rischi di produrre valutazioni condivise, riducendo la propensione al confronto tra chi la pensa in modo difforme. Per questo, precisiamo che l'omogeneità interna al gruppo nominale (da cui, ad esempio, sono state escluse organizzazioni della comunità locale non coinvolte nella gestione dei progetti) è una decisione precisa del gruppo di ricerca: valorizzare il sapere esperto di chi ha implementato i progetti, coerentemente con un disegno di ricerca che mette al centro la valutazione partecipata e privilegia il sapere degli attori coinvolti nei

progetti, rispetto a una valutazione indipendente che invece avrebbe dato priorità alla credibilità esterna del giudizio. Inoltre, come illustrato nelle pagine successive, il gruppo nominale è stato particolarmente animato, al punto che i giudizi valutativi hanno generato dissensi consistenti in merito al conseguimento degli obiettivi strategico-general del programma. Dunque, si è rivelato infondato il timore che venisse meno un confronto schietto tra chi la pensava diversamente.

Operativamente, ciascun partecipante al gruppo nominale si è espresso individualmente in merito al raggiungimento degli obiettivi strategico-general del programma di sviluppo castelbuonese, utilizzando una scala a 5 livelli, in ordine crescente: l'obiettivo generale è fallito; l'obiettivo generale è stato quasi per nulla raggiunto; l'obiettivo generale è stato raggiunto solo in parte; l'obiettivo generale è stato raggiunto secondo le attese iniziali; l'obiettivo generale è stato pienamente raggiunto, anche oltre le attese iniziali. L'analisi aggregata delle risposte ha classificato gli obiettivi strategico-general secondo la logica del consenso valutativo, tenendo distinti: gli obiettivi sul cui conseguimento si è osservato un dissenso marcato (quando tra i partecipanti al gruppo nominale non si è costituita una maggioranza, ad esempio due gruppi di quattro o tre partecipanti ciascuno hanno espresso posizioni antitetiche); gli obiettivi rispetto ai quali è emerso un dissenso moderato (quando quattro partecipanti hanno espresso lo stesso giudizio valutativo e altri tre partecipanti hanno manifestato un giudizio difforme); gli obiettivi sul cui raggiungimento il gruppo si è espresso in modo quasi unanime (quando almeno cinque partecipanti hanno scelto la stessa risposta sulla scala valutativa).

Successivamente, i partecipanti hanno argomentato in gruppo, confrontandosi sui motivi del dissenso, marcato o moderato, in merito al raggiungimento degli obiettivi strategico-general. Così facendo, si è valutato in modo partecipato l'impatto del programma, dando priorità agli obiettivi su cui si è registrato dissenso all'interno del gruppo. Gli esiti della ricerca avvalorano questa strategia, grazie alle interessanti posizioni, talvolta molto distanti tra loro, che aiutano a far luce sui processi che hanno condizionato il raggiungimento di alcuni obiettivi.

La discussione ha seguito l'ordine decrescente di dissenso. Per ciascuno degli otto obiettivi su cui non si era raggiunto un consenso univoco, il ricercatore ha presentato le due posizioni prevalenti e ha dato la parola alternativamente ai sostenitori dell'una e dell'altra, vincolando ogni nuovo intervento all'apporto di argomentazioni inedite rispetto a quelle già emerse. Di seguito si presentano le argomentazioni relative a due obiettivi in particolare: il

rafforzamento della collaborazione fra le organizzazioni del territorio e gli enti pubblici locali, e quello tra le sole organizzazioni del territorio.

Riguardo al conseguimento dell'obiettivo di rafforzare la collaborazione tra le organizzazioni del territorio e gli enti pubblici locali, il gruppo ha espresso un dissenso marcato: una parte ha sostenuto che l'obiettivo non è stato per nulla raggiunto, l'altra che lo è stato solo parzialmente. Tra chi si è espresso in modo netto sul mancato raggiungimento dell'obiettivo, prevale la posizione dell'omesso sostegno da parte dell'amministrazione comunale all'attuazione dei progetti finanziati. L'amministrazione ha rivolto critiche forti, anche in occasioni pubbliche, al programma di sviluppo e alle organizzazioni destinatarie dei finanziamenti. C'è stata una politicizzazione degli interventi finanziati: i progetti sono stati presentati pubblicamente come l'esito di una matrice politico-ideologica distante dall'amministrazione in carica, portatori di valori lontani dalla loro idea di società e di comunità. Si segnala che nella fase di coprogettazione (prima dell'avvio dei progetti), l'amministrazione comunale aveva dato supporto all'intervento; sostegno che è venuto a mancare con il cambio di amministrazione, che si è distanziata dai progetti approvati dalla giunta politica precedente. La presenza del partner pubblico nei partenariati è stata dannosa, poiché in alcuni casi ha rallentato e ostacolato le attività del programma. La responsabilità andrebbe attribuita alla decisione di progettare un programma di sviluppo locale con il coinvolgimento dell'amministrazione comunale.

Tra chi ha sostenuto un parziale conseguimento dell'obiettivo di migliorare la collaborazione tra le organizzazioni del territorio e gli enti locali, si riportano alcune esperienze di collaborazione tra l'amministrazione locale e le organizzazioni coinvolte. In generale, l'argomentazione a favore di un parziale raggiungimento dell'obiettivo sostiene che la collaborazione tra enti pubblici locali e organizzazioni non profit è essenziale per un programma di sviluppo locale. Il coinvolgimento dell'amministrazione pubblica locale non è solo auspicabile, ma anche opportuno per accelerare le procedure amministrative (ad esempio l'ottenimento di un'autorizzazione necessaria allo svolgimento di un'attività progettuale). Nel caso di Castelbuono, ciò non è avvenuto a causa del deterioramento dei rapporti personali tra gli amministratori locali e i referenti di progetto, con effetti negativi sullo svolgimento delle attività e sul loro impatto sul territorio. Dunque, secondo questa posizione, la mancata collaborazione non era dovuta alla politicizzazione dei progetti stessi, bensì a cattivi rapporti tra le persone.

Per approfondire il mancato o parziale conseguimento di un obiettivo così delicato, che ha diviso i partecipanti al gruppo nominale, è stato intervistato il sindaco in carica, per comprendere il suo punto di vista sull'intero

programma di sviluppo locale¹². Si riscontra l'apprezzamento dell'amministrazione comunale per la bontà degli interventi, soprattutto in ambito culturale, che testimoniano l'importanza che Fondazione con il Sud riveste per il Mezzogiorno, in qualità di soggetto privato e come fondazione per lo sviluppo locale. Nonostante ciò, viene criticata aspramente la decisione di finanziare organizzazioni private per pagare gli stipendi dei lavoratori e avviare start-up con l'acquisto di mezzi, attrezzature e strumenti tecnici di vario tipo, che, una volta terminato il progetto, sono rimasti di proprietà delle organizzazioni private. Pur considerando nobili le finalità dei progetti, i finanziamenti – a detta del sindaco – andavano orientati all'acquisto di mezzi e attrezzature che, una volta conclusi i progetti, sarebbero dovuti diventare beni pubblici della comunità, utilizzabili gratuitamente da chiunque, e non beni privati nella disponibilità di pochi. Di contro, l'amministrazione pubblica rivendica il proprio ruolo di regolatore nei processi di sviluppo, mettendo al centro il valore del bene pubblico, fissando regole e stabilendo priorità su ciò di cui il territorio necessita, evitando di alimentare le richieste di pochi a scapito dell'interesse di tutti.

Sul rafforzamento della collaborazione tra le organizzazioni del territorio, la maggior parte dei partecipanti al gruppo nominale ha dichiarato che l'obiettivo è stato raggiunto secondo le attese iniziali; una parte, invece, ha sostenuto che è stato raggiunto solo parzialmente. A sostegno del pieno raggiungimento dell'obiettivo generale del programma, vi sono le argomentazioni di chi ritiene che il rafforzamento della collaborazione tra le organizzazioni dello stesso partenariato sia stato un elemento centrale per la buona riuscita del programma di sviluppo locale. Per raggiungere gli obiettivi specifici, il responsabile del partenariato ha rafforzato la collaborazione con i partner del progetto, dotati di capacità e competenze precise, complementari a quelle del capofila. Si è trattato soprattutto di collaborazioni preesistenti al partenariato, che si sono rafforzate per attuare le azioni progettuali. L'organizzazione capofila ha sostituito i partner esterni a Castelbuono, con cui collaborava in passato per lo svolgimento di alcune attività, con i partner interni al partenariato, nato nell'ambito del programma di sviluppo. La condivisione dell'obiettivo comune e le esperienze positive precedenti tra le organizzazioni sono state fattori determinanti che hanno portato i referenti dei progetti a sostenere che l'intervento ha consolidato la collaborazione tra le organizzazioni dello stesso partenariato, rafforzandole, soprattutto quando queste provenivano da esperienze passate di progetti già pianificati e implementati.

¹² L'amministrazione comunale non ha partecipato all'incontro del gruppo nominale, sebbene sia stata invitata.

assieme nel corso degli anni. Dunque, il programma di sviluppo di Castelbuono ha rafforzato e consolidato l'infrastrutturazione sociale grazie al potenziamento di organizzazioni che già operavano assieme. Meno incisivo è stato l'effetto dell'intervento quando il partenariato è stato costruito *ad hoc* per il progetto. Questa risultanza di ricerca solleva una questione interpretativa centrale: i cambiamenti osservati sono attribuibili al programma o riflettono dinamiche relazionali già radicate nel territorio? I partecipanti al gruppo nominale, che hanno giudicato positivamente il miglioramento della collaborazione tra le organizzazioni dello stesso partenariato, sono unanimi nel ritenere che i due fattori abbiano interagito, influenzandosi reciprocamente. Sul territorio esistevano già organizzazioni che collaboravano in modo soddisfacente, ma il rafforzamento di queste collaborazioni è dovuto all'intervento della Fondazione, poiché i finanziamenti erogati hanno permesso investimenti necessari al consolidamento proprio di tali collaborazioni, investendo in risorse materiali (come l'acquisto di mezzi e attrezzature tecniche) e immateriali (come la formazione necessaria al loro uso), grazie alle quali le organizzazioni si sono consolidate nella professionalizzazione dei loro operatori¹³.

Inoltre, la discussione ha prodotto argomentazioni a sostegno del raggiungimento parziale dell'obiettivo. Tra queste, è interessante riportare la posizione di chi sostiene che le organizzazioni finanziate avrebbero dovuto fare di più in termini di collaborazioni di secondo livello, cioè tra le organizzazioni inserite in partenariati diversi. Pur non essendo prevista né formalmente come obiettivo del programma né tra gli obiettivi dei singoli progetti, i referenti dei tre principali progetti avrebbero dovuto, a loro avviso, sviluppare una rete di comunità in cui far confluire le organizzazioni dei tre progetti del programma di sviluppo locale, per conseguire gli obiettivi generali in modo coordinato e strategico. Questa forma di meta-collaborazione tra partenariati avrebbe creato le condizioni per rispondere più efficacemente ai bisogni del territorio, affrontare con maggiore consapevolezza gli obiettivi strategici del programma e definirne di nuovi.

A detta di molti, la barriera principale è stata l'assenza di comunicazione e di coordinamento comune. Terminata la fase di coprogettazione, in cui la motivazione alla partecipazione è stata alta, i tre progetti hanno preso direzioni diverse, senza l'accompagnamento da parte di Fondazione con il Sud.

¹³ Sul punto, sarebbe interessante ricostruire la genesi sociale delle relazioni territoriali tra queste organizzazioni. Il ruolo svolto dalle collaborazioni pregresse e l'eventuale centralità già detenuta dagli attori che oggi occupano posizioni strategiche nella rete. Tale approfondimento non è presente in questo saggio ma sarà oggetto di un futuro sviluppo di questa stessa ricerca.

«Il programma di sviluppo locale messo su a Castelbuono non è diventato un progetto di comunità. Ciascun progetto è rimasto un progetto a sé stante, lavorando in compartimenti stagni. Questo è il limite del mancato raggiungimento degli obiettivi e dei risultati del progetto o del programma». Tra i partecipanti al gruppo nominale è diffusa l'idea che la responsabilità di non essere riusciti a collaborare tra loro sia delle organizzazioni stesse e non di Fondazione con il Sud, che presumibilmente avrebbe accolto con piacere una decisione così strategica. «Devo essere onesto: la Fondazione si aspettava da noi un salto di qualità».

4. Dalla valutazione partecipata all'apprendimento collettivo: osservazioni e prospettive valutative

L'integrazione tra l'analisi delle reti collaborative e le argomentazioni sollevate nel gruppo nominale consente di formulare alcune osservazioni valutative sul programma. I risultati della SNA hanno evidenziato un incremento della densità relazionale complessiva e una riduzione della frammentazione strutturale. Tuttavia, le risultanze del gruppo nominale hanno evidenziato una criticità sfuggita alla SNA: la segmentazione settoriale delle collaborazioni. Le reti attivate dai tre progetti si sono sviluppate come cluster distinti, ciascuno imperniato su specializzazioni tematiche e su nuclei organizzativi separati. Questa configurazione ha favorito l'intensificazione delle collaborazioni all'interno di ciascun partenariato, limitando lo sviluppo di sinergie trasversali.

La mancata realizzazione di una meta-collaborazione tra le organizzazioni coinvolte in progetti diversi è un'opportunità mancata. Durante il gruppo nominale, diversi stakeholder hanno sottolineato che l'assenza di occasioni di confronto tra i beneficiari dei tre progetti ha limitato le possibilità di condividere idee e strategie comuni. L'esperienza di Castelbuono suggerisce che l'efficacia degli interventi di sviluppo locale dipende sia dalla capacità di attivare singole iniziative sia dalla costruzione di un'infrastruttura relazionale capace di favorire processi di apprendimento collettivo e di integrazione tra ambiti tematici diversi (Picciotto e Wiesner, 2002; Trigilia, 2005).

Una prima lezione appresa riguarda l'opportunità di prevedere, nei futuri programmi di sviluppo locale, occasioni strutturate e frequenti di confronto tra beneficiari di progetti diversi, durante lo svolgimento delle attività e non solo a valle del programma. Tali occasioni potrebbero costituire piattaforme di scambio di buone pratiche o tavoli tematici trasversali, finalizzati a stimolare la riflessione collettiva su obiettivi integrati. La creazione di questi spazi

potrebbe contribuire a un rafforzamento sostanziale della capacità collaborativa del territorio, orientata a logiche di sviluppo sistemiche e multisettoriali.

Un secondo nodo critico riguarda il ruolo del Comune. Il reticolo mostra che l'amministrazione comunale è fortemente connessa tramite relazioni dirette, senza tuttavia assumere un ruolo altrettanto rilevante nei percorsi di connessione indiretta. Il Comune è formalmente coinvolto in numerosi partenariati, ma svolge un ruolo di coordinamento meno marcato rispetto ad altri attori centrali, come il Museo Civico Castello o il Consorzio Agrietica. Appare più come un nodo di passaggio – in virtù delle sue funzioni amministrative – che come promotore di sinergie collaborative. Al riguardo, gli stakeholder hanno espresso argomentazioni contrastanti: alcuni hanno sostenuto che il Comune è stato poco coinvolto nei processi decisionali e operativi dei progetti; altri hanno rilevato che la sua presenza si è tradotta in forme di condizionamento che hanno rallentato le attività e introdotto logiche di appartenenza estranee alla natura del programma. Questa duplice critica – troppo poco coinvolto per alcuni, troppo presente per altri – segnala una tensione nota nella letteratura sullo sviluppo locale (Trigilia, 2005; Bagnasco, 2006): la necessità di coinvolgere l'amministrazione pubblica, in quanto detentrica di funzioni regolative, a fronte del rischio che tale coinvolgimento si traduca in logiche di controllo politico o in una burocratizzazione eccessiva. Una seconda lezione appresa riguarda l'opportunità di ripensare le modalità di coinvolgimento dell'attore pubblico. L'esperienza suggerisce che la presenza del Comune non dovrebbe essere imposta come requisito formale, ma potrebbe essere subordinata all'assunzione di impegni istituzionali espliciti, come protocolli d'intesa che ne stabiliscano il ruolo e le responsabilità operative, e forme di cofinanziamento che ne attestino un impegno sostanziale. In assenza di tali condizioni, le organizzazioni del privato sociale potrebbero essere lasciate libere di definire autonomamente le proprie alleanze progettuali, senza che la presenza pubblica sia considerata un requisito formale. In linea con l'approccio della governance collaborativa (Ansell e Gash, 2008), il coinvolgimento del pubblico potrebbe essere orientato al rafforzamento della capacità collettiva di azione, anche del privato sociale. Una terza osservazione critica riguarda il ruolo della valutazione a posteriori, che ha prodotto apprendimenti rilevanti sia per la Fondazione sia per gli stakeholder locali, rendendo visibili trasformazioni strutturali, ruoli di intermediazione e criticità organizzative che erano rimasti in larga misura impliciti. Tuttavia, la valutazione a posteriori ha limitato la possibilità di tradurre tali apprendimenti in azioni correttive durante lo svolgimento dei progetti. Un accompagnamento valutativo avrebbe potuto offrire occasioni di riflessività collettiva, consentendo di identificare tempestivamente le criticità, di correggere le strategie inefficaci e di rafforzare le capacità organizzative

attraverso processi di apprendimento adattivo. Questa osservazione richiama la valutazione orientata al *capacity building* (UNDP, 2008; Patton, 2011), utile a rafforzare le competenze tecniche, gestionali e relazionali degli attori coinvolti, e suggerisce che l'accompagnamento valutativo si configura come uno spazio istituzionalizzato di riflessione collettiva, in cui le scelte strategiche e le criticità possono essere discusse e orientate in modo condiviso. Nel caso di Castelbuono, l'accompagnamento valutativo avrebbe potuto assumere molteplici forme: monitoraggio partecipato con momenti periodici di riflessione collettiva; analisi delle reti collaborative condotte in fasi intermedie per individuare dinamiche emergenti; laboratori di *capacity building* orientati al rafforzamento delle competenze progettuali, gestionali e collaborative. Tali attività non sostituiscono la valutazione finale, ma la integrano, producendo una base informativa più ricca e un processo di apprendimento distribuito nel tempo. Una terza lezione appresa riguarda l'opportunità di prevedere occasioni intermedie di riflessione valutativa come componente strutturale dei futuri programmi di sviluppo locale. Da un lato, l'esperienza suggerisce di concepire la valutazione come un investimento nel rafforzamento delle capacità degli attori locali. Questo implica la previsione di risorse dedicate, l'identificazione di spazi valutativi intermedi e la costruzione di relazioni collaborative tra valutatori e beneficiari, orientate non alla verifica formale, bensì al supporto sostanziale. Dall'altro lato, i beneficiari potrebbero essere coinvolti attivamente nei processi valutativi in itinere, non come oggetti passivi ma come coproduttori di conoscenza riflessiva sulle proprie pratiche. L'adozione di un approccio valutativo formativo e partecipato rafforza la capacità del territorio di apprendere dall'esperienza, di riconoscere e valorizzare i propri punti di forza e di affrontare tempestivamente le criticità. In questo senso, la valutazione si configura come uno strumento di empowerment organizzativo e territoriale, coerente con la logica generativa che orienta l'approccio di Fondazione allo sviluppo locale.

Dichiarazioni

Gli autori dichiarano che il Software Claude (versione Sonnet 5) 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

Amaturo, E., Punziano, G. (2016). *I "mixed methods" nella ricerca sociale*. Roma: Carocci.

- Ansell, C., Gash, A. (2008). *Collaborative Governance in Theory and Practice*. Journal of Public Administration Research and Theory, 18(4): 543-571. DOI: 10.1093/jopart/mum032
- Bagnasco, A. (2006). *Tracce di comunità*. Bologna: Il Mulino.
- Banca d'Italia (2022). *Il divario Nord-Sud: sviluppo economico e intervento pubblico*. Collana "Seminari e Convegni", n. 25, Roma.
- Bertin, G. (2011). *Con-sensus method. Ricerca sociale e costruzione di senso*. Milano: FrancoAngeli.
- Bezzi, C. (2013). *Fare ricerca con i gruppi*. Milano: FrancoAngeli.
- Bezzi, C. (2021). *Manuale di ricerca valutativa*. Milano: FrancoAngeli.
- Cartocci, R. (2007). *Mappe del tesoro. Atlante del capitale sociale in Italia*. Bologna: Il Mulino.
- Chen, H.T. (1990). *Theory-Driven Evaluations*. Newbury Park (CA): Sage Publications.
- Cousins, J.B., Whitmore, E. (1998). *Framing Participatory Evaluation*. New Directions for Evaluation, 80: 5-23. DOI: 10.1002/ev.1114.
- Denzin, N.K. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods*. New York: McGraw-Hill.
- Donati, P., Colozzi, I. (2006). *Terzo Settore e valorizzazione del capitale sociale in Italia: Luoghi e attori*. Milano: FrancoAngeli.
- Foresti, M. (2003). *La partecipazione in ambito valutativo: mito o realtà? Teoria e pratica degli approcci partecipati in valutazione*. Rassegna Italiana di Valutazione, 25: 69-93.
- Greene, J.C. (2007). *Mixed methods in social inquiry*. San Francisco: Jossey-Bass/Wiley.
- Johnson, R.B., Onwuegbuzie A.J., Turner L.A. (2007). *Toward a Definition of Mixed Methods Research*. Journal of Mixed Methods Research, 1(2): 112-133. DOI: 10.1177/1558689806298224
- Lo Presti, V. (2008). *Valutare i programmi di sviluppo locale*. Milano: FrancoAngeli.
- Lo Presti, V., Dentale M., Celardi E. (2018). *Ricerca valutativa e Case Study Research: un caso "positivo" di valutazione dell'infrastrutturazione sociale a Palermo*. Rassegna Italiana di Valutazione, a. XXII, n. 71: 9-30. DOI: 10.3280/RIV2018-071002
- Long, J.C., Hibbert, P., Braithwaite, J. (2016). *Structuring successful collaboration: a longitudinal social network analysis of a translational research network*. Implementation Science, 11: 19. DOI: 10.1186/s13012-016-0381-y.
- Martire, F., Parra Saiani, P., Cataldi, S. (2023). *La ricerca sociale e le sue pratiche*. Roma: Carocci.
- Palumbo, M. (2001). *Il processo di valutazione: decidere, programmare, valutare*. Milano: FrancoAngeli.
- Palumbo, M. (2003). *La valutazione partecipata e i suoi esiti*. Rassegna Italiana di Valutazione, 25: 71-88.
- Palumbo, M. (2011). *La partecipazione che cambia la valutazione (e forse anche i valutatori?)*. Rivista Trimestrale di Scienza dell'amministrazione, (2010/4). DOI: 10.3280/SA2010-004003
- Parra Saiani, P. (2004). *Triangolazione e privato sociale. Strategie per la ricerca valutativa*. Acireale-Roma: Bonanno.
- Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.
- Picciotto, R., Wiesner, E. (a cura di) (2002). *Evaluation and development: The institutional dimension*. New Brunswick, NJ: Transaction Publishers.
- Putnam, R.D. (1993). *Making Democracy Work. Civic Traditions in Modern Italy*. Princeton (NJ): Princeton University Press.
- Stame, N. (2004). *Theory-Based Evaluation and Types of Complexity*. Evaluation: 10(1), 58-76.
- Stame, N. (2016). *Valutazione pluralista*. Milano: FrancoAngeli.

SVIMEZ (2024). *Rapporto SVIMEZ 2024 - L'economia e la società del Mezzogiorno*. Associazione per lo Sviluppo dell'Industria nel Mezzogiorno, Roma.

Tomei, G. (2004). *Valutazione partecipata della qualità. Il cittadino-utente nel giudizio sugli interventi di politica e servizio sociale*. Milano: FrancoAngeli.

Triglia, C. (2005). *Sviluppo locale. Un progetto per l'Italia*. Roma-Bari: Laterza.

UNDP (2008). *Capacity Development Practice Note*. New York: United Nations Development Programme.

Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, L. (1966). *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: The University of Chicago Press.

Weiss, C.H. (1995). *Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families*. In J.P. Connell et al. (Eds.), *New Approaches to Evaluating Community Initiatives*. Washington DC: Aspen Institute.

Partecipazione e valutazione nelle politiche universitarie di Public Engagement: verso modelli dialogici e generativi

Emanuela Pascuzzi, Stefania Chimenti***

This paper analyses an experience of participatory evaluation during the initial phase of an Italian university's Public Engagement (PE) policy. It integrates social research, evaluation, and co-design. The aim is to understand how PE is reflected in the practices and perceptions of the academic community, and to build a shared vision for its future development.

The case study demonstrates how participation can represent an epistemological and political dimension of evaluation, generating knowledge in pluralistic manner, fostering a shared narrative of the university's public role, transforming internal relationships, and laying the foundations for a more structured, inclusive, and socially impactful PE policy.

Keywords: participatory action research; academic community; third mission; appreciative inquiry; participatory evaluation.

Introduzione

Negli ultimi decenni, la partecipazione è divenuta una parola chiave nel lessico delle politiche pubbliche, come anche in quello della valutazione (Forrest, 2003; Gregory, 2000; Riquier, 1997; Rebien, 1996; Guba e Lincoln, 1989). Pur se talvolta intesa riduttivamente come un insieme di tecniche per coinvolgere i portatori di interesse, gestire conflitti o legittimare decisioni già prese, sono presenti posizioni che concepiscono la partecipazione come una dimensione epistemologica e politica dei processi decisionali e valutativi, riconoscendo che la conoscenza utile per valutare e decidere è

* Emanuela PASCUZZI, Università della Calabria (emanuela.pascuzzi@unical.it)

** Stefania CHIMENTI, Università della Calabria (stefania.chimenti@unical.it)

Invio Proposta: 29-04-2026. Accettazione: 22-07-2026.

distribuita, situata, plurale, e che la qualità democratica delle politiche dipende anche da come questa conoscenza viene mobilitata, discussa e messa in comune (Valastro, 2010; Palumbo e Torrigiani, 2009, 2013).

Il contributo analizza un'esperienza che si colloca nella fase istitutiva della politica di Public Engagement di un'università del Sud Italia, integrando ricerca sociale, valutazione partecipativa e co-ideazione di policy. Nello specifico, viene analizzato il percorso partecipativo di ricerca e sviluppo del Public Engagement (PE) dell'Università della Calabria, progettato per rispondere a una duplice esigenza: da un lato, comprendere in che modo il PE fosse presente, in forma più o meno visibile, nelle pratiche e nelle rappresentazioni della comunità accademica; dall'altro, costruire una visione condivisa del suo sviluppo futuro, in un'ottica di missione sociale dell'università.

L'analisi del caso risponde a una domanda di ricerca centrale: in che modo un processo di valutazione partecipativa può contribuire a definire una politica universitaria di Public Engagement nella sua fase istitutiva? La rilevanza della domanda deriva dal carattere emergente e ancora instabile del PE nel contesto italiano, che richiede approcci capaci di far emergere conoscenze diffuse, orientare scelte organizzative e sostenere la costruzione di una visione condivisa sul ruolo pubblico dell'università.

L'articolo si propone di mostrare come la partecipazione possa assumere un ruolo costitutivo nella produzione di conoscenza valutativa e nella definizione delle priorità di policy.

Per farlo, dopo aver delineato il quadro teorico di riferimento (par. 1), si descrivono il contesto istituzionale in cui nasce il percorso (par. 2) e il disegno metodologico adottato (par. 3), si analizzano i principali risultati e le trasformazioni generate (par. 4), discutendo, infine, le implicazioni dell'esperienza per la valutazione partecipativa nel policy making universitario (par. 5) e avanzando alcune considerazioni conclusive.

1. Università e missione pubblica in trasformazione

Le università europee e italiane sono al centro di un nuovo paradigma che assegna loro un ruolo che oltrepassa le tradizionali missioni di formazione e ricerca, chiamandole a divenire attori pubblici che contribuiscono in modo diretto allo sviluppo sociale, culturale ed economico dei territori in cui operano (Compagnucci e Spigarelli, 2020; Meyer e Sporn, 2019; Zomer e Benneworth, 2011). La cosiddetta Terza Missione (TM) ha reso esplicita questa dimensione, riconoscendo il valore delle attività di valorizzazione della

ricerca, prima nella dimensione del Trasferimento Tecnologico e, più recentemente, in quella del Public Engagement (Arnaldi, 2020; Boffo, Moscati e Rostan, 2024).

La crescente richiesta di “conoscenza utile” ha portato a una maggiore attenzione verso la valutazione pubblica della ricerca e della TM, focalizzandosi tanto sulla produzione di sapere, quanto sulla dimostrazione del valore economico e sociale delle conoscenze¹ (Blasi, 2023). In Italia, i sistemi di valutazione periodica delle università (AVA, VQR) hanno assunto un ruolo sempre più rilevante nell’allocazione delle risorse e nella definizione delle strategie di ateneo (Blasi, 2023; Fasanella e Martire, 2020). Come mostrano alcuni studi (Balduzzi e Vaira, 2018; Boffo, Moscati e Rostan, 2024), tali sistemi possono innescare processi di isomorfismo istituzionale (Di Maggio e Powell, 1983), spingendo le università ad adeguarsi a modelli organizzativi e comportamenti ritenuti legittimi dal centro politico-amministrativo. In questo contesto, il rischio è che le attività accademiche, nelle diverse forme ed espressioni, vengano interpretate prevalentemente come adempimenti da rendicontare, piuttosto che come spazi di sperimentazione democratica e di co-produzione di valore pubblico. Le politiche di valutazione, se concepite in modo strettamente tecnocratico, inoltre, possono contribuire a verticalizzare i processi decisionali, riducendo gli spazi di partecipazione interna e di confronto con la società (Regonini, 2001).

La letteratura sulla valutazione partecipativa ha messo in luce come il giudizio valutativo sia un processo sociale che implica scelte di valore, interpretazioni, negoziazioni (Scardigno *et al.*, 2019; Palumbo, 2001; Guba e Lincoln, 1989, 1982). In questa prospettiva, la partecipazione rappresenta una modalità di produzione e valorizzazione della conoscenza che riconosce la pluralità dei punti di vista e la natura situata dei saperi, non riducendosi, dunque, a una semplice tecnica di coinvolgimento degli stakeholder (Greenwood e Levin, 1998). Approcci come la *developmental evaluation* (Patton, 1994) o le pratiche di co-valutazione (Torrighiani, 2025; De Ambrogio e Guidetti, 2023) insistono sull’importanza di includere i soggetti coinvolti nelle politiche nei processi di definizione delle domande valutative, di raccolta e interpretazione dei dati, di costruzione dei giudizi. La partecipazione assume, così, una dimensione epistemologica, influenzando su cosa viene considerato

¹ Nel contesto della valutazione universitaria italiana, a partire dal Bando VQR 2020-2024 dell’ANVUR, l’espressione “Terza missione” viene sostituita o integrata da quella di “Valorizzazione delle conoscenze”. La nuova terminologia ha inteso ridefinire il concetto superando l’idea di attività marginali o limitate al Trasferimento Tecnologico, per includere un impatto più ampio delle università sulla società.

“rilevante”, su quali evidenze vengono raccolte, su come queste vengono interpretate e su quali azioni ne derivano.

Nel caso delle università, la partecipazione nei processi valutativi può contribuire a rendere visibili pratiche e significati che sfuggono a indicatori standardizzati, nonché a rafforzare il senso di appartenenza e di responsabilità condivisa rispetto alla missione pubblica.

La necessità di rafforzare una responsabilità condivisa è tanto più urgente se si considera la natura organizzativa degli atenei. Le università sono istituzioni *loose coupling* (Fonseca e Nieth, 2021; Weick, 1982) con strutture interne complesse e frammentate (Goddard e Vallance, 2014). Le strutture decentrate, come i dipartimenti, rischiano di restare ai margini del processo che alimenta la definizione e la valutazione delle politiche di ateneo e di ripiegare la propria azione entro uno spazio eccessivamente autoreferenziale.

Il progressivo orientamento verso la valutazione partecipativa nel policy making universitario italiano si colloca all'intersezione tra l'evoluzione del quadro normativo nazionale sulla performance pubblica e l'affermazione, nella letteratura internazionale, di modelli di governance improntati alla co-produzione del valore pubblico. Le Linee guida del Dipartimento della Funzione Pubblica (DFP, 2019), elaborate in attuazione del d.lgs. 150/2009 e del d.lgs. 74/2017, formalizzano il principio secondo cui gli stakeholder devono essere coinvolti in modo strutturato nei processi di definizione, monitoraggio e valutazione delle politiche pubbliche. Parallelamente, le analisi sull'*open government* (OECD, 2025; 2023) mostrano come la partecipazione informata e deliberativa contribuisca a migliorare la qualità delle decisioni, a rafforzare la legittimazione istituzionale e a generare fiducia nei confronti delle organizzazioni pubbliche. In tale prospettiva, le università - istituzioni complesse, multilivello e ad alta intensità di conoscenza - sono chiamate a integrare nei propri processi decisionali il contributo di studenti, personale, comunità scientifiche e attori territoriali, riconoscendoli come co-valutatori e co-produttori delle politiche accademiche, nel rispetto del cosiddetto requisito di inclusione, suggerito da House e Howe (2000). La valutazione partecipativa assume così una funzione strategica nel ripensare la governance universitaria in chiave di accountability, trasparenza e orientamento al valore pubblico, soprattutto in relazione alle sfide contemporanee che investono la didattica, la ricerca e la terza missione.

Il Public Engagement si presta particolarmente a essere letto in chiave partecipativa. Nel contesto italiano – e nel presente contributo – esso è inteso come componente della Terza Missione orientata al valore sociale, educativo

e culturale, in linea con l'inquadramento formale di ANVUR² (Boffo, Moscati e Rostan, 2024; ANVUR, 2018). Si tratta, tuttavia, di una categoria concettualmente fluida e in evoluzione (Falbo *et al.*, 2025), i cui confini dipendono dai sistemi di governance, dalle culture istituzionali e dai dispositivi di valutazione. Nelle interpretazioni più ampie, coincide con l'impegno civico complessivo dell'università (Goddard e Vallance, 2014; Watson *et al.*, 2011), mentre altri approcci lo descrivono come orientamento collaborativo e co-produttivo che attraversa ricerca, didattica e servizio pubblico (Mazzonetto, 2025; Beere, Votruba e Wells, 2011; Furco, 2010).

Questa natura plurale e non rigidamente definita genera implicazioni valutative rilevanti, che richiedono approcci capaci di cogliere processi, relazioni e significati, come evidenziato dagli studi sulla valutazione partecipativa (Palumbo, 2001; Guba e Lincoln, 1989).

2. Il contesto universitario e la fase istituyente del Public Engagement

Nel caso dell'Università della Calabria, la strutturazione della TM ha origine nei primi anni Duemila, con un orientamento iniziale prevalentemente rivolto al Trasferimento Tecnologico. Nel 2003 vengono istituiti il Liaison Office e il TechNest, segnando l'avvio di una politica di valorizzazione della ricerca applicata e del dialogo con il sistema produttivo. Per diversi anni, questa dimensione rimane centrale, mentre le attività di PE sul versante culturale o sociale, pur presenti in forma diffusa, non trovano un riconoscimento formale né un coordinamento unitario.

Una svolta significativa si verifica con l'insediamento della nuova governance nel 2019, che introduce per la prima volta la "Missione sociale" come componente strategica della TM. Questo cambiamento viene formalizzato attraverso l'assegnazione di una delega rettorale specifica e l'inserimento del PE nel Piano Strategico 2020-2022, accanto al Trasferimento Tecnologico. Il PE viene così riconosciuto come ambito di azione rilevante, ma la sua traduzione operativa è ancora tutta da costruire.

All'avvio del percorso, il quadro che emerge è quello di un campo frammentato e poco visibile: le iniziative di PE sono numerose ma episodiche, spesso legate all'impegno individuale di singoli docenti o gruppi di ricerca; solo alcuni dipartimenti hanno nominato referenti dedicati; manca un

² Va detto, tuttavia, che a fronte di una definizione iniziale ampia e aperta, ANVUR ha progressivamente ristretto il perimetro del PE, circoscrivendone i campi di azione alla divulgazione scientifica e alle attività di interazione con le scuole (Carrera *et al.* 2025).

linguaggio condiviso per descrivere il PE, e molti attori faticano a riconoscere come tali attività, che pure svolgono da anni (laboratori con le scuole, progetti con associazioni, iniziative culturali, ecc.), ricadano nel campo specifico della Missione sociale.

Sul piano organizzativo, non esiste un sistema strutturato di monitoraggio e valutazione delle attività di PE, le informazioni sono disperse, non comparabili e difficilmente utilizzabili per fini strategici. Sul piano culturale, la consapevolezza del PE come componente della missione istituzionale è ancora debole e a prevalere è la percezione che si tratti di attività aggiuntive e facoltative rispetto a didattica e ricerca.

In questo scenario, l'ateneo esclude una definizione top-down della politica di PE e sceglie di attivare un percorso partecipativo di ricerca e sviluppo. L'idea di fondo è che, per rendere il PE una componente strutturale e condivisa della missione universitaria, sia necessario partire dall'ascolto delle esperienze esistenti, dalle rappresentazioni della comunità accademica e dalle sue aspirazioni. Si apre, così, un percorso che ha una duplice natura: a) valutativa, in quanto mira a comprendere lo stato dell'arte del PE, le sue potenzialità e criticità, le condizioni che ne favoriscono o ostacolano lo sviluppo; b) progettuale, in quanto orientato a co-ideare azioni di sistema e proposte per rafforzare la collaborazione tra università e società, definendo priorità, strumenti e criteri di azione.

Per sostenere questo percorso, viene istituito un sistema di leadership diffusa articolato su due livelli: il Gruppo di Regia e Supporto (GRES), con funzioni di coordinamento centrale, e il Gruppo di Azione per il PE (GrAPE), composto da rappresentanti dei 14 dipartimenti e di alcuni centri, come ad esempio il Sistema Museale di ateneo, includendo sia docenti-ricercatori sia personale tecnico-amministrativo. Il GrAPE diventa il luogo privilegiato di confronto, restituzione e co-ideazione.

3. Il disegno metodologico

Il percorso è progettato come ricerca valutativa partecipativa, orientata a far emergere le rappresentazioni, le pratiche e le aspettative della comunità accademica rispetto al Public Engagement (PE) in una prospettiva costruttivista (Greenwood e Levin, 1998; Guba e Lincoln, 1989). La scelta di una metodologia partecipativa risponde alla natura fluida del PE e al suo carattere che è al contempo radicato nell'esperienza degli accademici ma emergente come missione istituzionale dell'ateneo, suggerendo che siano gli attori universitari ad esprimerne significati, priorità e condizioni di sviluppo. La

strategia adottata si ispira alla *community participative action research* (Israel *et al.*, 2012; Reason e Bradbury, 2008). L'idea è che la produzione di conoscenza sul PE non sia separabile dall'attivazione di processi di cambiamento poiché la ricerca è già intervento e la valutazione è già trasformazione (Guba e Lincoln, 1989).

La strategia di ricerca-azione partecipativa è costruita intorno a due principali metodi: il Confronto Creativo (CC) e l'Appreciative Inquiry (AI). Il Confronto Creativo (Sclavi, 2021; Sclavi e Susskind, 2011) viene utilizzato come dispositivo dialogico, in particolare nel lavoro del GrAPE. Esso si fonda su due pilastri - ascolto e creatività - mettendo in atto pratiche di ascolto attivo, sospensione del giudizio, esplorazione di mondi possibili e moltiplicazione delle opzioni prima di giungere all'adozione di decisioni co-progettate e condivise (Fig. 1). Nel contesto Unical, il Confronto Creativo permette, innanzitutto, di mettere in dialogo prospettive disciplinari e ruoli diversi (docenti di aree scientifiche differenti, personale tecnico-amministrativo, governance) e di affrontare tensioni e resistenze rispetto al PE, trasformandole in occasioni di apprendimento. In maniera decisiva, lo spazio di dialogo mediato dall'approccio del Confronto Creativo consente di co-costruire un linguaggio condiviso per parlare di PE, superando incomprensioni e stereotipi.

L'adesione alle "regole del gioco" del CC (Fig. 2) - tutti esploratori di mondi possibili, nessuna proposta esclusa a priori, attenzione alle condizioni di implementabilità - contribuisce, inoltre, a creare un clima di fiducia e di responsabilità condivisa.

Fig. 1 - Come navigare nella complessità attraverso il CC

Per avviare il processo è fondamentale costruire un clima di fiducia personale tra tutti i partecipanti. Lo strumento decisivo è l'Ascolto Attivo, inteso come occasione per comprendere più a fondo le preoccupazioni reciproche e la diversità dei punti di vista, senza concentrarsi sul chi abbia ragione o torto. Nella fase dedicata alla generazione di nuove opzioni emergono due principali difficoltà. 1) Il compito può sembrare irrealizzabile: nessuna delle proposte iniziali appare adeguata e non si intravedono alternative. In questo caso è necessario cambiare prospettiva, sospendendo la discussione su ostacoli e vincoli e lavorando insieme alla costruzione di una visione condivisa di futuro desiderabile (di solito già abbozzata nella sintesi della fase precedente), che diventa la bussola per individuare idee e pratiche accettabili per tutti. 2) Può subentrare l'ansia di arrivare rapidamente a una conclusione, con il timore che ampliare le opzioni faccia perdere tempo. Qui possono aiutare un tono leggero, un po' di spirito e il ricordo di casi analoghi in cui, proprio quando sembrava non esserci via d'uscita, è emersa una soluzione inattesa. La presenza e l'accompagnamento di un/a facilitatore/trice garantisce la tenuta del processo.

Fonte: Sclavi, 2021.

Fig. 2 - Le regole base del Confronto Creativo in Unical

- Un processo partecipativo orientato dal CC è aperto a tutti gli interessati ma richiede che ogni partecipante sappia come funziona, voglia farne parte e accetti le regole di base.
- Quando un partecipante nomina un proprio sostituto al lavoro del gruppo, si impegna a tenere aggiornato quest'ultimo tanto nel merito quanto nel metodo della discussione.
- Il faccia a faccia del piccolo gruppo aiuta a dare concretezza e pragmaticità alla discussione.
- Ogni singolo partecipante può prendere la parola sia in plenaria che nel piccolo gruppo, dicendo brevemente il suo punto di vista. Non dilungarsi, così da dare spazio ad altri.
- I diversi punti di vista e le divergenze sono considerati importanti contributi per aumentare la conoscenza. Ognuno e ogni piccolo gruppo è consapevole che il tutto è più della somma delle parti e che il suo è un contributo che deve essere letto alla luce di tutti gli altri.
- È fondamentale fare emergere tutte le posizioni e formulare domande non per criticarle, ma per capire le ragioni che le sottendono (Ascolto Attivo).
- Ognuno si fa carico di cercare quali interessi e preoccupazioni comuni sono alla base delle posizioni (anche divergenti) e si impegna nell'esplorazione di possibili soluzioni di mutuo gradimento (moltiplicazione delle opzioni, buone pratiche).
- Separare il momento dell'invenzione (moltiplicazione delle opzioni) da quello della decisione ("Proibito adottare una soluzione prima di aver moltiplicato collettivamente le opzioni!").
- Molti processi di CC comportano decisioni che, almeno in parte, fanno affidamento su nozioni tecniche e specialistiche di qualche tipo. In questo caso, occorre elencare con chiarezza le questioni tecniche su cui si ritiene necessaria una consulenza, individuare una gamma di esperti ai quali il gruppo può ricorrere, presentare e discutere con questi ultimi anche le implicazioni politiche delle soluzioni da essi proposte.
- Gli approfondimenti tecnici possono essere più agevolmente condotti attraverso il lavoro in sottogruppi. I risultati rappresentano un punto di partenza del confronto in plenaria e sono utili a facilitare la decisione nel gruppo più ampio.
- La decisione è favorita dall'elaborazione di proposte ponte e pacchetti di proposte, che cercano di includere al meglio le esigenze di tutti e assume la forma di un testo unico, anche provvisorio e rivedibile, che riassume i diversi punti di vista e le soluzioni di mutuo gradimento individuate.

Fonte: Nostro adattamento da "Prontuario delle regole di base del Confronto Creativo" in Sclavi e Susskind, 2011.

L'Appreciative Inquiry viene adottata come approccio centrale del percorso. Nata in ambito organizzativo (Cooperrider e Whitney, 2005; Cooperrider, Whitney e Stavros, 2003; Whitney e Trosten-Bloom, 2003; Cooperrider e Srivastva, 1987), l'Appreciative Inquiry si basa sull'assunto che ogni organizzazione possieda un "nucleo positivo" fatto di valori, esperienze e risorse che, se riconosciuti e valorizzati, possono attivare processi generativi di cambiamento.

Il modello delle 4-D (*Discovery, Dream, Design, Destiny/Delivery*) struttura il percorso in quattro fasi integrate (Fig. 3). La fase di *Discovery* riguarda l'esplorazione delle esperienze positive già esistenti, che costituiscono la base conoscitiva da cui partire. Su queste fondamenta si innesta la fase di *Dream*, dedicata all'immaginazione collettiva di un futuro desiderato, in cui i partecipanti costruiscono visioni condivise e aspirazioni comuni. A partire da tali visioni prende forma la fase di *Design*, in cui si co-ideano soluzioni e strategie capaci di tradurre le aspirazioni in azioni concrete. Infine, la fase di *Destiny/Delivery* riguarda l'attuazione dei cambiamenti che consolidano il percorso e aprono la strada a nuove possibilità di apprendimento e sviluppo.

L'efficacia dell'approccio è sostenuta da principi fondanti derivati dal costruzionismo sociale, che sottolineano come la realtà sia costruita attraverso l'interazione sociale e il linguaggio, come l'immaginario futuro influenzi le scelte del presente e come ogni ricerca sia già intervento (Fig. 4).

Fig. 3 - Modello delle 4-D dell'Appreciative Inquiry

Discovery Esperienze positive	Dream Futuro desiderato
Design Soluzioni condivise	Destiny Cambiamenti attuati o prospettati

Fonte: Nostro adattamento da Cooperrider, Whitney e Stavros, 2003.

Fig. 4 - Principi dell'Appreciative Inquiry

Principio	Descrizione
Costruzionista	Un'organizzazione è una costruzione sociale. <i>Words create worlds!</i>
Di simultaneità	Ricerca e intervento procedono di pari passo. La ricerca è cambiamento!
Poetico	La storia di un'organizzazione è un racconto in fieri che ciascuno contribuisce ad arricchire
Di anticipazione	L'immagine del futuro guida il comportamento nel presente. L'immaginario ispira l'azione
Positivo	Le organizzazioni evolvono nella direzione delle cose che studiano. Domande positive creano un cambiamento positivo

Fonte: Nostro adattamento da Cooperrider, Whitney e Stavros, 2003.

Nel caso presentato, l'Appreciative Inquiry viene utilizzata anche in chiave valutativa (Coghlan, Preskill e Catsambas, 2003; Lo Presti, 2009). Le interviste raccolgono contenuti descrittivi, invitando i partecipanti a riflettere

in profondità sulle esperienze di PE che considerano più significative e sulle ragioni che le rendono tali. Allo stesso tempo, li sollecitano a individuare i fattori che ne hanno favorito il successo, gli ostacoli incontrati lungo il percorso e le modalità con cui sono riusciti ad affrontarli. L'esplorazione si estende poi alle aspirazioni per il futuro dell'ateneo, chiedendo ai partecipanti quale forma di PE desidererebbero vedere sviluppata e quali condizioni ritengano necessarie per rendere possibile tale visione. In questo modo, la valutazione è orientata non tanto a individuare "mancanze" o "fallimenti", quanto a "localizzare l'energia per il cambiamento" (Elliott, 1999), partendo dai punti di forza per affrontare anche le criticità.

La letteratura ha evidenziato alcuni rischi dell'AI, come la possibile riproduzione dello status quo o l'eccessiva enfasi sul positivo (Patton, 2002). Nel percorso Unical, si è cercato di bilanciare questa dimensione dando spazio anche all'espressione delle criticità, ma scegliendo di partire dalle esperienze positive per costruire un terreno generativo comune.

Per quanto riguarda il disegno empirico del percorso partecipativo, tra il 2021 e il 2022 vengono condotte 93 interviste in profondità, utilizzando un campionamento a valanga (Di Franco, 2010). Il percorso inizia coinvolgendo i direttori di dipartimento e i delegati alla TM e/o al PE delle strutture, estendendosi progressivamente a docenti-ricercatori e personale tecnico-amministrativo con esperienze significative di PE.

La traccia di intervista, ispirata ai principi dell'Appreciative Inquiry, si articola in quattro nuclei tematici. Il primo invita i partecipanti a rileggere la propria storia universitaria, mettendo in relazione la dimensione biografica con l'evoluzione dell'ateneo. Il secondo chiede di raccontare la "storia di Public Engagement più positiva" vissuta o osservata, così da far emergere esperienze significative e fattori di successo, oltre che ostacoli e strategie di fronteggiamento. Il terzo apre uno spazio di immaginazione dedicato all'esplorazione del "sogno di Public Engagement" per l'Unical, mentre il quarto sollecita a tradurre tale visione in proposte concrete, individuando azioni, strumenti e possibili cambiamenti organizzativi.

I risultati delle interviste vengono analizzati e restituiti in diversi momenti e contesti: negli appuntamenti del GrAPE, dove costituiscono la base per discussioni e attività di co-progettazione, e in incontri nei Consigli di Dipartimento, presentando il profilo specifico del PE emerso con riferimento alla struttura. Sul portale web di ateneo, inoltre, viene definito uno spazio per la pubblicazione del catalogo delle iniziative/attività di PE a partire dal 2019, alimentato dai processi di monitoraggio nel frattempo implementati. I risultati dell'analisi e le proposte vengono poi condivise in incontri dedicati in vista della definizione dell'agenda di policy dei candidati alla carica di

Rettore. Questi momenti di restituzione svolgono una funzione valutativa cruciale, permettendo di verificare la rispondenza delle interpretazioni, arricchire le analisi con nuovi contributi, trasformare i risultati in leve per il cambiamento.

4. Risultati ed effetti della valutazione partecipativa

4.1 La visione collettiva emersa: sogni e disegni di un'università civica

Nella visione collettiva emersa dalle interviste e approfondita nei momenti di restituzione, l'università “si sporca le mani”, esce dal campus e si radica nei territori, dialoga con le scuole, con le comunità fragili, con le amministrazioni locali, con le imprese. Essa viene descritta come un motore civico, culturale e sociale per il territorio, capace di attivare processi di rigenerazione urbana e sociale, di agire come partner educativo lungo un continuum formativo che va dalla scuola all'università, grazie a progetti condivisi, percorsi di orientamento e laboratori congiunti. I partecipanti attribuiscono, inoltre, all'ateneo il compito di promuovere iniziative rivolte a migranti, disoccupati, ex detenuti e persone in condizioni di fragilità, attraverso percorsi di alfabetizzazione linguistica e dei diritti, attività formative e azioni di accompagnamento. Parallelamente, alcuni vedono l'università come un alleato delle amministrazioni locali e delle piccole imprese, in grado di offrire consulenze, formazione e supporto all'innovazione. Infine, nella visione degli intervistati, l'ateneo assume il ruolo di promotore di cittadinanza scientifica, grazie a progetti di *citizen science*, eventi di divulgazione e spazi di dialogo tra saperi esperti e saperi profani.

Questa proiezione del futuro riflette in larga misura l'esperienza vissuta, ma la libera dalle difficoltà sperimentate e immagina un modello di impegno pubblico che superi l'approccio spontaneistico e occasionale, orientandosi verso forme più strutturate e sistemiche.

Nel significato prevalente assegnato dagli intervistati, il PE è un campo di valori più che un catalogo rigido di attività: i confini non sono rigidamente definiti, ma vengono interpretati in modo flessibile nelle diverse strutture e nei diversi ambiti disciplinari, con alcuni tratti comuni: l'orientamento a generare utilità sociale e il coinvolgimento del territorio. Questa eterogeneità conferma che il PE è un campo concettualmente poroso, che si compone di pratiche che attraversano ricerca, didattica e *public service*, rendendo difficile una delimitazione netta e richiedendo un framework interpretativo capace di cogliere tale fluidità.

Inoltre, dalle interviste emerge che il PE contribuisce a rafforzare il senso di comunità interna, perché favorisce la collaborazione tra docenti, ricercatori e personale tecnico-amministrativo (PTA) e alimenta una cultura organizzativa fondata sulla condivisione e sulla responsabilità collettiva; al tempo stesso, gli intervistati segnalano ostacoli culturali e strutturali che ne limitano lo sviluppo, indicando la necessità di forme di valorizzazione istituzionale e professionale che rendano queste pratiche sostenibili e pienamente integrate nella missione dell'ateneo.

In quanto ai *sogni* emersi dal percorso partecipativo, questi non vanno intesi come semplici esercizi immaginativi, ma come dispositivi di agency (Adams e Groves, 2007; Appadurai, 2004) che mobilitano la comunità accademica, orientano la definizione di priorità condivise e permettono di portare istanze dal basso nelle sedi centrali e nazionali di confronto. Quando tradotti in obiettivi condivisi, questi sogni diventano disegni: visioni operative capaci di guidare l'azione dell'ateneo.

4.2 Cosa serve allo sviluppo del PE

Sebbene l'approccio adottato abbia privilegiato una prospettiva generativa, il percorso partecipativo ha fatto emergere anche le criticità sperimentate. Nelle interviste, i partecipanti hanno discusso difficoltà organizzative, occasioni mancate e condizioni che hanno limitato l'efficacia delle iniziative di PE. In questo senso, il processo ha permesso di identificare fattori che ostacolano la collaborazione con i territori, trattandoli come leve di apprendimento e miglioramento (Guba e Lincoln, 1989). A tal riguardo è utile soffermarsi su tre criticità sistemiche emerse dalle interviste, che rappresentano barriere all'istituzionalizzazione del PE e all'impegno pubblico degli accademici (OECD, 2024; Watermeyer, 2015; Beere, Votruba e Wells, 2011). La prima riguarda il mancato riconoscimento nella progressione di carriera e l'assenza di incentivi che limitano il coinvolgimento di una parte della comunità accademica: l'attuale sistema di valutazione non valorizza le attività di PE, alimentando tra gli accademici (specialmente tra i più giovani) la sensazione che tale impegno sia privo di un *ritorno* o che, addirittura, sottraendo tempo alle pubblicazioni, ostacoli il proprio percorso verso l'acquisizione di ruoli strutturati o verso l'avanzamento di carriera. Il secondo nodo critico è legato all'assenza di finanziamenti dedicati. Le iniziative di PE, infatti, sono raramente sostenute da risorse finanziarie, riducendo di fatto il loro raggio d'azione e le opportunità di coinvolgimento di altri collaboratori. Infine, grava sul personale accademico un carico di lavoro e burocratico poco

sostenibile: chiamati a destreggiarsi tra didattica, servizio agli studenti, ricerca e compiti istituzionali, docenti e ricercatori – privi di adeguato supporto organizzativo – faticano a trovare il tempo aggiuntivo necessario per fare PE in maniera strutturata e non episodica, con il rischio che le attività di *community engagement* costino sacrificio (Jump, 2015).

4.3 Trasformazioni nella governance, nella cultura organizzativa e nella valutazione

Una valutazione del percorso di ricerca e sviluppo del Public Engagement, condotta attraverso analisi documentale e organizzativa³, consente di osservare effetti generati su diversi piani, riconducibili non solo ai contenuti emersi dalle interviste, ma anche al processo complessivo attivato. In particolare, sono emersi almeno tre ordini di effetti – organizzativi, culturali e valutativi – che hanno segnato in modo significativo il percorso.

Sul piano organizzativo, il lavoro di ascolto e confronto periodico con i delegati al PE di tutti i dipartimenti ha favorito un loro coinvolgimento attivo e propositivo. Il loro ruolo ha contribuito a rendere più visibile il PE nelle agende dipartimentali, a creare canali di comunicazione bidirezionale tra il centro e le strutture e a distribuire in modo più equilibrato la responsabilità della Terza Missione, evitando che restasse concentrata in poche figure. Il GrAPE è diventato uno spazio stabile di confronto, in cui si sono discusse linee guida, si sono condivise esperienze e si sono elaborate proposte comuni. In questo senso, il percorso ha favorito la costruzione di una leadership diffusa, in cui la governance centrale non è stata l'unica legittimata a definire la politica di PE, ma ha dialogato con una rete di referenti e pratiche.

Sul piano culturale, il percorso ha contribuito a far emergere e a rafforzare la consapevolezza del PE come componente fondamentale della missione universitaria. Le interviste hanno permesso a molti partecipanti di rileggere attività che svolgevano da anni (iniziative di divulgazione, collaborazioni con enti locali, scuole o associazioni, etc.) come pratiche di PE, riconoscendone il valore e la coerenza con la missione dell'ateneo. La ricostruzione delle "storie positive" ha generato anche effetti in termini di sviluppo di comunità, rafforzando il senso di appartenenza e il riconoscimento di radici comuni, mostrando che il PE, lungi dall'essere un'invenzione recente, ha

³ Sono stati analizzati: le note etnografiche di accompagnamento alle interviste, i contenuti dei report delle riunioni periodiche del GRES e del GrAPE, l'evoluzione del portale di Ateneo, la piattaforma di ateneo dedicata al monitoraggio del Public Engagement.

rappresentato una vocazione già presente, seppur poco riconosciuta. Allo stesso tempo, ha reso visibili ostacoli e frustrazioni: la mancanza di riconoscimento nei percorsi di carriera, la scarsità di risorse, la complessità burocratica, la difficoltà di comunicare le iniziative all'esterno. Questa consapevolezza ha alimentato una domanda di cambiamento e ha condotto molti docenti e membri del PTA a esprimere il desiderio che il PE non restasse affidato all'iniziativa individuale, ma diventasse una pratica strutturata, sostenuta e valorizzata a livello istituzionale.

Sul piano valutativo, in un più ampio processo di allineamento delle università italiane agli indirizzi dei sistemi nazionali di valutazione (AVA3 e VQR), il percorso partecipativo ha contribuito a orientare e rendere più consapevoli le trasformazioni, favorendo l'adozione di priorità, assetti organizzativi e strumenti (Balduzzi e Vaira, 2018; Di Maggio e Powell, 1983). Nello specifico, la base conoscitiva condivisa sulle pratiche di PE esistenti è stata utilizzata per avviare un sistema di monitoraggio permanente a supporto degli indicatori strategici di ateneo, sviluppando una piattaforma digitale disegnata in maniera collaborativa e implementata dall'ufficio ICT dell'ateneo. Tale base conoscitiva è custodita in un repository accessibile a tutte le strutture, rivelandosi uno strumento utile sia per la pianificazione strategica sia per il monitoraggio e la mappatura delle iniziative di PE promosse dal personale interno.

Inoltre, il percorso ha contribuito a una riflessione sulle dimensioni di valutazione del PE che vanno oltre il mero conteggio degli eventi. È emersa la necessità di analizzare le ricadute sociali e culturali delle attività, nonché il loro impatto accademico. Tutto ciò ha implicato – in termini di policy della valutazione universitaria e sul piano organizzativo – una infrastrutturazione di competenze e risorse a sostegno della valutazione di impatto, attivando un dialogo con il Presidio di Qualità e la governance di ateneo per l'assunzione di atti regolativi e strumenti operativi in questa direzione.

5. La partecipazione come dimensione epistemologica e politica della valutazione

L'esperienza UNICAL permette di riflettere in modo concreto su come la partecipazione possa assumere una valenza epistemologica e politica nei processi di valutazione e di policy making universitario. Con riferimento agli attori del processo, il percorso ha coinvolto una pluralità di soggetti – governance, direttori di dipartimento, delegati, docenti-ricercatori, personale tecnico-amministrativo – ognuno con differenti competenze, conoscenze

disciplinari, esperienza di gestione, capacità relazionali o organizzative, conoscenza dei territori. La partecipazione ha permesso di mettere in circolo questi saperi, riconoscendo che la valutazione del PE richiede il contributo di chi il PE lo pratica quotidianamente e, dunque, non può essere appannaggio di un solo gruppo (ad esempio, la governance o gli uffici amministrativi).

L'esperienza analizzata consente anche di osservare la capacità trasformativa della partecipazione sulla valutazione. A questo riguardo, è possibile, tracciare tre tipi di effetti.

Sul piano delle domande valutative, le interviste e i momenti di confronto hanno permesso di far emergere questioni che difficilmente sarebbero state formulate in un processo top-down (ad esempio, come il PE contribuisce al senso di comunità interna, quali ostacoli culturali e strutturali ne limitano lo sviluppo, quali riconoscimenti sono necessari per renderlo sostenibile).

Sul piano delle evidenze, la valutazione ha affiancato ai dati quantitativi (numero di eventi, partecipanti, ecc.), la valorizzazione di narrazioni ed esperienze soggettive che hanno permesso di cogliere dimensioni qualitative del PE (relazioni, fiducia, empowerment, ecc.). Sul piano degli esiti, i risultati della valutazione hanno contribuito - mediante le azioni di advocacy del GRES, in coerenza con il Manifesto APEnet (2025) - ad alimentare processi di cambiamento organizzativo (nomina di delegati, creazione di gruppi di lavoro, avvio di sistemi di monitoraggio) e culturale (nuove rappresentazioni del PE, maggiore consapevolezza del suo valore). In questo senso, la partecipazione ha reso la valutazione un processo di apprendimento collettivo, in cui la produzione di conoscenza appare inseparabile dalla trasformazione delle pratiche. La partecipazione ha, inoltre, contribuito a ridurre, almeno in parte, le asimmetrie di potere e di informazione che caratterizzano spesso i processi decisionali universitari. La scelta di ascoltare in profondità docenti e PTA, di restituire i risultati, di discutere le proposte in sedi collegiali, ha reso più aperto e trasparente il processo di definizione della politica di PE.

Ciò non significa che le gerarchie siano scomparse, né che tutte le richieste siano state accolte o attuate⁴. Il percorso ha, però, mostrato che è possibile

⁴ Il piano di mandato del nuovo Rettore eletto sul finire del 2025, in verità, recepisce diverse indicazioni provenienti dal percorso di ricerca e sviluppo del PE, condivise con il candidato nei mesi precedenti all'elezione. Viene, infatti, confermata e consolidata la direttrice strategica della Terza Missione, rinominata come "Collaborazione con la società", dando spazio al suo interno a una linea strategica per l'Impegno sociale e culturale e ad obiettivi orientati sia alla divulgazione scientifica che alla co-creazione di iniziative ad alto impatto sociale, anche nell'ambito della rigenerazione sociale e urbana. Il piano, inoltre, assegna un budget annuale a questa linea strategica, si impegna a rafforzare la struttura tecnico-organizzativa di supporto, a adottare un sistema di incentivi e premialità, nonché a potenziare le attività di monitoraggio e valutazione di impatto.

costruire spazi di negoziazione in cui le decisioni non siano semplicemente “calate dall’alto”, ma risultino da un confronto ampio tra attori diversi. In questo senso, la partecipazione ha avuto una valenza politica, redistribuendo, almeno in parte, il potere di definire cosa sia il PE, quali obiettivi debba perseguire, quali risorse debba mobilitare.

L’esperienza UNICAL mette in luce anche alcuni limiti e alcune condizioni necessarie per consolidare la partecipazione come dimensione strutturale della valutazione. Dalle narrazioni e dai momenti di confronto emerge, innanzitutto, la necessità di stabilizzare nel tempo i ruoli e le responsabilità - come quelli dei delegati e del gruppo di confronto tra delegati e referenti di ateneo e dei dipartimenti - così da evitare che il percorso dipenda da singole persone e possa invece contare su una continuità organizzativa. Si evidenzia, inoltre, l’esigenza di disegnare processi valutativi che contemplino un adeguato tempo per la partecipazione e che gli atenei si dotino di figure professionali con competenze di facilitazione ed expertise nel campo della valutazione partecipativa in ambito universitario.

Conclusioni

Alcune condizioni che influenzano lo sviluppo del PE dipendono da scelte regolative e da incentivi definiti a livello centrale, che orientano comportamenti e priorità degli atenei (Bleiklie, Enders e Lepori, 2017). In questo senso, la piena integrazione del PE nell’azione istituzionale universitaria richiede non solo iniziative locali, ma anche un ripensamento sistemico delle politiche nazionali di valutazione, finanziamento e carriera accademica, così da rendere sostenibile e riconosciuto l’impegno pubblico delle università. Come evidenziato da Capano (2018), gli atenei operano entro spazi di progettazione definiti dal centro politico-amministrativo, e il rafforzamento della Terza Missione non può prescindere da interventi riformatori a livello nazionale e da un sostegno più convinto dei governi⁵.

⁵ Al momento della stesura del presente saggio, lo scenario registra un segnale in controtendenza rispetto al processo di rafforzamento della Terza Missione nel sistema universitario italiano, associato alla riformulazione delle attività in capo ad ANVUR nel nuovo regolamento concernente la struttura e il funzionamento dell’Agenzia (DPR 12/2026). Si legge, infatti, che la stessa è chiamata a valutare la qualità complessiva delle attività di didattica e ricerca, mentre la valorizzazione della conoscenza, prima sinonimo di Terza missione e ora campo in cui quest’ultima viene inclusa – generando una certa confusione semantica se non addirittura un cosiddetto “errore categoriale” (Marradi, 2007; Ryle, 1938) – diviene oggetto di valutazione “anche su richiesta del Ministero”, sollevando ambiguità e incertezze sul futuro.

Ciò non riduce, tuttavia, l'importanza di promuovere processi di valutazione partecipativa all'interno degli atenei. In un contesto organizzativo complesso, in cui ogni dipartimento gode di un certo grado di autonomia in didattica, ricerca e, in misura ancora maggiore nel campo della Terza Missione, l'introduzione di una nuova policy di ateneo – come quella sul Public Engagement – attraverso un processo decisionale centralizzato e verticale rischia di generare adattamenti passivi e di ignorare tanto le difficoltà, quanto le idee, competenze e risorse presenti nelle periferie.

La valutazione partecipativa consente invece di far emergere conoscenze distribuite, di rendere visibili pratiche altrimenti disperse, di costruire una narrazione condivisa sul ruolo pubblico dell'ateneo e di definire criteri e priorità più inclusivi e orientati all'impatto sociale (Foresti, 2003; Palumbo, 2001; House e Howe, 2000; Orton e Weick, 1990). Lo studio di caso dell'Università della Calabria mostra come il coinvolgimento attivo di docenti e personale tecnico-amministrativo nella fase valutativa e nella co-ideazione delle politiche possa orientare il policy making interno, rafforzando la legittimità delle scelte e migliorando la loro implementazione. Inoltre, le istanze emerse dal basso possono essere portate dall'ateneo nelle sedi di confronto nazionale – come la CRUI, o le reti APEnet e NETVAL che dialogano con ANVUR e con il Ministero – contribuendo a influenzare l'agenda delle politiche universitarie e a rappresentare bisogni e prospettive maturate nei contesti locali. L'esperienza conferma che la partecipazione è al tempo stesso una dimensione epistemologica della valutazione, perché influisce sui modi in cui si produce conoscenza sulle politiche, e una dimensione politica, perché incide sui processi decisionali. Al contempo, mostra che la partecipazione non è una soluzione semplice né universale: richiede tempo, risorse e competenze, e deve confrontarsi con vincoli regolativi e strutturali. Proprio per questo, la capacità degli atenei di attivare processi partecipativi interni e di rappresentarne gli esiti nei luoghi decisionali nazionali costituisce un elemento cruciale per sostenere l'evoluzione della Terza Missione nel sistema universitario italiano.

Dichiarazioni

Le autrici dichiarano di aver utilizzato i software Microsoft Copilot e Gemini (versione 3.5 Flash-Lite) esclusivamente per attività di ricerca bibliografica. Le autrici si assumono la piena responsabilità dei contenuti, delle analisi e della versione finale del manoscritto.

Riferimenti bibliografici

Adams, B.E., Groves, C.R. (2007). *Future matters: action, knowledge, ethics*. The Study of Time / Supplements, No. 3. Leiden: Brill.

ANVUR (2018). *Linee guida per la compilazione della Scheda Unica Annuale Terza Missione e Impatto Sociale SUA-TM/IS per le Università*.

APEnet (2025). *Manifesto. Per il valore pubblico della conoscenza*. Testo disponibile al sito: https://www.apenetwork.it/wp-content/uploads/2025/06/Manifesto_APenet_2025_def.pdf

Appadurai, A. (2004). The Capacity to Aspire: Culture and the Terms of Recognition. In Rao V., Walton M. (Eds) *Culture and Public Action*, Washington DC, USA: The World Bank.

Arnaldi, S. (2020). Le politiche della ricerca e la Terza missione nelle Università: legami e trasformazioni. *Autonomie locali e servizi sociali*, 42 (1): 31-48. DOI: 10.1447/97468

Balduzzi, G., Vaira, M. (2018). La Terza Missione dell'Università come campo organizzativo e politico: Tre studi di caso in atenei e territori del Nord d'Italia. *Scuola Democratica*, 9 (3): 455-474. DOI: 10.12828/92142.

Beere, C.A., Votruba, J.C., Wells, G.W. (2011). *Becoming an engaged campus: A practical guide for institutionalizing public engagement*. New York: John Wiley & Sons.

Blasi, B. (2023). *Società e Università: Valutazione e impatto sociale*. Milano: FrancoAngeli.

Boffo, S., Moscati, R., Rostan, M. (2024). *La Terza Missione nell'università italiana: Politiche e attività in dodici atenei*. Milano: Guerini e Associati.

Bleiklie, I., Enders, J., Lepori, B., (2017). *Managing universities: policy and organizational change from a Western European comparative perspective*. London: Palgrave Macmillan.

Capano, G. (2018). Policy design spaces in reforming governance in higher education: the dynamics in Italy and the Netherlands. *Higher Education*, 75 (4): 675-694. DOI: 10.1007/s10734-017-0158-5.

Carrera, L., Castelnuovo, I., Cavaglià, 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 dei casi. *Rassegna Italiana di Valutazione*, 92: 123-144. DOI: 10.3280/RIV2025-092007

Coghlan, A.T., Preskill, H., Catsambas. T.T. (2003). An overview of appreciative inquiry in evaluation. *New Directions for Evaluation*, 100: 5-22. DOI: 10.1002/EV.96

Compagnucci, L., Spigarelli, F. (2020). The third mission of the university: A systematic literature review on potentials and constraints. *Technological Forecasting and Social Change*, 161. DOI:10.1016/j.techfore.2020.120284

Cooperrider, D.L., Srivastva, S. (1987). Appreciative Inquiry in organizational life. In Woodman R. e Pasmore W., a cura di, *Research in organizational change and development, volume 1*. Greenwich, CT: JAI Press.

Cooperrider, D.L., Whitney, D., Stavros, J.M. (2003). *Appreciative Inquiry handbook*. Bedford Heights, OH: Lakeshore Publishers.

Cooperrider, D., Whitney, D. (2005). *Appreciative inquiry: A positive revolution in change*. San Francisco: Berrett-Koehler.

De Ambrogio, U., Guidetti, C. (2023). E se provassimo a covalutare? *Impresa sociale*, 4: 47-49. <https://www.rivistaimpresasociale.it/rivista/articolo/e-se-provassimo-a-co-valutare>

Di Franco, G. (2010). *Il campionamento nelle scienze umane: Teoria e pratica*. Milano: FrancoAngeli.

Di Maggio, P.J., Powell, W.W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2): 147-160. DOI: 10.2307/2095101

Dipartimento della Funzione Pubblica (2019). *Linee guida sulla valutazione partecipativa nelle amministrazioni pubbliche*, n. 4. Testo disponibile al sito: https://performance.gov.it/system/files/LineeGuidaeRifNorm/LG_Valutazione_partecipativa_28-11.pdf

Elliott, C. (1999). *Locating the energy for change: An introduction to Appreciative Inquiry*. Winnipeg: International Institute for Sustainable Development. Testo disponibile al sito: <http://iisd1.iisd.ca/pdf/appreciativeinquiry.pdf>

Falbo, C., Giacosa, E., Perego, A. (2025). Concezioni e denominazioni nell'ambito della Terza missione. In: CRUI, a cura di, *Il ruolo delle università per il progresso della società. Dalla Terza missione alla Valorizzazione delle conoscenze*. Testo disponibile al sito: <https://www.cruir.it/archivio-notizie/il-ruolo-delle-universita%C3%A0-per-il-progresso-della-societa%C3%A0.html>

Fasanella, A., Martire, F., a cura di (2020). *Valutazione della ricerca e ricerca sulla valutazione. Riflessioni, analisi e proposte per la VQR*. Milano: FrancoAngeli.

Fonseca, L., Nieth, L. (2021). The role of universities in regional development strategies: A comparison across actors and policy stages. *European Urban and Regional Studies*, 28(3): 298-315. DOI: 10.1177/0969776421999743

Foresti, M. (2003). La partecipazione in ambito valutativo: mito o realtà? Teoria e pratica degli approcci partecipati in valutazione. *Rassegna italiana di valutazione*, 25: 45-70. DOI: 10.1400/66480

Furco, A. (2010). The engaged campus: Toward a comprehensive approach to public engagement. *British Journal of Educational Studies*, 58(4), 375-390. DOI: 10.1080/00071005.2010.527656

Goddard, J., Vallance P. (2014). *The university and the city*. Londra: Routledge

Greenwood D.J., Levin, M. (1998). *Introduction to Action Research*. Londra: Sage Publications.

Gregory, A. (2000). Problematizing Participation: A critical review of Approaches to Participation in Evaluation Theory. *Evaluation*, 6 (2): 179-199. DOI: 10.1177/13563890022209208

Guba, E.G., Lincoln, Y.S. (1989). *Fourth generation evaluation*. New York: Sage Publications.

Guba, E.G., Lincoln, Y.S. (1982). The place of value in Needs Assessment. *Educational Evaluation and Policy Analysis*, 4 (3): 311-320. DOI: 10.3102/0162373700400331

House, E.R., Howe, K.R. (2000). Deliberative democratic evaluation. *New Directions for Evaluation*, 85: 3-12. DOI:10.1002/ev.1157

Israel, B.A., Eng, E., Schulz A.J., Parker, E.A. (2012). *Methods for Community-Based Participatory Research for Health*. Hoboken: John Wiley e Sons.

Jump, P. (2015). *Public engagement means 'sacrificing' academic career*. Testo disponibile al sito: <https://www.timeshighereducation.com/news/public-engagement-means-sacrificing-academic-career>

Marradi, A. (2007). *Metodologia delle scienze sociali*. Bologna: il Mulino.

Mazzonetto, M., a cura di, (2025). *Mutual learning exercise on public engagement in ReI. Final report*. European Commission.

Meyer, M., Sporn, B. (2019). Leaving the ivory tower: Universities' third mission and the quest for legitimacy. In: Kövér Á. e Franger G., a cura di, *University and society*. Cheltenham: Edward Elgar.

OECD (2025). *Recommendation of the Council on Open Government*, OECD/LEGAL/0438. Testo disponibile al sito: <https://legalinstruments.oecd.org/public/doc/359/359.en.pdf>

OECD (2024). *Knowledge exchange and collaboration between universities and society in Italy: The ITA.CON Project*. Testo disponibile al sito: https://www.oecd.org/en/publications/knowledge-exchange-and-collaboration-between-universities-and-society-in-italy_ae130814-en.html

- OECD (2023). *Open Government for Stronger Democracies: A Global Assessment*. Paris: OECD Publishing. DOI: 10.1787/5478db5b-en
- Orton, J.D., Weick, K.E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review*, 15(2), 203-223. DOI: 10.2307/258154
- Palumbo, M. (2001). *Il processo di valutazione: Decidere, programmare, valutare*. Milano: FrancoAngeli.
- Palumbo, M., Torrigiani, C. (2013). Participatory evaluation in the field of social policies: who, what, why, how and when. *Working Paper Series, FPpV* 25.
- Palumbo, M., Torrigiani, C., a cura di (2009). *La partecipazione tra ricerca e valutazione*. Milano: FrancoAngeli.
- Patton, M.Q. (1994). Developmental Evaluation. *Evaluation Practice*, 15(3): 311-319. DOI: 10.1177/109821409401500312
- Patton, M.Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage Publications.
- Reason, P., Bradbury, H., a cura di (2008). *The SAGE Handbook of Action Research*. New York: SAGE Publications Ltd.
- Rebien, C.C. (1996). Participatory Evaluation of Development Assistance: Dealing with Power and Facilitative learning. *Evaluation*, 2 (2): 151-172. DOI: 10.1177/135638909600200203
- Regonini, G. (2001). *Capire le politiche pubbliche*. Bologna: Il Mulino.
- Riquier, G. (1997). *Stakeholders' Participation in Public Policy Evaluations: impact on Organizational Learning*, paper presentato al congresso dell'European Evaluation Society, Stockholm.
- Ryle, G. (1938). Categories. In: Aa. Vv. *Proceedings of the Aristotelian Society*, 38:189-206.
- Scardigno, A.F., Blasi, B., Pastore, S., Guaragno, F. (2019). Valutare la Terza Missione dell'Università: processi partecipativi e meccanismi di consultazione. *Rassegna Italiana di Valutazione*, 74 (1): 55-76. DOI: 10.3280/RIV2019-074004
- Scavi, M. (2021). La città, la democrazia deliberativa e i cittadini come costruttori di comunità/2. Il Bo Live. Testo disponibile al sito: <https://ilbolive.unipd.it/it/news/societa/citta-democrazia-deliberativa-cittadini-come-0>
- Scavi, M., Susskind, L.E. (2011). *Confronto creativo. Dal diritto di parola al diritto di essere ascoltati*. Milano: Et Al Edizioni.
- Torrigiani, C. (2025). Co-programmare, co-progettare e valutazione delle politiche sociali. *Rassegna italiana di valutazione*, (29)91: 30-47. DOI: 10.3280/RIV2025-091003
- Valastro, A. (2010). Partecipazione, politiche pubbliche, diritti. Le regole della democrazia partecipativa. In: Valastro A., a cura di, *Itinerari per la costruzione di un metodo di governo*. Napoli: Jovene editore.
- Watermeyer, R. (2015). Lost in the 'third space': the impact of public engagement in higher education on academic identity, research practice and career progression. *European Journal of Higher Education*, 5(3), 331-347. DOI: 10.1080/21568235.2015.1044546
- Watson, D., Hollister, R., Stroud, S.E., Babcock, E. (2011). *The engaged university: International perspectives on civic engagement*. London: Routledge.
- Weick, K.E. (1982). Management of organizational change among loosely coupled elements. In: Goodman P.S. e Associates, a cura di, *Change in organizations*. San Francisco, CA: Jossey-Bass.
- Whitney, D., Trosten-Bloom, A. (2003). *The power of Appreciative Inquiry*. San Francisco, CA: Berrett-Koehler.
- Zomer, A., Bennenworth, P. (2011). The rise of the third mission. In: Enders J., de Boer H.F., Westerheijden D.F., a cura di, *Reform of higher education in Europe*. Rotterdam: Sense.

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 (anna fausta.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 conflige 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 imprevisi, 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.

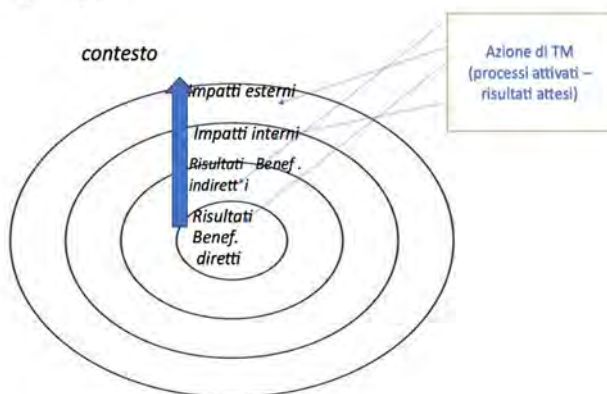
2. Una proposta per la valutazione

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 dal "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., Castelnovo, 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 CenturyText*. 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

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.



Management, finanza,
marketing, operations, HR

Psicologia e psicoterapia:
teorie e tecniche

Didattica, scienze
della formazione

Economia,
economia aziendale

Sociologia

Antropologia

Comunicazione e media

Medicina, sanità



Architettura, design,
territorio

Informatica, ingegneria
Scienze

Filosofia, letteratura,
linguistica, storia

Politica, diritto

Psicologia, benessere,
autoaiuto

Efficacia personale

Politiche
e servizi sociali



FrancoAngeli

La passione per le conoscenze

La *Riv. Rassegna Italiana di Valutazione* è Rivista scientifica ANVUR Classe A per i settori concorsuali: area 14/C1 (Sociologia generale), area 14/C2 (Sociologia dei processi culturali e comunicativi), area 14/C3 (Sociologia dei fenomeni politici e giuridici), area 14/D1 (Sociologia dei processi economici, del lavoro, dell'ambiente)

La Rivista utilizza i criteri del processo di referaggio indicati dal Coordinamento delle riviste italiane di sociologia (CRIS). Le proposte di pubblicazione sulla RIV devono essere sottoposte secondo le modalità di cui al sito *FrancoAngeli* (www.francoangeli.it)

La Rivista è indicizzata su: Catalogo italiano dei periodici/Acnp, Ebsco Discovery Service, Essper, Google Scholar, JournalTOCS, ProQuest Summon, Torrossa - Casalini Full Text Platform

La *Rassegna Italiana di Valutazione* esce dal 1996:

- le annate dal 1996 al 2001 sono esaurite
- nel 2002 la *Riv* non è uscita
- dal 2003 esce per i tipi di FrancoAngeli
- dal 2026 esce in open access

Publicato con licenza *Creative Commons Non commerciale-Non opere derivate 4.0 Italia* (CC-BY- NC-ND 4.0 IT)

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito: <https://creativecommons.org/licenses/by-nd/4.0/deed.it>

Sono riservati i diritti per Text and Data Mining (TDM), AI training e tutte le tecnologie simili.

I link attivi e gli eventuali QR code inseriti nel volume sono forniti dall'autore. L'editore non si assume alcuna responsabilità su link e QR code ivi contenuti che rimandano a siti non appartenenti a FrancoAngeli.

Redazione, amministrazione, distribuzione, abbonamenti: Viale Monza, 106 - 20127 Milano - e-mail: riviste@francoangeli.it

Autorizzazione Tribunale di Milano n. 187 del 17 marzo 2003 - Quadrimestrale – Direttore responsabile: Stefano Angeli ISSN 1120-1972-5027.

Copyright © 2026 by FrancoAngeli s.r.l. - Stampa: Geca - Divisione Libri di Ciscra Spa, Via Belvedere 42, 20862 Arcore (MB).

II quadrimestre 2026 – prima pubblicazione settembre 2026

Rassegna Italiana di Valutazione

XXX - 95/2026

La Riv. Rassegna Italiana di Valutazione è la rivista dell'AIV ed è pubblicata da FrancoAngeli in open access.

Ha l'obiettivo di elevare il livello del dibattito valutativo in Italia, favorendo il dialogo tra le diverse anime disciplinari e professionali della valutazione.

I temi affrontati sono oggi al centro del processo attraverso cui la valutazione si sta consolidando nel settore pubblico, privato e del terzo settore: gli obiettivi e il funzionamento delle istituzioni pubbliche con finalità valutative; il ruolo dei nuclei di valutazione; il quadro deontologico e le prassi metodologiche consolidate per la valutazione dei programmi; la definizione di standard di qualità e di sistemi di accreditamento e di valutazione; il mercato della valutazione; l'obbligatorietà dei processi valutativi e altri snodi cruciali per lo sviluppo della cultura valutativa nel Paese.

I contenuti e le idee espresse dagli autori e dalle autrici non riflettono necessariamente gli orientamenti dell'Associazione Italiana di Valutazione né dei Comitati di direzione e editoriale, che intendono garantire uno spazio aperto a orientamenti e approcci diversi, purché scientificamente argomentati e non in contrasto con i principi contenuti nel Codice deontologico associativo.

