

Disegni orientati alla teoria: una guida per la valutazione dei programmi di occupazione e innovazione

Mita Marra

Working Papers della Fondazione Giacomo Brodolini

Direttore scientifico della collana

Anna Maria Simonazzi, Presidente della Fondazione Giacomo Brodolini

Direttore scientifico della serie "Scuola Europea di Relazioni Industriali - SERI"

Michele Faioli, Università Cattolica del Sacro Cuore

Fondazione Giacomo Brodolini

00185 Roma - Via Goito, 39

tel. 0644249625

fax 0644249565

info@fondazionebrodolini.it

www.fondazionebrodolini.it

ISBN: 9788895380605

*Pubblicazione realizzata in collaborazione con l'Università degli Studi di Napoli
Federico II.*

***Disegni orientati
alla teoria: una guida
per la valutazione dei
programmi di occupazione
e innovazione***

Mita Marra

22

Sommario

INTRODUZIONE.....	1
GLOSSARIO.....	3
1. OLTRE I LUOGHI COMUNI SULLA VALUTAZIONE.....	6
1.1 I programmi: oltre agli indicatori c'è di più	6
1.2 Programmi non standardizzabili e diversità delle reazioni dei beneficiari	9
1.3 Bandi di valutazione da segmentare per affrontare la complessità dei programmi	13
1.4 Dal quadro logico alla ricostruzione del processo	14
1.5 Il monitoraggio non è la valutazione di processo	17
2. LE DOMANDE VALUTATIVE.....	21
2.1 Collegare le domande agli obiettivi dello studio	21
2.2 Le caratteristiche delle domande valutative	24
2.3 L'importanza del contesto	25
2.4 Le domande sulla causalità	27
3. CAUSALITÀ E DISEGNI VALUTATIVI.....	28
3.1 La causalità sequenzialista e approccio controfattuale	28
3.2 Regolarità e approccio statistico	30
3.3 Le causalità multiple negli studi di caso	31
3.4 La causalità generativa e theory-based	32
3.5 Una panoramica sui disegni valutativi non sperimentali	33

4. LA VALUTAZIONE ORIENTATA ALLA TEORIA.....	37
4.1 Ricostruzione della teoria del programma	38
4.2 Ricostruzione della teoria dell'implementazione	41
4.3 La teoria del cambiamento	42
4.4 La valutazione realistica: una specifica forma di theory-based evaluation	43
4.5 Limiti e vantaggi associati al theory-based	45
4.6 L'approccio theory-based nelle politiche attive del lavoro	49
5. COMPLESSITÀ E VALUTAZIONE.....	62
5.1 I principi della scienza della complessità rilevanti per la valutazione	62
5.2 Le dimensioni di complessità nei programmi e nei contesti	63
5.3 La politica di Smart Specialization: questioni valutative	67
5.4 Le carenze delle valutazioni realizzate	69
5.5 Come valutare una politica di innovazione?	73
CONCLUSIONI.....	84
RIFERIMENTI BIBLIOGRAFICI.....	86

INTRODUZIONE

La complessità dei problemi e dei contesti sociali che le odierne politiche pubbliche provano ad affrontare richiede informazioni affidabili e rilevanti sui processi e sugli effetti degli interventi. La trasformazione ecologica e tecnologica, esige, inoltre, una serrata osservazione dell'implementazione e dei risultati dei programmi di transizione alla sostenibilità, anche al fine di introdurre correttivi tempestivi e pertinenti a ridefinire le strategie d'intervento, laddove necessario.

La presente guida metodologica si propone di orientare valutatori, amministratori, manager e studenti nella definizione di **disegni valutativi basati sulla teoria** nel campo delle politiche attive del lavoro e delle strategie di specializzazione intelligente ricomprese nell'ambito della più ampia politica di coesione europea, con riferimento in particolare alla valutazione dei programmi finanziati dal Fondo Sociale Europeo e del Fondo Europeo per lo Sviluppo Regionale. Tra i vari approcci metodologicamente robusti e politicamente rilevanti, la valutazione orientata alla teoria aiuta a comprendere se, come, perché, in quali circostanze e per quali beneficiari le politiche attive del lavoro funzionano. Incorporando costrutti e principi mutuati dalla scienza della complessità, la valutazione orientata alla teoria aiuta a mettere a fuoco le complementarità e le interdipendenze che le politiche d'innovazione intendono sviluppare in contesti socioeconomici eterogenei.

Esaminando le premesse che sottendono i disegni valutativi orientati alla teoria, la presente nota mette in evidenza i vantaggi – ma anche i limiti – dell'impianto *theory-based* e offre spunti concettuali e metodologici per progettare valutazioni di processo e di impatto utili a diffondere e consolidare la cultura della valutazione nell'ambito della pubblica amministrazione e del terzo settore in Italia (Marra, 2017).

Prima di addentrarsi nelle specificità dell'approccio orientato alla teoria e nelle sue applicazioni nell'ambito delle politiche del lavoro e dell'innovazione (v. capitoli 4 e 5), la guida si sofferma su alcuni luoghi comuni sulla valutazione (v. capitoli 1, 2 e 3). Si tratta di pratiche valutative e routine amministrative che finiscono per semplificare la complessità dei fenomeni da esplorare, riducendo l'utilità e l'innovatività delle conoscenze valutative a vantaggio della programmazione delle politiche e degli investimenti per la sostenibilità. La guida prova a scardinare idee preconcepite sulla natura dei programmi e sulla verifica degli impatti che avallano valutazioni di bassa

qualità - carenti sul piano della raccolta delle informazioni di prima mano e della formulazione di giudizi sul merito e sul valore degli interventi. La guida intende perorare, piuttosto, l'idea di una valutazione che è analisi multidisciplinare delle politiche e dei contesti e critica costruttiva delle scelte pubbliche - contro la mentalità burocratica dell'adempimento ma anche contro i tecnicismi della misurazione per i soli addetti ai lavori.

Il presente lavoro è organizzato in 5 capitoli. A partire dall'individuazione di alcuni luoghi comuni sulla valutazione, il primo capitolo mette in evidenza le questioni salienti che saranno affrontate nei capitoli successivi. Il secondo capitolo si sofferma sulle domande valutative come passaggio essenziale per sviluppare il disegno di uno studio di valutazione orientato alla teoria. Il terzo capitolo mette a fuoco le diverse interpretazioni della causalità che sottendono i disegni di valutazione di impatto. Il quarto capitolo si focalizza sull'approccio orientato alla teoria e lo contestualizza nell'ambito delle politiche attive del lavoro. Il quinto capitolo ricostruisce alcune dimensioni chiave della scienza della complessità e del pensiero sistemico che possono informare le valutazioni delle strategie di specializzazione intelligente per l'innovazione e lo sviluppo regionale. In ultimo, le conclusioni ripercorrono sinteticamente i temi trattati nella presente guida.

GLOSSARIO

Attivazione: l'OCSE definisce le strategie di attivazione quelle misure tese ad aumentare la partecipazione al mercato del lavoro al fine di contrastare gli effetti della disoccupazione e migliorare l'occupabilità.

Approccio o disegno di valutazione: insieme di principi che guidano i processi di creazione della conoscenza e che orientano l'analisi e la raccolta dei dati rispetto a dimensioni d'interesse da valutare.

Bando per l'affidamento di studi valutativi: è la richiesta di servizi valutativi che le amministrazioni pubbliche e le organizzazioni del terzo settore esprimono formalmente in ossequio ai regolamenti comunitari e nazionali vigenti.

Complessità: è un termine tecnico che denota l'ambiguità dei risultati dei programmi con finalità sociale e l'incertezza che caratterizza i fenomeni che investono i contesti in cui si attuano i programmi e che incidono sull'esito degli stessi.

Contesto: luogo fisico e insieme di regole codificate, norme sociali e culture che influenzano tanto le scelte dei decisori quanto i comportamenti degli attori e dei destinatari dei programmi.

Emergenza: è un termine tecnico che caratterizza il fenomeno del sorgere di un cambiamento socioeconomico, culturale o politico-istituzionale con proprietà nuove a partire da un intervento.

Embeddedness: è un termine inglese che viene tradotto con radicamento e che si riferisce alle reti socioeconomiche che influenzano le scelte pubbliche e le relazioni tra gli agenti che operano in un contesto di scala variabile.

Impatto: ovvero il cambiamento di medio-lungo periodo ottenuto (risultato) o dovuto ad un intervento di politica pubblica.

Input: risorse da investire o costi da sostenere al fine di realizzare un programma o un investimento pubblico.

Metodo: insieme di regole e di principi che guidano la raccolta delle informazioni e dei dati e l'analisi degli stessi nei processi di valutazione.

Output: realizzazione fisica o prodotto o servizio che introducono un cambiamento come effetto di un programma o di una politica. La valutazione dell'output si concentra sulle realizzazioni fisiche di un programma o sugli indicatori finanziari che indicano il livello di erogazione della spesa.

Outcome: il cambiamento introdotto dall'output in termini di risultato finale che offre vantaggi diretti. La valutazione dell'outcome o di risultato esamina il cambiamento intervenuto come effetto di un programma.

Politiche attive del mercato del lavoro: Le politiche attive del lavoro sono tutte le iniziative messe in campo dalle Istituzioni, nazionali e locali, per promuovere l'occupazione, l'inserimento e il reinserimento lavorativo.

Processo (valutazione di): analisi del processo di attuazione di un intervento teso a generare un cambiamento nel contesto. La valutazione di processo esamina in che modo un intervento è stato implementato, se l'attuazione ha seguito fedelmente il piano, se sono intervenute modifiche, perché e quali risultati si sono generati nella struttura del programma

Programma: insieme di risorse materiali e immateriali intenzionalmente tese a generare un cambiamento socioeconomico, culturale e/o politico-istituzionale.

Razionalità limitata: in un contesto caratterizzato da informazioni incomplete, l'essere umano risponde agli stimoli ambientali e alle opportunità offerte dalle politiche secondo processi limitatamente razionali, in particolare in condizioni di incertezza e complessità. Secondo le acquisizioni sperimentali della psicologia sociale, l'individuo presenta caratteristiche comportamentali automatiche e inconsapevoli che lo portano a decidere in modo sub-ottimale anche in situazioni di limitata complessità e incertezza.

Razionalità strumentale o mezzi-fini (o scelta razionale): secondo il modello di razionalità decisionale, una politica o un programma assumono il comportamento razionale dell'essere umano o del destinatario che rispondono a incentivi e sanzioni secondo un calcolo economico, assumendo informazioni complete.

Specializzazione intelligente (*smart specialization*): si tratta di strategie di innovazione nazionali o regionali che definiscono le priorità delle politiche pubbliche e degli investimenti produttivi allo scopo di creare un vantaggio competitivo. Le strategie di specializzazione intelligente mirano a potenziare ricerca e innovazione secondo le esigenze imprenditoriali, al fine di rispondere alle opportunità emergenti a livello territoriale e agli sviluppi del mercato locale e globale.

Teoria del cambiamento: processo rigoroso e partecipativo nel quale differenti portatori di interesse nel corso di una pianificazione articolano i loro obiettivi di lungo termine [impatto] e identificano le condizioni che essi reputano debbano dispiegarsi [pre-conditions] affinché tali obiettivi siano raggiunti. Tali condizioni sono schematizzate nei cambiamenti [outcome] che si vogliono ottenere e sono organizzate graficamente in una struttura causale [catena dei risultati].¹ La teoria del cambiamento si compone della teoria del programma e della teoria dell'implementazione.

¹ Dana H. Taplin, *Helène Clark (2013) Theory of Change basics, Act Knowledge, New York.*

1. OLTRE I LUOGHI COMUNI SULLA VALUTAZIONE

1.1 I programmi: oltre agli indicatori c'è di più

È noto che, affrontando un problema o un disagio sociale, una politica pubblica mira a formulare e attuare una soluzione che migliori le condizioni socioeconomiche di una comunità a livello locale, nazionale o globale. Gli obiettivi di una politica esprimono le finalità di un cambiamento sociale desiderato rispetto allo *status quo* – ad esempio, l'aumento degli occupati in uno specifico settore produttivo o una maggiore partecipazione delle donne al mercato del lavoro. Le misure che vengono attuate nei contesti sociali mobilitano risorse umane, finanziarie, organizzative e temporali che si presume possano trasformare le condizioni di vita o le modalità di produzione e di lavoro.

Nell'approccio valutativo tradizionale si assume che le politiche pubbliche siano un insieme di misure o strumenti con gradi diversi di coercizione e volontarietà, coerenti con gli obiettivi dichiarati. Ad esempio, se l'obiettivo è contrastare l'evasione fiscale, gli strumenti di politica pubblica possono essere di natura fiscale (es. aliquote fiscali ridotte), o legati alla legislazione (es. un aumento delle sanzioni per gli evasori fiscali) o, ancora, basati sulla comunicazione (es. norme sociali utilizzate per informare una comunità locale sul comportamento virtuoso di altre comunità vicine o distanti). Il primo strumento farà appello all'incentivazione, il secondo alla coercizione, il terzo all'informazione e alla sensibilizzazione. Lo stesso obiettivo di policy può, quindi, essere raggiunto con diversi strumenti che possono presentare modalità d'intervento specifiche e sostanzialmente differenti. Gli interventi, a loro volta, influenzano una o più determinanti del comportamento per cui l'individuo di fronte a certe norme o informazioni si presume sia sollecitato a comportarsi come desiderato dal decisore. Dovrebbe, cioè, essere razionale per l'individuo, destinatario di un intervento, decidere secondo le finalità perseguite dalla politica o dal programma, anche se l'effetto atteso può anche non avverarsi. È noto, infatti, che le conseguenze inattese delle azioni umane sono la regola piuttosto che l'eccezione (Callon *et al.*, 2009).

Eppure, secondo i modelli di pianificazione razional-comprensivi, gli strumenti di policy sono subordinati agli obiettivi strategici al fine di conseguire risultati attesi come, ad esempio, l'aumento del numero di imprese internazionalizzate o delle famiglie al di fuori della soglia della povertà. I risultati attesi sono articolati in un arco temporale di

breve, di medio e di lungo periodo e misurati secondo metriche di natura finanziaria, di realizzazione (output) e d'impatto (outcome), al fine di documentare il progresso compiuto rispetto agli obiettivi. Gli indicatori sono, quindi, misure utili ad intercettare e caratterizzare, sia sul piano quantitativo che sul piano qualitativo, il cambiamento che si verifica in un contesto, presumibilmente, come effetto di una politica o di un programma. Gli indicatori possono essere stime probabilistiche dei risultati attesi – misure approssimative (*proxies*), descrittori degli effetti osservati o osservabili nei contesti sociali – oppure definiscono gli standard da raggiungere durante l'attuazione e al termine del programma.

Spesso un programma viene espresso in termini definiti o definibili e riproducibili attraverso gli indicatori di risultato o di impatto: ad esempio, una modifica dei limiti di velocità in autostrada di un intervento di regolazione come il codice della strada o la diminuzione dei carichi di lavoro di un operatore sociale o l'aumento o il decremento dei pagamenti nell'erogazione di un sussidio (Weiss, 1998). Negli esempi riportati, l'intervento si caratterizza attraverso le misure delle dimensioni del cambiamento da osservare. La valutazione dell'intervento consiste, quindi, nella misurazione del cambiamento secondo scale e metriche variabili – come, ad esempio, l'aumento del numero di casi da gestire o dei sussidi di disoccupazione da erogare, la diminuzione degli incentivi assegnati, la variazione dell'ammontare delle sovvenzioni contro la povertà o l'incremento degli affitti delle case popolari. In questi casi, occorre precisare che si tratta di verifiche di monitoraggio degli indicatori prescelti piuttosto che di una valutazione dei risultati o degli impatti.

Numerosi bandi per l'affidamento della valutazione del Fondo Sociale Europeo e del Fondo Europeo di Sviluppo Regionale richiedono, infatti, di verificare e aggiornare gli indicatori di monitoraggio dei risultati. È importante ricordare che valutare gli esiti di un programma d'inserimento lavorativo non consiste nel misurare la variazione degli indicatori dell'occupazione, o di altri indicatori selezionati, prima e dopo l'intervento realizzato; e non basterà verificare l'andamento degli indicatori – come, ad esempio, la variazione percentuale dell'occupazione femminile e dei tassi di attività rispetto ai tassi di occupazione e di disoccupazione – per valutare gli esiti o l'impatto di medio-lungo termine di una politica attiva del mercato del lavoro. La verifica degli indicatori prima e dopo l'intervento sono utili a monitorare il processo di attuazione, ma non ad attribuire gli effetti osservati al programma, secondo una relazione di causalità diretta.

Gli indicatori sono "qualcosa" di osservabile/rilevabile/misurabile che "sta per" ossia viene utilizzato al posto di "qualcos'altro". Di per sé, gli indicatori non consentono di

ricostruire la concatenazione di effetti attesi e inattesi di un intervento complesso che mira a trasformare le condizioni e le dinamiche del mercato del lavoro. Anzi, talvolta, gli indicatori possono celare importanti dimensioni di interesse collegate ad altri fenomeni, da esaminare in maniera più approfondita, facendo ricorso ad una pluralità di fonti di informazione. Ad esempio, l'indice di dipendenza associato al progressivo invecchiamento della società italiana non rivela il fatto che, con un sistema di welfare che fa leva sul supporto familiare il, gli anziani, nel loro ruolo di nonni, garantiscono una rete di sicurezza talora indispensabile per le giovani coppie che provano a conciliare tempi di lavoro e tempi di cura della famiglia.

I bandi con cui si richiedono i servizi di valutazione nell'ambito dei programmi cofinanziati dai fondi strutturali finiscono, sovente, per accordare molta enfasi alla verifica degli indicatori di monitoraggio nell'ambito della valutazione di risultato o di impatto. Come già sottolineato precedentemente, ai valutatori si chiede di aggiornare le metriche che caratterizzano un dato programma operativo regionale o nazionale, attraverso la verifica dei parametri quantitativi. Ma i parametri prescelti non sempre rendono conto della variabilità delle reazioni dei beneficiari e delle dimensioni immateriali che concorrono a generare il cambiamento culturale e sociale che i programmi perseguono.

Quando il focus della valutazione si concentra prevalentemente sull'aggiornamento degli indicatori di monitoraggio, di fatto, si rischia di trascurare l'individuazione degli effetti inattesi, che dovrebbero essere invece al centro del processo di valutazione. È importante non sottovalutare la novità delle trasformazioni organizzativo-istituzionali e socioeconomiche che – incrementali o radicali che siano – sono spesso imprevedibili, date le condizioni di elevata incertezza che caratterizzano i sistemi socioeconomici odierni. Si perpetuano, in tal modo, culture organizzative ancora poco sensibili e capaci di apprendere dall'esperienza di implementazione degli interventi. Il rischio è di introdurre errori e distorsioni nel giudizio – soprattutto, in presenza di programmi sperimentali o partecipati che prevedono il coinvolgimento di molteplici attori impegnati nella programmazione e nell'attuazione. A monte, le richieste delle amministrazioni non mettono in questione la significatività degli indicatori utilizzati così come la qualità e la rappresentatività dei dati disponibili. A valle, con le esigue metriche da considerare, le valutazioni realizzate finiscono per fornire prove poco adeguate ad informare le decisioni che investono programmi complessi e di ampia scala.

1.2 Programmi non standardizzabili e diversità delle reazioni dei beneficiari

Secondo i modelli di pianificazione razionale (razionalità mezzi-fini), i programmi e le attività che li compongono sono strumenti scelti coerentemente con gli obiettivi. Si assume che le attività realizzate siano responsabili degli effetti da misurare attraverso gli indicatori selezionati. In condizioni di non elevata incertezza, la valutazione *ex ante* stima, in maniera probabilistica, gli effetti attesi ed esamina *ex post* gli scostamenti tra gli effetti verificabili e quelli pianificati o desiderati.² L'impatto è quantificato nel saldo netto tra effetti osservati ed effetti che si sarebbero verificati in assenza del programma, per cui si assume che le attività realizzate determinino i mutamenti del contesto a prescindere dalla reazione soggettiva dei beneficiari e degli altri attori coinvolti.

Nella pratica delle organizzazioni sia pubbliche che private, i programmi sono frutto dell'esperienza di manager, professionisti, esperti e operatori che spesso intuitivamente assumono decisioni o intraprendono corsi di azione per risolvere i problemi che incontrano nel lavoro quotidiano. Come sostengono Carol Weiss (1998) e Patricia Rogers (2008) – due eminenti studiose della valutazione a livello internazionale – i programmi non scaturiscono da un calcolo razionale che si articola in una serie di affermazioni logicamente chiare e consequenziali che giustificano perché, a fronte dei costi da sostenere, i benefici ripagheranno gli sforzi profusi. Nelle situazioni reali, le emozioni e le limitate facoltà computazionali della mente umana orientano la pratica delle organizzazioni per far fronte ai problemi o aiutano a sviluppare capacità di adattamento rispetto a situazioni di crisi. In tali circostanze, i valutatori dovranno ricostruire ciò che i programmi intendono realizzare e come intendano farlo – in altri termini, la valutazione è chiamata a comprendere gli obiettivi espliciti e impliciti, dichiarati e taciti, ma anche le modalità con cui gli obiettivi vengono raggiunti e gli effetti che si sono concretamente dispiegati. È un lavoro di analisi e ricerca, non solo di verifica degli indicatori di monitoraggio o di impatto.

Considerare i programmi come strumenti coerenti rispetto agli obiettivi rischia di farci perdere di vista la pluralità dei significati che diversi beneficiari e attori attribuiscono agli effetti dei programmi stessi. Un effetto può essere percepito in maniera differente a seconda dei beneficiari e dei contesti sociali ove programmi identici possono presentare specificità anche significative nelle modalità di organizzazione e attuazione degli interventi. I programmi sono, infatti, iniziative che investono

² *In situazioni complesse, quando i risultati d'interesse sono molteplici e incerti, la capacità computazionale, secondo le norme della razionalità formale, si riduce notevolmente.*

aspirazioni e percezioni, idee e credenze, procedure e strutture organizzative, attività e allocazioni finanziarie ed esprimono una forte intenzionalità. Un programma non mobilita esclusivamente risorse materiali, tangibili e misurabili, ma anche e soprattutto componenti immateriali, talora incommensurabili, come le aspettative degli attori coinvolti, gli stili di lavoro o di interazione tra i membri di un'organizzazione, le routine e le norme sociali che permeano un'istituzione o un insieme di istituzioni, una comunità o più comunità a livello regionale/nazionale. Anche le dimensioni non tangibili possono influenzare i comportamenti individuali e collettivi in quanto le reazioni intuitive, emozionali e affettive dei destinatari dei programmi esercitano un potere di trasformazione sociale ed economica³ – una capacità sovente sottovalutata o considerata come effetto trascurabile di disturbo che va superato potenziando il ragionamento (probabilistico) e il controllo.

Se un intervento è standardizzabile, può essere attuato nello stesso modo in tutti i contesti in cui esso viene realizzato. In tali circostanze, il programma è relativamente semplice da valutare (Funnell e Rogers, 2011). Un intervento semplice non è necessariamente banale o di facile attuazione, ma è più facile valutare se esso funziona come pianificato (v. tabella 1). Ad esempio, l'amministrazione dei sussidi per la disoccupazione può essere considerata come un intervento semplice e uniformemente applicato in tutti i territori del nostro paese. I criteri di ammissibilità per le prestazioni dell'assicurazione contro la disoccupazione sono standardizzati, il livello di pagamento è definito in anticipo per diversi gruppi target, esiste solo un'organizzazione responsabile del trasferimento dei fondi e lo stesso importo viene erogato ovunque per ciascun gruppo target. Eppure non è detto che l'effetto dei sussidi sia identico per tutti i destinatari.

La maggior parte degli interventi, specialmente nell'ambito delle politiche sociali, del lavoro e dell'innovazione (come sarà evidente più oltre), non prevedono azioni

³ *La decisione umana è fortemente influenzata da fattori emozionali e affettivi. L'affettività è una guida euristica nella raccolta delle informazioni e nella elaborazione della scelta. Ragioni di tipo affettivo ci portano a non considerare i costi passati per una iniziativa come costi sommersi per concentrarsi solo su quelli futuri e sui relativi benefici. I costi già affrontati condizionano le nostre scelte, anche se l'investimento fatto non è più recuperabile e non porterà ad un beneficio. Ciò porta spesso ad una escalation irrazionale di investimenti su progetti fallimentari e al successivo fallimento. Si veda sul punto Hirschman (1967) e il Principio della mano che nasconde in Marra (2020) disponibile su <https://uni-tuebingen.de/en/fakultaeten/wirtschafts-und-sozialwissenschaftliche-fakultaet/faecher/fachbereich-sozialwissenschaften/politikwissenschaft/institut/lehrende/comparative-public-policy-professor-seeleib-kaiser/journal-of-european-social-policy>*

standardizzabili. Essi presentano differenze significative nella natura delle misure messe in campo, nelle modalità di attuazione e, quindi, anche nei processi di valutazione. In tal senso, i programmi sono sistemi complessi che interagiscono con i beneficiari e gli attori in gioco. Non sono le specifiche attività di un programma a produrre un effetto – atteso o inatteso – di cambiamento, ma le risposte che i destinatari e le comunità più ampie manifestano quando si relazionano con il mondo amministrativo o con il mondo delle imprese in contesti socioeconomici anche molto differenziati. Per questo motivo, per valutare un programma occorre esplorare le reazioni dei beneficiari e degli attori che, modificando i propri comportamenti, trasformano i contesti (v. più oltre causalità generativa). E poiché sia gli interventi realizzati che le risposte dei beneficiari sono sovente differenziati, i valutatori dovranno rendersi conto del livello di complessità da affrontare in considerazione della variabilità delle reazioni da esaminare.

TAB. 1 CLASSIFICAZIONE DELLE DIMENSIONI SEMPLICI, COMPLICATE E COMPLESSE PER LA VALUTAZIONE DI PROCESSO E DI IMPATTO DEI PROGRAMMI

Obiettivi	Dimensioni semplici	Dimensioni complicate	Dimensioni complesse
	Singoli	Multipli	Emergenti
Natura dell'intervento	Standardizzata	Obiettivi e risultati sono valutati in maniera differente	Non standardizzabile, adattiva, evolutiva e emergente
Modalità di implementazione	Una sola agenzia	Molteplici organizzazioni già identificate con ruoli specificati	Molteplici organizzazioni con ruoli emergenti e imprevedibili
Procedure di attuazione	Procedure identiche nei vari contesti di attuazione	Procedure differenti a seconda dei contesti e degli attori coinvolti	Procedure sensibili alle condizioni iniziali dei contesti
Valutazione	Valutazione di impatto	Valutazione di impatto e di processo	Valutazione di processo

Fonte: Rielaborazione da Funnell e Rogers (2011)

A tal proposito, Funnell e Rogers (2011) classificano i programmi come semplici, complicati e complessi in relazione alla natura dell'intervento, alle modalità di implementazione, alle procedure di attuazione, ai processi di valutazione.

La classificazione evidenziata nella tabella 1 ci permette di ricostruire un quadro sfaccettato delle relazioni che esistono tra i programmi e le agenzie di attuazione, gli altri attori coinvolti, i beneficiari che possono reagire in maniera diversa alle attività previste. Un intervento presenta aspetti complicati quando i destinatari, gli attori o le organizzazioni di attuazione valutano gli esiti da prospettive divergenti, quando i risultati previsti sono molteplici e potenzialmente concorrenti, o quando gli impatti possono emergere, in maniera imprevedibile, a diversi livelli di analisi.

Le politiche per l'occupazione, ad esempio, presentano una pluralità di aspetti complicati che emergono a causa del basso grado di consenso tra gli attori rispetto agli obiettivi e agli impatti attesi. Si pensi alle misure di sostegno al reddito dei disoccupati e agli effetti di dipendenza che si presume siano generati dai sussidi rispetto all'obiettivo di partecipazione al mercato del lavoro. Le agenzie di attuazione delle misure di attivazione sono molteplici e presentano, sovente, interessi concorrenti: pensiamo, ad esempio, ai servizi pubblici e privati di intermediazione tra domanda e offerta di lavoro, alla pluralità di casse di disoccupazione, ai sindacati, ai datori di lavoro, ecc... Inoltre, i mercati del lavoro evolvono in relazione ai cicli economici e alle specificità locali, aumentando il grado di complessità dell'impatto atteso.

Un programma presenta aspetti complessi quando fa parte di un intervento più articolato che funzionerà, secondo le attese, solo se altre componenti entrano in gioco secondo la pianificazione adottata (Funnell e Rogers 2011). In condizioni di elevata incertezza, tuttavia, i risultati dei programmi non sono affatto prevedibili. Le politiche di specializzazione intelligente finanziate dalla Commissione Europea, che approfondiremo in seguito, si propongono di infittire le interdipendenze tra settori industriali, imprese e sistemi di produzione radicati a livello territoriale. Si tratta di interventi che coinvolgono molteplici attori e si dispiegano secondo potenziali processi trasformativi che solo in parte sono intenzionalmente orientati verso la direzione di cambiamento atteso. In circostanze di elevata incertezza, in cui l'azione si svolge attraverso dinamiche di auto-organizzazione imprenditoriale, i valutatori non potranno che osservare il processo di attuazione per comprendere gli effetti c.d. emergenti (v. più oltre, il concetto di emergenza), facilitando l'apprendimento adattivo, offrendo feedback in tempo reale e provando a migliorare e/o a correggere la rotta laddove si palesano disfunzioni inattese. Il punto è se gli amministratori e gli attori coinvolti nel processo di valutazione saranno, al pari dei valutatori, sensibili ad interpretare i programmi come sistemi complessi e ad apprezzare il contributo di studi non tradizionali adeguati a gestire le dimensioni complicate e complesse degli interventi.

1.3 Bandi di valutazione da segmentare per affrontare la complessità dei programmi

L'idea che i programmi siano sistemi complessi da esplorare stenta ancora a permeare la cultura della valutazione delle organizzazioni pubbliche e private italiane⁴. Un problema ricorrente dei bandi per l'affidamento degli studi nell'ambito dei programmi operativi regionali e nazionali concerne la scarsa capacità di orientare il lavoro valutativo intorno a domande specifiche. Come già sottolineato precedentemente, le autorità di gestione richiedono la verifica degli indicatori finanziari, di realizzazione e di risultato o valutazioni tematiche o di impatto. In tali circostanze, la raccolta e l'analisi di dati di fonte primaria nonché la formulazione di un giudizio di merito rischiano di essere sacrificati in valutazioni descrittive che si limitano a contabilizzare l'elevato numero di interventi che compongono un programma operativo regionale o nazionale. Sebbene le valutazioni tematiche abbiano introdotto una necessaria segmentazione del lavoro valutativo, la prassi ricorrente sembra preferire affidamenti di notevole entità finanziaria in cui l'evaluando è l'intero programma operativo. Si impone, quindi, una riflessione su:

- la richiesta di servizi valutativi ancora poco focalizzati su quesiti di ricerca che possano orientare la raccolta di informazioni di fonte primaria e l'analisi delle informazioni quali-quantitative;
- la natura dei programmi da valutare attraverso approcci teorico-metodologici in grado di affrontare l'elevato grado di complessità e incertezza dei contesti socioeconomici odierni;
- l'entità finanziaria dei bandi onnicomprensivi che possono non essere commisurati alla mole del lavoro valutativo da realizzare.

Per migliorare sia la pratica valutativa che le scelte amministrative, occorrerà approfondire uno sforzo pedagogico che vede i valutatori dialogare e collaborare con i committenti al fine di sviluppare maggiore consapevolezza dei bisogni conoscitivi da soddisfare attraverso le valutazioni. Considerando l'importanza di indagare le reazioni eterogenee degli attori in contesti eterogenei, i bandi dovrebbero stimolare i valutatori ad esplorare la variabilità che investe i comportamenti e le aspettative, le cognizioni e le preferenze dei destinatari e delle comunità di riferimento per ricostruire le condizioni

⁴ Secondo Naldini (2013), le più frequenti debolezze delle valutazioni sono una insufficiente attenzione ai risultati, un disegno valutativo spesso inadeguato a causa di una carente teoria del programma, un utilizzo non rigoroso di metodi e tecniche, valutazioni d'impatto su interventi abbandonati e quindi di scarsa utilità, una insufficiente capacità, o volontà, di comunicare i risultati della valutazione e di proporre cambiamenti delle politiche.

di implementazione dei programmi. È un impegno da assumere non solo sul piano cognitivo, ma anche sul piano etico e politico-economico. Da un lato, occorre superare una logica che riduce la valutazione alla stima probabilistica dei benefici attesi o alla verifica degli scostamenti tra effetti osservati ed effetti desiderati. Dall'altro, occorre riconoscere il contributo all'apprendimento organizzativo che la valutazione offre alle amministrazioni capaci di formulare domande valutative pertinenti, selezionando valutatori competenti.

Quale strategia possono sviluppare i valutatori per affrontare il lavoro che li attende? Come sarà evidente più oltre, occorrerà focalizzare l'attenzione sulle domande valutative (v. Capitolo 2) al fine di individuare le questioni più salienti da indagare empiricamente, predisponendo la raccolta delle informazioni e dei dati di fonte primaria. L'analisi d'impatto ruota intorno all'identificazione dei nessi causali che collegano input, output e outcome di un intervento che coinvolge risorse materiali ma anche aspettative, prassi, routine e tanta creatività, sovente non adeguatamente riconosciuta e apprezzata.

1.4 Dal quadro logico alla ricostruzione del processo

Al fine di ricostruire le relazioni tra input, output e outcome di un programma, il quadro logico, nel suo modello standard, esamina come le attività dell'intervento conducono ai risultati attesi, come i risultati ottenuti contribuiscono al raggiungimento delle finalità del progetto e come le finalità del progetto contribuiscono al raggiungimento degli obiettivi strategici di un programma o di una politica pubblica. Il quadro logico è uno strumento analitico utile alla condivisione delle informazioni rispetto ad un determinato intervento. Tuttavia, come già sottolineato a proposito degli indicatori, non bisogna confondere il quadro logico con la valutazione della teoria di un programma (v. più oltre).

Il quadro logico esprime un'idea plausibile di come funziona un programma in determinate condizioni spazio-temporali. Esso può offrire una narrazione convincente che prefigura la performance attesa di un programma, raccontando agli stakeholder, in forma discorsiva, schematica o infografica, il problema su cui si concentrano le attività del programma e le modalità con cui sono effettuate le operazioni sulla base delle risorse impiegate in relazione ai risultati attesi. Sia i manager che i valutatori possono preparare un modello logico in qualsiasi momento del ciclo di vita di un programma e sovente rivedono il modello man mano che vengono raccolte

informazioni più circostanziate sullo stesso. Si tratta di un processo che, se condotto in maniera appropriata, contribuisce a formulare le domande valutative in maniera iterativa, come ad esempio:

- Il programma ha funzionato?
- Come e perché il programma ha funzionato?
- In quali circostanze (spazio-temporali) il programma ha funzionato?
- Il programma può funzionare in contesti differenti?

A partire dal quadro logico, si possono individuare gli input, le attività, gli output o le realizzazioni fisiche o finanziarie di un programma, nonché la performance dello stesso. Ma è attraverso la ricostruzione della teoria del programma e dell'implementazione (Weiss, 1998) che il valutatore può mettere in relazione da un punto di vista causale input, output, risultati intermedi e impatti di medio-lungo termine di un programma. È un'operazione tutt'altro che facile nelle circostanze in cui i programmi non prevedono attività standardizzabili secondo meccanismi ben consolidati e rodati nella pratica.

Nelle circostanze in cui il programma fosse standardizzabile, come ad esempio un servizio di mammografia, si potrà ricostruire il processo di erogazione, specificando dove e chi effettua il servizio e a chi viene somministrato. Eppure, anche il modo in cui l'operatore si porge nei confronti dell'utente non è un aspetto standardizzabile se è vero che l'empatia fa la differenza nella qualità della prestazione. Come già sottolineato precedentemente, nell'ambito di misure attive del mercato del lavoro, lo sviluppo di un curriculum per l'apprendimento di competenze digitali all'interno di un programma di formazione finalizzato all'occupabilità e all'inclusione sociale presenta solo alcune operazioni standardizzabili: gli istruttori possono attuare il programma secondo modalità differenziate e i valutatori non possono assumere che la documentazione formale che descrive il nuovo curriculum corrisponda alla pratica dell'insegnamento. Piuttosto, i valutatori dovranno verificare in che modo gli istruttori adottano il materiale didattico e interagiscono con i discenti in aula.

Nella maggior parte degli interventi da esaminare, i valutatori devono sforzarsi di ricostruire il contenuto dei programmi e le modalità di attuazione in quanto le operazioni realizzate nella pratica sono differenti a seconda dei contesti organizzativi e sociali e degli attori coinvolti. Ad esempio, la scelta di delegare alle agenzie di lavoro interinale l'intermediazione tra domanda e offerta di lavoro suggerisce un'elevata differenziazione nell'organizzazione, nel finanziamento e negli standard di qualità dei servizi di *job-matching* che soluzioni più regolamentate potrebbero ridurre, definendo prestazioni aderenti ai piani formulati dalle amministrazioni. La formazione

programmata ed erogata da un ente dello Stato, per esempio, seguirà procedure di rilevanza pubblica mentre un erogatore privato, anche accreditato, potrà godere di maggiori margini di flessibilità con standard di qualità talora molto variabili. Come già sottolineato precedentemente, lo strumento di intervento manifesta molteplici aspetti di volontarietà che non sono standardizzabili. Gli stessi programmi di sviluppo regionale collegati alle strategie di *Smart Specialization* per la digitalizzazione delle imprese richiedono un approfondimento puntuale sui settori industriali, sulle specializzazioni produttive emergenti nei sistemi territoriali, sulla diffusione delle tecnologie abilitanti, nonché sulle caratteristiche aziendali che spiegano l'adozione degli strumenti digitali e la velocità di transizione verso gli standard dell'Industria 4.0. In tal caso, la differenziazione degli strumenti adottati investe i settori produttivi, i contesti territoriali e le singole imprese: almeno tre diverse unità di analisi presentano un'elevata variabilità da considerare nel disegno di uno studio valutativo.

Orientando l'attenzione sul processo di attuazione, i valutatori dovranno ricostruire come il programma funziona in un'organizzazione o rispetto ad una comunità di riferimento, esaminando l'impianto logico-concettuale che sottende l'intervento, le fasi e le difficoltà incontrate durante l'implementazione e i risultati insperati o indesiderati. In maniera riflessiva e partecipata, si tratta di mettere in campo una valutazione di processo che richiede di prendere in esame gli strumenti e gli obiettivi del programma e rintracciarne l'evoluzione secondo traiettorie culturali, politico-economiche e istituzionali che mutano a seconda dei contesti sociali. Come già precedentemente sottolineato, ciò significa anche indagare i significati che gli agenti attribuiscono ad un risultato o ad un impatto, esaminando le preferenze e le motivazioni che si formano nelle interazioni sociali che hanno potere causale. Si tratta di cogliere l'endogeneità delle preferenze, considerando che le preferenze non sono date (esogenamente) ma vengono formate nel corso delle interazioni sociali e possono mutare a seconda dell'evoluzione delle relazioni tra gli attori.

Quand'anche il focus della valutazione fosse orientato alla performance e ai risultati di medio-lungo periodo, occorre dedicare del tempo alla comprensione del contenuto del programma e delle modalità attraverso cui si sono generati gli esiti dello stesso. Se i valutatori non hanno alcuna idea di come il programma si è svolto nella pratica dei contesti, potrebbero porre domande valutative inappropriate e erroneamente attribuire i risultati osservati agli interventi previsti. Se i valutatori accettassero supinamente la descrizione dei programmi desunta dalla documentazione ufficiale potrebbero incorrere nel fallimento della valutazione. Lo studio potrebbe attribuire gli effetti (o l'assenza degli effetti attesi) ad un programma che non è mai stato realizzato o che è stato attuato con un basso livello di competenza o secondo modalità talmente

differenti, che sarebbe una forzatura ricondurlo al programma formalmente descritto. Anche se l'intervento avesse prodotto un esito particolarmente favorevole, è sempre utile indagare le ragioni e i motivi per cui il programma potrebbe o dovrebbe essere finanziato nuovamente. I valutatori sono chiamati a "disvelare" le logiche reali dei programmi, non avallando le retoriche o le informazioni imprecise e inesatte che sovente alimentano la loro reputazione.

Si pensi, ad esempio, ad una scuola di imprenditorialità che introduce problematiche di etica e di impatto sociale nell'ambito della formazione curricolare, ma non tutti i docenti considerano tale scelta appropriata e non tutti la perseguono concretamente. Si pensi ad un'attività di intermediazione tra domanda e offerta di lavoro che non incoraggi adeguatamente la ricerca di lavoro tra i percettori del sussidio. Negli esempi riportati, non occorre raccogliere dati sugli effetti dei programmi attuati perché i programmi non operano secondo le finalità ufficialmente perseguite. Diventa cruciale, quindi, ricostruirne le modalità di attuazione nella pratica.

Un'importante fonte di informazioni sui programmi da valutare è il personale che lavora direttamente o indirettamente nell'ambito di un dato programma. Direttori, manager, professionisti, consulenti, assistenti possono offrire spunti molto preziosi e veicolare sapere tacito e contestuale che non è divulgato in nessun documento, indicatore o dato già codificati. I valutatori possono apprendere informazioni rilevanti semplicemente osservando come i manager e gli altri attori coinvolti nell'attuazione del programma operano e che cosa realizzano. Anche in maniera informale, effettuare sopralluoghi o incontrare il personale, gli utenti, gli amministratori, i beneficiari di un programma permette di raccogliere utili indizi sull'implementazione – una conoscenza organizzativa che non è ancora esplicitamente veicolata nei documenti ufficiali e attraverso i dati disponibili.

1.5 Il monitoraggio non è la valutazione di processo

Una procedura più sistematica di acquisizione di informazioni organizzative è il monitoraggio, che più volte richiamata in precedenza, è simile alla valutazione di processo ma sostanzialmente finalizzata ad un altro scopo. Il monitoraggio è realizzato dall'agenzia che finanzia e gestisce il programma o che supervisiona le operazioni progettuali rispetto agli standard. Secondo Rossi e Freeman (1993), il monitoraggio è teso a portare a galla:

- se e come il programma riesce a raggiungere la popolazione target;
- se le attività realizzate sono coerenti con le attività pianificate;
- se l'erogazione della spesa e di tutte le altre risorse allocate nell'ambito del programma avviene in maniera appropriata.

Apprendere le informazioni appena menzionate permette ai finanziatori e alle agenzie di attuazione di effettuare le operazioni di sorveglianza sulle attività realizzate. Il monitoraggio presuppone una serie di standard da raggiungere – ad esempio, categorie di destinatari da servire; tipologie di attività da realizzare; ammontare di risorse finanziarie da erogare – e verifica se il programma funziona secondo gli standard, misurando il progresso che questo realizza rispetto agli obiettivi. Per i valutatori, il monitoraggio è una preziosa fonte di informazioni sul programma e può orientare la valutazione di processo.⁵

La valutazione di processo ricostruisce in che modo il programma viene attuato, gli attori che prendono parte all'implementazione e la stessa logica dell'implementazione rispetto a quanto è stato pianificato. La valutazione di processo descrive i servizi, le attività, le scelte e le procedure di un programma. Fornisce un feedback tempestivo sull'attuazione del programma, sulle difficoltà incontrate e sulle modifiche necessarie. La valutazione di processo è volta a spiegare perché i risultati sono stati o meno raggiunti. Pertanto, la valutazione di processo non si realizza né esclusivamente attraverso un'analisi desk, né attraverso la verifica degli indicatori di monitoraggio, come spesso è richiesto dai bandi delle valutazioni dei programmi regionali. Come già accennato precedentemente, la richiesta di verifica degli indicatori di monitoraggio svaluta il potenziale della valutazione, quel sapere di prima mano che i valutatori generano attraverso la ricerca sul campo e l'esplorazione dei contesti.

La valutazione di processo è cruciale nella ricostruzione dei nessi causali che collegano input, output e outcome/impatti in programmi complessi, per cui occorre l'osservazione diretta dei contesti, la conoscenza della logica degli interventi, delle

⁵ Secondo i principi formulati dalla Commissione Europea, un sistema di monitoraggio ben progettato dovrebbe essere governato dai seguenti principi: completezza – vale a dire che copre tutti gli obiettivi dell'intervento; proporzionalità – vale a dire che riflette i costi di raccolta delle informazioni e l'importanza attribuita ai diversi aspetti dell'intervento; sovrapposizione minima per evitare duplicazioni e raccolte di dati non necessarie; tempestività: non tutte le prove devono essere raccolte contemporaneamente, ma dovrebbero essere disponibili quando si dà avvio ad una valutazione pianificata; accessibilità – in linea di principio, tutte le prove dovrebbero essere messe a disposizione del pubblico con informazioni chiare sulle loro specificità e limitazioni; v. Better Regulation Guidelines, pag. 40, https://commission.europa.eu/system/files/2021-11/swd2021_305_en.pdf

condizioni di attuazione e delle reazioni dei destinatari. Sviluppare una familiarità e, anzi, approfondire la conoscenza del programma è, quindi, imprescindibile per comprendere *dove* osservare e *chi* interpellare nella fase di raccolta dei dati. Si tratta di una competenza che si acquisisce con il tempo, utile a disvelare le routine, le ipocrisie, le negligenze e le infrazioni delle regole nell'agire organizzativo, ma anche le pratiche informali che spesso permettono di risolvere i problemi inattesi.

Al fine di sviluppare una valutazione di processo, Carol Weiss (1998) suggerisce di:

- *Formulare domande valutative rilevanti e incisive* e, quindi, raccogliere dati e informazioni appropriati per generare risposte adeguate.
- *Comprendere i dati* che emergono durante la fase di raccolta e analisi e che talora possono sembrare incongruenti e difficili da interpretare. Se i valutatori sono a conoscenza di come il programma funziona e come i manager interagiscono con i beneficiari e tutti gli altri attori coinvolti, essi saranno in grado d'interpretare le informazioni che ricevono e avanzare nuove piste di indagine.
- *Interpretare i risultati delle analisi* e il significato degli effetti osservati empiricamente. Se, ad esempio, il programma ha subito modifiche durante il corso di attuazione, conoscere ciò che ha influenzato o prodotto i cambiamenti intervenuti aiuta a comprendere cosa funziona o meno del programma da valutare.
- *Proporre validi suggerimenti*: i manager e i decisori politici si attendono di ricevere dalla valutazione non solo dati e informazioni di prima mano ma anche consigli e raccomandazioni su come migliorare o orientare la programmazione futura. Conoscere il programma permetterà ai valutatori di avanzare suggerimenti pratici, praticabili, accettabili politicamente e sostenibili sul piano finanziario.
- *Riportare i risultati e le conclusioni dell'indagine valutativa*: i report valutativi non dovrebbero esclusivamente fare riferimento esclusivamente ai programmi, ma specificare come essi sono stati formulati, gestiti e recepiti dal contesto politico-socio-organizzativo; ciò aiuta i valutatori ad evitare generalizzazioni non empiricamente fondate e a prendere in considerazione le caratteristiche specifiche dei contesti.
- *Effettuare meta-analisi*: ovvero una sintesi delle risultanze di studi valutativi già condotti in merito alla stessa tipologia di programma; da qualche tempo diverse procedure statistiche riescono a combinare dati e informazioni di studi diversi al fine di stimare sinteticamente gli effetti dei programmi, come ad es. i programmi di inserimento lavorativo. La meta-analisi richiede numerose informazioni sull'attuazione dei programmi, sugli esiti degli stessi e anche sulla loro classificazione nell'ambito di un settore di policy.

Come sarà evidente nel prosieguo del lavoro, la valutazione di processo è cruciale quando si adottano disegni valutativi orientati alla teoria, quando, cioè, l'analisi delle relazioni che collegano input, output e outcome dei programmi si basa sulla ricostruzione di una o più teorie di cambiamento secondo cui un intervento può produrre effetti attesi e inattesi. Ad esempio, gli interventi di adozione della fibra ottica per la realizzazione delle reti telematiche prevedono una complessa fase di infrastrutturazione. Il completamento del cantiere con lo scavo e la cablaggio delle città sono fasi imprescindibili per aumentare la connettività e diminuire il *digital divide*. La valutazione del processo permette, quindi, di comprendere come il programma di digitalizzazione è stato attuato e quali risultati intermedi ha raggiunto durante la sua implementazione. Si comprende, quindi, che frequentemente i risultati intermedi sono passaggi cruciali da verificare nella teoria dell'impatto di un programma. In tal senso, valutazione di processo e di impatto sono logicamente collegate.

2. LE DOMANDE VALUTATIVE

2.1 Collegare le domande agli obiettivi dello studio

Il primo passo di qualsiasi studio valutativo prevede la formulazione di domande valutative pertinenti e congruenti rispetto agli obiettivi conoscitivi espressi dai committenti. Occorre comprendere se lo studio intenda esaminare:

- gli effetti o l'impatto di medio-lungo periodo;
- l'attuazione del programma o la performance dell'organizzazione;
- i costi e i benefici dell'intervento secondo un criterio di efficienza allocativa;
- la sostenibilità del programma;
- i fattori chiave del programma per la trasferibilità e / o la scalabilità.

La tabella 2 riporta alcune delle domande più ricorrenti nelle valutazioni di processo e di impatto. Come già precedentemente menzionato, nella valutazione di processo, il focus è centrato sugli obiettivi del programma e sulle modalità di implementazione. Nella valutazione d'impatto, l'attenzione si rivolge al contesto al fine di intercettare il cambiamento osservabile e gli effetti attesi e inattesi del programma.

TAB. 2 - ESEMPI DI DOMANDE VALUTATIVE A SECONDA DEL TIPO DI VALUTAZIONE

Tipo di valutazione	Esempi di domande valutative
Valutazione di processo	Sono stati raggiunti gli obiettivi del programma? In che misura e quali le motivazioni del successo? Il processo di attuazione è aderente al piano prestabilito? I costi sostenuti sono commisurati ai benefici ottenuti? Quali indicatori caratterizzano il processo di attuazione? In quale orizzonte temporale emergono i risultati del programma?
Valutazione d'impatto	Qual è l'impatto del programma nel tempo (breve o medio periodo) In che misura un impatto specifico può essere attribuito all'intervento? Quali gruppi sociali sono beneficiari dell'intervento o hanno sperimentato effetti controproducenti? In che modo l'intervento ha trasformato il contesto di riferimento? L'intervento funzionerà altrove?

Le domande valutative si specificano ulteriormente a seconda dei criteri adottati per formulare giudizi sul processo o sull'impatto del programma da valutare. I criteri valutativi forniscono un quadro normativo teso ad esaminare il merito o il valore di un intervento (politica, strategia, programma, progetto o attività). Vale la pena sottolineare che le domande e i criteri valutativi definiscono il piano operativo della valutazione. Tra i criteri per la valutazione d'impatto e di processo, la Commissione europea pone l'accento su: la rilevanza, l'efficacia, l'efficienza, la coerenza e il valore aggiunto del programma, articolando le domande valutative come mostrato nella tabella 3⁶.

TAB. 3 DOMANDE E CRITERI VALUTATIVI

Criteri valutativi	Esempi di domande valutative
Rilevanza	In che misura il programma risponde a un'esigenza identificata? In che misura il programma si allinea con le priorità del governo e delle amministrazioni locali? Il programma definisce una funzione legittimamente realizzata dal governo?
Efficacia	In che misura il programma sta raggiungendo i risultati attesi, a breve, medio e lungo termine? In che misura il programma sta producendo risultati utili (output, outcome) e/o raggiungendo ciascuno dei suoi obiettivi?
Efficienza	I risultati del programma sono commisurati alle risorse impiegate? In che misura il rapporto tra input e output è tempestivo, efficiente e conforme agli standard previsti? L'intervento è stato efficiente (rispetto alle alternative)? Il programma garantisce il miglior utilizzo delle risorse?
Coerenza (interna e esterna)	Gli obiettivi del programma sono coerenti rispetto ai bisogni dei beneficiari? Gli strumenti e le attività del programma sono coerenti rispetto agli obiettivi? Gli obiettivi del programma rispondono alle esigenze del contesto? Il programma è coerente con le politiche dell'UE sul cambiamento climatico e sulla sostenibilità?
Valore aggiunto (o sussidiarietà)	Qual è il valore aggiunto del programma, ad es. in termini di occupazione, innovazione, cooperazione, ecc.? I risultati o gli impatti individuati sarebbero stati raggiunti senza l'intervento?

Fonte: Rielaborazione da [betterevaluations.org](https://www.betterevaluations.org) e [Better Regulations Toolbox](https://www.betterregulations.eu)

6 V. https://commission.europa.eu/system/files/2021-11/swd2021_305_en.pdf

La rete del Development Assistance Committee (DAC) dell'Organizzazione per la Cooperazione e lo Sviluppo Economico (OCSE) nell'ambito della valutazione dei programmi di sviluppo (EvalNet)⁷ ha definito i sei criteri della pertinenza, della coerenza, dell'efficacia, dell'efficienza, dell'impatto e della sostenibilità, da contestualizzare rispetto all'evaluando e agli attori coinvolti.

Si possono aggiungere altri criteri, come la sostenibilità ambientale o l'equità sociale o di genere, la democraticità delle decisioni o l'impatto sociale. L'adozione dei criteri da seguire non è meccanicistica ma dipende dallo scopo della valutazione e dalle esigenze degli attori e del contesto in cui si realizza uno studio. In genere, le domande valutative orientano la scelta dei criteri e la loro interpretazione anche per impostare lo studio come valutazione di processo, di risultato o di impatto.

Il primo passo per formulare domande pertinenti – che possano meglio specificare le richieste contenute nei bandi di affidamento dei servizi valutativi – richiede di stabilire una relazione con i committenti e gli attori interessati al programma e alla valutazione. Occorre comprendere (i) le ragioni che hanno motivato la richiesta di uno studio di valutazione, (ii) in che modo si prevede saranno utilizzate le risultanze dello studio e (iii) chi sono i principali utilizzatori previsti. È anche importante verificare che la valutazione possa essere condotta con le risorse disponibili (es. fondi, personale) in considerazione del modello logico che descrive le componenti del programma. La formulazione delle domande valutative è, quindi, una fase cruciale che sollecita i valutatori ad intraprendere un dialogo serrato e iterativo con i committenti al fine di delimitare e circoscrivere le questioni salienti, gli ambiti e i temi da approfondire. In considerazione delle richieste formalizzate nel bando e consultando gli attori chiave coinvolti nell'attuazione del programma e nel processo di valutazione, i valutatori cominciano, quindi, ad esplorare:

- la logica dell'intervento;
- le motivazioni alla base della valutazione;
- le questioni su cui concentrare l'attenzione;
- le conclusioni emerse nei precedenti studi valutativi condotti sullo stesso programma (es. valutazione *ex ante*, monitoraggio o valutazione intermedia);

A partire dalla consultazione dei committenti, dei responsabili del programma e dei principali stakeholder, le versioni iterative delle domande valutative dovranno tener conto dei problemi sollevati dai testimoni privilegiati consultati all'inizio della valutazione e delle aspettative degli attori con cui il valutatore si rapporta durante

⁷ [V. https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm](https://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm)

il processo di raccolta e analisi dei dati. Al fine di formulare domande pertinenti, i valutatori possono riferirsi al quadro logico, incoraggiando lo staff del programma e tutti gli attori interessati a formulare domande specifiche sulle operazioni realizzate, sull'andamento dell'implementazione e sui risultati già ottenuti. Occorre assicurarsi che i vari interessi in gioco nonché le preoccupazioni e le aspettative degli attori coinvolti siano rappresentati nel processo di definizione delle domande più rilevanti.

Un esempio di domande valutative è reperibile nell'archivio del Ministero dell'Istruzione⁸, ove le domande sono classificate in base all'efficacia della strategia, agli effetti attesi e osservati, ai gruppi target e agli stakeholder coinvolti e alle cause dei risultati osservati. Va da sé che non bisogna formulare un numero troppo elevato di domande valutative cui non si potrà dare risposta: l'esempio consultabile online presenta evidentemente un numero elevato di quesiti che rischiano di rimanere evasi. Gli esperti di valutazione suggeriscono di non porre più di 5-7 domande principali, da specificare in base al tipo di valutazione, agli utilizzatori attesi, agli usi previsti (scopi) e ai criteri di valutazione utilizzati (appropriatezza, efficacia, efficienza, ecc.). Può essere utile immaginare diversi scenari in cui prefigurare le risposte verificando se le domande investono le questioni che la valutazione intende affrontare. Si tratta di un processo interattivo e iterativo – come già sottolineato precedentemente – che permette anche di verificare i dati già disponibili e la loro attendibilità.

2.2 Le caratteristiche delle domande valutative

La valutazione è una forma di ricerca applicata ma è anche differente dalla ricerca applicata in quanto indaga questioni relative al valore, al merito e al significato delle politiche o, in termini più semplici, alla "bontà di un programma" (Scriven, 2015). La valutazione non descrive semplicemente ciò che sta accadendo ad un programma e quali effetti sono riconducibili agli interventi realizzati. Come nota Gullickson (2018), sovente una domanda di valutazione è un quesito di ricerca che richiede risposte di natura descrittiva o causale. Domande del genere forniscono informazioni utili ad affrontare questioni valutative specifiche. Tuttavia, ciò che differenzia una domanda valutativa da un quesito di ricerca è il fatto che una domanda valutativa identifica dimensioni precipuamente legate al valore dell'evaluando, come ad esempio, la qualità, il costo, la rilevanza politica, la soddisfazione, il successo, l'efficacia, l'efficienza, la sostenibilità, la scalabilità, l'esportabilità o la trasferibilità, le aree di miglioramento, i risultati e l'impatto (compresi gli effetti non intenzionali). Pertanto, occorre verificare

⁸ V. https://archivio.pubblica.istruzione.it/fondistrutturali/allegati/domande_valutazione.pdf

che le domande valutative affrontino le questioni appena richiamate che implicano ed esplicitano un giudizio e non semplicemente una ricerca su un programma. La tabella 4 evidenzia come variano le domande di ricerca dalle domande di valutazione e propone alcuni esempi su come le domande di ricerca possano essere trasformate in domande di valutazione.

TAB. 4 DIFFERENZA TRA DOMANDE DI RICERCA E DOMANDE DI VALUTAZIONE

Domande di ricerca	Domande di valutazione
Quali esigenze intende affrontare il programma da valutare? Come è stato implementato il programma?	Qual è la rilevanza dei bisogni cui il programma intende rispondere?
Quali sono i risultati?	In che misura è stato implementato il programma (fedeltà al piano, qualità dell'erogazione, costo rispetto ad erogatori simili)?
I cambiamenti sono statisticamente significativi?	Quali sono i risultati o gli impatti più importanti del programma da valutare?
Quali sono le componenti del programma da valutare, e come si collegano ai risultati attesi?	Secondo quali standard e secondo quali portatori di interesse?
	Rispetto a quali evidenze dovrebbe essere verificato l'impatto?
	Quali componenti del programma hanno contribuito maggiormente agli impatti osservati?

Fonte: Rielaborazione a partire da Gullickson (2018).

2.3 L'importanza del contesto

Nell'ambito delle domande valutative, è cruciale il riferimento al contesto in cui si è realizzato l'intervento, ove indagare i significati che gli attori attribuiscono ad un risultato o ad un impatto, vale a dire, le preferenze e le motivazioni che si formano nelle interazioni sociali che – come già sottolineato – hanno potere causale. Per contesto non si intende esclusivamente il luogo in cui il programma è stato realizzato, ma anche e soprattutto l'insieme di regole, istituzioni, culture e norme sociali che influenzano l'agire dei beneficiari e delle comunità. Occorre pertanto porre domande del tipo:

- Qual è il contesto in cui il programma ha funzionato?
- Quali sono i tratti tangibili e intangibili che caratterizzano il contesto in cui il

programma è stato attuato?

- Qual è il contesto in cui il programma ha maggiori probabilità di funzionare?
- In che modo la progettazione del programma incorpora le opportunità e i vincoli presentinel contesto?
- Quali gruppi sociali hanno maggiori probabilità di essere influenzati dal programma?

Rispondere alle domande appena esposte non sarà affatto facile nelle fasi iniziali di avvio di un programma. Diversamente quando il programma è già in una fase avanzata di attuazione o ha completato le attività previste, si potranno affrontare domande del tipo:

- Quanto tempo impiegherà il programma a generare i risultati attesi?
- Qual è la traiettoria dell’impatto?
- Qual è la teoria del programma in relazione ai risultati concretamente conseguiti?
- Cosa possiamo imparare dalle valutazioni precedenti in relazione agli esiti del programma o di una specifica tipologia di programmi?

In presenza di programmi che si attuano in una molteplicità di contesti amministrativi e socioeconomici, le domande valutative dovranno tenere conto della complessità e dell’eterogeneità dei contesti. La complessità aumenta in presenza di programmi multi-obiettivo, multi-attore e multi-settore di natura pluriennale o nelle circostanze in cui gli interventi si propongono di affrontare i problemi più impellenti del nostro tempo – quali la povertà, il cambiamento climatico, le epidemie, le disuguaglianze sociali – in contesti sociali differenziati, in cui occorre individuare i fattori che, pur non essendo sotto il diretto controllo del programma, ne possono influenzare il successo o il fallimento.

Due tipi di contesto influenzano la progettazione e la realizzazione di un programma: fattori antecedenti e fattori di mediazione (Harrell et al., 1996). I fattori antecedenti sono quelli con cui il programma viene avviato, come le caratteristiche dell’utenza, le variabili geografiche e i fattori socioeconomici e demografici. I fattori di mediazione sono le influenze che emergono durante lo svolgimento del programma, come i cambiamenti dello staff dello stesso, nuove politiche pubbliche, una recessione o una fase di espansione dell’economia nonché i programmi concorrenti e complementari da considerare in fase di valutazione. È particolarmente importante esaminare, ad esempio, come determinate caratteristiche degli utenti e dei beneficiari potrebbero influenzare il risultato di un programma. Nel caso in cui un programma di formazione fosse progettato per aumentare le competenze degli immigrati, che non hanno avuto già esperienze di lavoro nel contesto italiano, il valutatore dovrebbe porre domande

relative alle modalità di apprendimento di adulti provenienti da altre pratiche organizzative/produttive al fine di considerare e rispettare le diversità culturali dei beneficiari.

2.4 Le domande sulla causalità

Quando le domande valutative sollevano quesiti del tipo *se, come e perché* i programmi funzionano, occorre addentrarsi nei meandri della causalità al fine di attribuire gli effetti osservati agli interventi realizzati. I valutatori hanno bisogno di aprire la scatola nera dell'intervento e ricostruire in che modo, per quali motivi, per chi e in quali circostanze gli effetti osservati sono attribuibili all'evaluando. Esiste, in altri termini, un problema di inferenza causale – un problema che richiede di adottare una teoria che spieghi la causalità, interpretando, in primo luogo, come si dispiega il processo di causazione. In secondo luogo, occorre formulare un'asserzione causale che descriva, cioè, il collegamento tra le cause ipotizzate e gli effetti osservati (Stern, 2016).

Secondo la guida alla valutazione d'impatto di Stern *et al.* (2016), l'interpretazione dell'inferenza causale può aderire a diversi modelli teorici cui si fa riferimento, specialmente nella valutazione d'impatto, vale a dire:

- la *causalità sequenziale o contro-fattuale* si basa sul confronto tra due o più casi in cui si osservano le cause potenziali e gli effetti ad esse collegati. L'inferenza causale si fonda sull'esperimento o su gli approcci quasi-sperimentali, secondo un impianto positivista.
- La causalità intesa come *regolarità* dei risultati (o impatti) d'interesse si basa sulla frequenza dell'associazione tra causa ed effetto per cui l'inferenza causale si sviluppa attraverso analisi statistiche.
- La *causalità multipla* dipende dalla combinazione di più cause che producono un effetto e l'inferenza si basa sugli approcci cosiddetti configurativi, tra cui la Qualitative Comparative Analysis (QCA) e la Contribution Analysis
- La *causalità generativa* dipende dall'identificazione dei meccanismi sociali che spiegano gli effetti, e l'inferenza si basa sugli approcci basati sulla teoria secondo un'interpretazione costruttivista, critico-realista [quest'approccio viene illustrato nel disegno *theory-based*] o collegata ai principi della complessità, come concisamente illustrato nei paragrafi che seguono.

Comprendere i modelli di inferenza causale è cruciale per selezionare gli approcci valutativi che incorporano diverse teorie e visioni della causalità.

3. CAUSALITÀ E DISEGNI VALUTATIVI

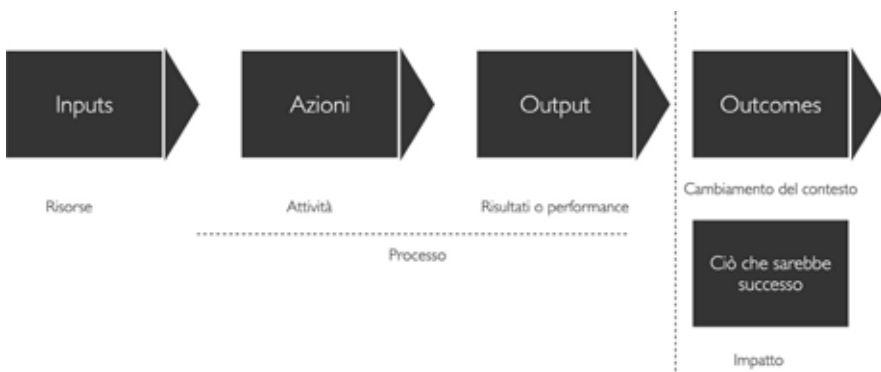
Le teorie o le interpretazioni dell'inferenza causale si associano a diversi disegni valutativi, come evidenziato nei paragrafi che seguono.

3.1 La causalità sequenzialista e approccio controfattuale

La teoria della causalità sequenzialista assume come causa del cambiamento osservato il programma, le cui attività producono gli effetti verificabili nel contesto di riferimento rispetto ai gruppi target. Al fine di ricostruire i nessi di causalità, il disegno di valutazione deve rispettare i criteri di validità interna ed esterna: attribuire o meno gli effetti all'evaluando (validità interna) e generalizzare i risultati dello studio (validità esterna).

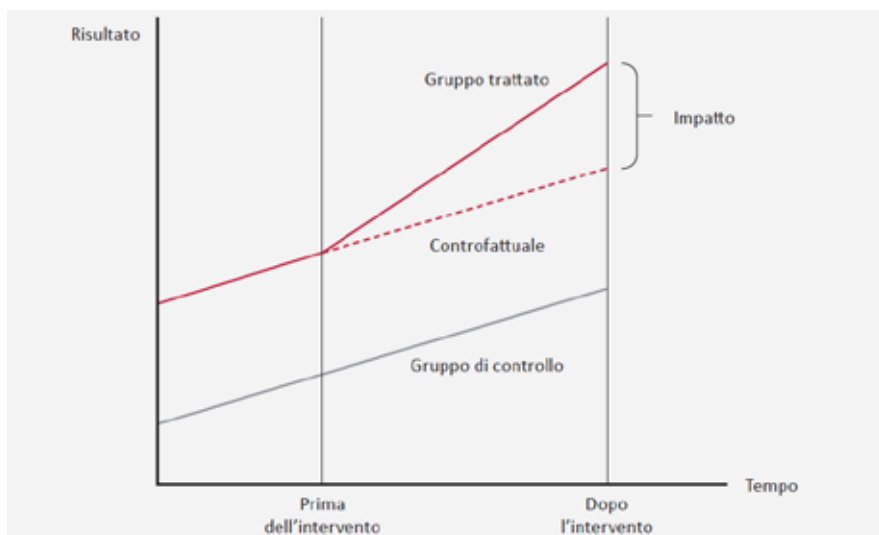
L'attribuzione causale è quel processo per cui le spiegazioni del cambiamento osservato sono collegate all'azione di un intervento. Ma per poter attribuire gli effetti osservati all'intervento in esame occorre confrontare la situazione in cui si attua un programma con la situazione ipotetica o reale in cui il programma non viene realizzato. Occorre, in altri termini, definire un controfattuale come presupposto dell'attribuzione causale che prefiguri ciò che sarebbe successo se l'intervento non fosse stato implementato. La figura 1 illustra i passaggi che si riferiscono all'attuazione del programma e isola gli outcome o risultati nell'ambito del contesto in cui il programma interagisce con altri fattori che possono trasformare le caratteristiche salienti. L'impatto del programma è dato dal confronto tra la situazione con la politica e la situazione senza politica.

FIG.1 CAUSALITÀ SEQUENZIALE



Nella valutazione d'impatto, gli esperimenti controllati e randomizzati (Randomized controlled trials - RCTs) sono tradizionalmente considerati come il disegno di valutazione metodologicamente più robusto per effettuare un'analisi causale, rispondendo alla domanda valutativa classica: il programma ha funzionato? La valutazione controfattuale esamina la differenza tra due gruppi di trattamento e di controllo per verificare la compresenza di una causa e degli effetti ipotizzati e osservati. In particolare, l'approccio sperimentale si basa su un'analisi controfattuale e attribuisce gli effetti osservati all'intervento realizzato attraverso la randomizzazione, vale a dire, l'assegnazione casuale dei soggetti ad un gruppo di trattamento (programma) e ad un gruppo di controllo (placebo), che vengono osservati nel tempo per comprenderne l'evoluzione e gli esiti (v. Figura 2).

FIG. 2 RAPPRESENTAZIONE DEL METODO CONTROFATTUALE



Fonte: Csillag Krekò, Scharle (2020)

Se i gruppi, alla fine del periodo di osservazione, presentano risultati differenti, questi ultimi potranno essere attribuiti all'intervento. Nelle circostanze in cui non sia possibile procedere all'assegnazione casuale dei soggetti ai gruppi di trattamento e controllo, i metodi di attribuzione causale rientrano nell'ampia famiglia degli approcci quasi-sperimentali che presentano una robustezza inferiore nell'analisi causale. Un

passo ulteriore che segue la verifica della validità interna (attribuzione causale) è la verifica della validità esterna che implica l'inferenza statistica, cioè la possibilità di generalizzare i risultati osservati nel campione di trattamento ad un universo più ampio, utilizzando dati statistici per trarre conclusioni riguardo la presenza o l'assenza di una relazione di causalità.

Se il controfattuale non è osservabile, occorre definire un "gruppo di controllo" credibile, che possa essere considerato simile al gruppo che partecipa al programma e che differisca unicamente per il fatto che i suoi membri non vi partecipano. Ad esempio nella valutazione controfattuale dell'impatto delle politiche sull'occupazione giovanile realizzata da INAPP, lo studio fornisce un'analisi dell'effetto causale delle misure previste su alcune variabili di risultato, come ad esempio la probabilità e la durata dello stato di disoccupazione, sulla base della stima dell'effetto della politica rispetto a quello che sarebbe stato il risultato se la politica non fosse stata introdotta, a parità di altre condizioni. Lo scenario controfattuale è costruito statisticamente al fine di procedere al confronto della situazione con e senza politica.

L'analisi controfattuale è idonea a rispondere alla domanda: L'intervento ha fatto una differenza nel contesto in cui è stato condotto? Diversamente essa risulta poco adeguata a rispondere alle domande di generalizzazione dei risultati (validità esterna), come ad esempio: l'intervento funzionerà in altri contesti? Come per le analisi statistiche, gli esperimenti sono particolarmente robusti nell'analisi di fattori causali indipendenti tra di loro, ma si rivelano meno significativi in presenza di elevata interdipendenza, soprattutto quando occorre considerare gli aspetti culturali, istituzionali, storici ed economici dei contesti di attuazione dei programmi.

3.2 Regolarità e approccio statistico

Secondo la teoria della regolarità, è necessario un elevato numero di casi da osservare per stabilire un nesso tra causa ed effetto. In altri termini, abbiamo bisogno di un campione con numerose osservazioni per intercettare correlazioni statisticamente significative che presentino cause multiple da considerare su un singolo effetto, una volta isolati i fattori di disturbo. L'inferenza causale secondo il modello delle regolarità è il processo attraverso il quale si utilizzano dati statistici (manifestazioni dei fenomeni di interesse) per trarre conclusioni riguardo la presenza o l'assenza di una relazione di causalità, sebbene la statistica convenzionale non distingua adeguatamente tra causazione e correlazione.

Numerosi modelli statistici più strutturati rispetto alle analisi convenzionali, come gli studi longitudinali e econometrici, vengono adottati al fine di registrare un numero di casi tale per cui è possibile intercettare variabili rilevanti e significative per l'analisi causale. Ad esempio, l'analisi longitudinale con stimatori fissi rappresenta una tecnica ampiamente usata dagli economisti per mitigare la distorsione imputabile all'omissione di variabili correlate con l'outcome e con la variabile causale. La logica di tale modello può essere sfruttata per eliminare fattori non osservati o non osservabili (sovente di natura contestuale) che influenzano i soggetti appartenenti ad un medesimo gruppo.

3.3 Le causalità multiple negli studi di caso

La teoria della causalità multipla adotta una prospettiva interpretativista o costruttivista secondo cui la realtà indagata è una costruzione sociale frutto dell'interpretazione degli attori coinvolti. Pertanto, le realtà costruite variano nella forma e nel contenuto a seconda degli individui, dei gruppi, delle culture, dei contesti spazio-temporali in cui si svolge la valutazione. In tal caso, i disegni valutativi più ricorrenti sono di tipo etnografico e basati su una "teoria situata" – nota con l'espressione inglese di *grounded theory*. Si può, però, adottare anche una prospettiva strutturalista che prevede disegni valutativi basati sull'analisi dei casi – come ad es. la Qualitative Comparative Analysis – QCA –, oppure basati su simulazioni e analisi di network. L'analisi causale si articola attraverso la comparazione dei fattori causali rilevanti a spiegare le relazioni d'interesse esistenti tra i casi e all'interno dei casi considerati. La generalizzazione analitica non è statistica, ma si basa sulla teoria. Ad esempio, il rapporto di valutazione d'impatto sociale del Museo Archeologico Nazionale di Napoli, relativamente alla verifica dell'impatto sociale del museo, individua alcune condizioni necessarie e sufficienti a generare un alto valore delle mostre, secondo le evidenze già emerse da altri studi di natura valutativa (Marra, 2023). La performance economica, l'affluenza di visitatori e la multidimensionalità del valore culturale sono variabili rilevanti identificate a partire dalla rassegna della letteratura sul valore dell'arte e dei beni culturali e della governance partecipativa. Il modello delle causalità multiple è indicato nelle circostanze in cui i casi analizzati mostrano livelli moderati di complessità e di interdipendenza. I modelli in esame sono meno adeguati ad individuare combinazioni di cause che presentano elevate complessità e interdipendenze.

3.4 La causalità generativa e *theory-based*

La teoria della causalità generativa considera il cambiamento frutto dell'interazione tra meccanismi causali e contesto in cui l'intervento viene realizzato. Non sono l'intervento o il programma in sé a provocare gli esiti osservati quanto la trasformazione che il contesto produce per effetto dei meccanismi innescati dal programma. L'espressione meccanismi di cambiamento si riferisce a processi che generano un risultato specifico non necessariamente generalizzabile. Come già sottolineato, i meccanismi sono leve di cambiamento che le attività del programma possono o meno azionare, il cui effetto dipende dalla reazione del contesto, dei beneficiari e degli attori che prendono parte al programma. Elster (1998) definisce i meccanismi come relazioni causali ricorrenti che possono condurre a conseguenze indeterminate.

La nozione di meccanismo di cambiamento è stata adottata anche da Pawson e Tilley (1997) nell'ambito della valutazione realistica. Gli autori sostengono che gli esiti dei programmi emergono non tanto come conseguenza degli interventi ma come effetto dei meccanismi di cambiamento che si attivano attraverso le reazioni cognitive, razionali, emotive, affettive e motivazionali degli attori. Il meccanismo non è sempre e comunque innescato dal programma su cui la valutazione tradizionale si concentra. Secondo le teorie generative, i risultati emergono come conseguenza delle reazioni dei beneficiari, delle comunità di riferimento, degli stakeholder interessati al programma che rispondono alle opportunità offerte dalle politiche pubbliche in maniera non sempre prevedibile. Se le risposte dei contesti variano a seconda dei meccanismi che entrano in gioco, la valutazione prova a ricostruire che cosa funziona, per chi e in quali circostanze, adoperando indicatori che definiscono relazioni causali specifiche, contestuali e irripetibili.

Secondo tale approccio, le politiche pubbliche sono sistemi complessi e stratificati, all'interno di un'ampia cornice politico-istituzionale che ingloba meccanismi multilivello. Una politica pubblica può attivare meccanismi simili con esiti divergenti, mentre risultati simili possono scaturire dall'azione di meccanismi diversi. Se potenzialmente le catene di causazione sono molteplici, le analisi per attribuire le cause agli effetti osservati e per generalizzare gli effetti sono intrinsecamente limitate. I disegni valutativi più utilizzati nell'ambito delle teorie generative ricostruiscono i processi causali attraverso la teoria del cambiamento. In tal senso, la causalità generativa è convincente sul piano della spiegazione dei processi causali, ma è meno efficace nella stima delle quantità o dell'ampiezza dell'impatto.

3.5 Una panoramica sui disegni valutativi non sperimentali

I disegni valutativi non sperimentali comprendono tutti gli approcci in cui le condizioni dell'esperimento, vale a dire il confronto dei risultati tra gruppo di trattamento e gruppo di controllo e l'assegnazione casuale dei soggetti ai gruppi appena menzionati, vengono a mancare. Sono inclusi in tale categoria di studi, le valutazioni che comparano i risultati prima e dopo il trattamento senza il confronto con un gruppo di controllo. Si annoverano in tale categoria di studi anche le analisi controfattuali che confrontano il gruppo di trattamento e il gruppo di controllo senza effettuare un'assegnazione casuale nella fase di campionamento. In considerazione delle diverse teorie della causalità finora presentate e dei disegni valutativi cui si è fatto cenno, è utile definire più precisamente che cosa si intende per disegno o approccio valutativo, distinguendo i disegni valutativi dai metodi di raccolta e analisi di dati utilizzati nella valutazione. La terminologia "approccio di valutazione" non ha una definizione condivisa. Ad esempio, il sito-web Better Evaluation – una delle risorse sulla valutazione più complete, tra quelle disponibili online – utilizza il termine *approccio* per riferirsi a un pacchetto integrato di opzioni, metodi o processi. Il sito illustra la definizione appena riportata riferendosi all'esperimento con gli studi randomizzati controllati per sottolineare che essi utilizzano una combinazione di campionamento casuale, gruppi di trattamento e controllo e indicatori e misure standardizzati come approccio alla valutazione di impatto di un programma. Un approccio di valutazione è, quindi, un modello analitico-concettuale, non un singolo metodo o una specifica una tecnica di analisi dei dati. L'approccio o il disegno valutativo sono un modo di strutturare l'analisi e la raccolta dei dati e si basano su una serie di principi metodologici e di specifiche dimensioni di interesse che guidano la valutazione.

Nel formulare un disegno valutativo è cruciale assumere la consapevolezza che non esiste l'approccio più robusto in assoluto. In alcune circostanze, un disegno valutativo può offrire una solida analisi della causalità; in altre circostanze, lo stesso approccio può presentare criticità che ne pregiudicano l'applicazione sia sul piano teorico che sul piano pratico. Come già accennato precedentemente, in presenza di un'elevata complessità, il disegno sperimentale non è la scelta più appropriata. Altri disegni valutativi riescono in maniera più efficace ad affrontare causalità multiple e non lineari, come ad esempio, gli approcci quasi sperimentali, gli studi statistici longitudinali, le analisi trasversali o cross-section, la valutazione basata sulla teoria, la QCA e gli approcci *developmental* più congeniali a programmi sperimentali e complessi (Patton, 2011). Il metodo sperimentale può presentare numerose problematiche anche di natura pratica, allorché può risultare complicato assegnare casualmente i soggetti ai gruppi di controllo e di trattamento e mantenerli separati durante tutta la durata

dell'osservazione. Pertanto, sebbene il metodo sperimentale e l'analisi controfattuale abbiano rappresentato uno standard di validità interna e esterna nell'ambito degli studi valutativi, quando si tratta di esplorare processi di cambiamento complessi nessun approccio è universalmente immune da deficienze metodologiche.

Prima di addentrarci nelle specificità degli approcci basati sulla teoria, la tabella 5 (che segue) ricapitola i principali disegni valutativi a seconda se il focus e/o l'esito della valutazione siano orientato su:

- *il processo*: in tal caso, il disegno più appropriato è la ricostruzione della teoria del programma e dell'implementazione (v. più oltre), ma è ricorrente anche il disegno outcome mapping quale specificazione della teoria del cambiamento secondo una prospettiva processualista;
- *le misure come indici sintetici*: in tal caso, il risultato della valutazione consiste nella quantificazione di uno o più indicatori – come nell'approccio del Social Return on Investment (SROI) – In cui lo SROI è l'indice che mette in rapporto i benefici e i costi di un intervento in termini finanziari (Marra, 2022; Yates e Marra, 2016);
- *specifici settori di intervento*: si tratta di valutazioni tematiche o settoriali, come ad esempio le analisi della povertà condotte dalle organizzazioni internazionali come la Banca mondiale o le Nazioni Unite (UNDP). Nelle analisi sulla povertà, si è recentemente assistito ad un crescente utilizzo dei metodi quantitativi per misurare l'impatto medio dei programmi di sviluppo. Tuttavia, l'interesse per la valutazione d'impatto ruota intorno a cosa funziona, perché e come e, come è noto, tale informazione non viene fornita semplicemente riportando l'effetto medio di un intervento. Di qui la dichiarazione della Rete delle Reti sulla valutazione d'impatto (NONIE) per cui l'approccio *theory-based* è sempre più utilizzato nelle valutazioni di processo e di impatto anche a livello settoriale.
- *l'analisi di performance*: in tal caso, l'attenzione è rivolta alle performance organizzative secondo modelli come il Balanced Scorecard che identifica gli elementi chiave come la visione, la strategia, i fattori di successo, le misure e gli indicatori significativi nonché i rendimenti di un'amministrazione o di un'impresa;
- *la complessità*: si tratta di valutazioni condotte in contesti e per programmi particolarmente innovativi o soggetti ad elevata incertezza che adottano approcci noti come developmental evaluation (Patton, 2011), systems thinking evaluation (Reynolds et al., 2016), complexity- sensitive evaluation (Rogers, 2008);
- *diritti umani e capacitazioni*: secondo cui l'attenzione è rivolta alla comprensione delle modalità in cui gli interventi salvaguardano i diritti umani e promuovono capacitazioni. Si tratta di una specificazione del *theory-based* in cui si annoverano

anche gli approcci femministi, sensibili al multiculturalismo, all'equità sociale e alla sostenibilità ambientale.

Volendo schematizzare secondo le più note categorie della letteratura sulla valutazione, la tabella che segue prova ad associare a ciascuna categoria identificata un insieme di disegni più comunemente adottati. Sovente, gli studi valutativi integrano più disegni o approcci per cui è possibile riscontrare, ad esempio, analisi SROI che si basano sulla teoria del cambiamento, o valutazioni della povertà che utilizzano indici sintetici, o ancora analisi di performance che si basano su valutazioni di processo o valutazioni che utilizzano la teoria dal programma e dell'implementazione per comprendere come si ampliano e si rafforzano i diritti umani e le capacitazioni. È interessante notare che le elaborazioni degli approcci *theory-based* sono state a lungo utilizzate anche nelle valutazioni sperimentali e quasi-sperimentali al fine di spiegare qualitativamente gli effetti quantificati. Ci sono addirittura studiosi come Rogers (2008) che sostengono che un approccio basato sulla teoria sia propedeutico a qualsiasi valutazione anche di tipo sperimentale. Nel capitolo che segue, l'attenzione si concentra sul disegno orientato alla teoria al fine di esplorare le dimensioni complicate e complesse dei programmi e dei contesti.

TAB. 5 DISEGNI VALUTATIVI NON SPERIMENTALI A CONFRONTO

Specificità della valutazione	Definizione	Esempi
Processo	Approcci che si focalizzano sul processo di generazione di un servizio o prodotto sociale, partendo da una revisione della logica di intervento	<i>Theory-based evaluation</i> Outcome mapping
Misure o indici sintetici	Modelli che guidano alla costruzione di un indicatore sintetico teso a misurare la creazione di valore sociale	SROI, local multiplier
Settore di intervento	Approcci sviluppati per settori specifici che focalizzano le dimensioni di valutazione sul settore di riferimento	Poverty and Social Impact Assessment (PSIA) nell'ambito della Banca mondiale, Health Impact Assessment (HIA) nell'ambito dell'Organizzazione mondiale della sanità
Performance	Modelli basati su indicatori tesi a misurare la performance dei risultati di un'organizzazione	Balanced scorecard
Complessità non-lineare	Approcci sviluppati in contesti caratterizzati da elevata incertezza per interventi sperimentali	Developmental Evaluation
Diritti umani o sviluppo umano	Metodi che considerano gli interventi di sviluppo come processi di ampliamento delle capacità e di salvaguardia dei diritti umani	Capability approach Human Rights Approach

4. LA VALUTAZIONE ORIENTATA ALLA TEORIA

La valutazione orientata alla teoria, altrimenti nota anche come teoria del programma (Bickman, 1990), logica del programma (Funnell, 1997), valutazione basata sulla teoria o teoria del cambiamento (Weiss, 1995, 1998), valutazione guidata dalla teoria (Chen, 1997), teoria dell'azione (Schorr, 1997), logica di intervento (Nagarajan e Vanheukelen, 1997), analisi del percorso di impatto (Douthwaite et al., 2004) e scienza della valutazione guidata dalla teoria del programma (Donaldson, 2005) si riferisce a una varietà di modi in cui si può sviluppare un'analisi causale che collega gli input e le attività del programma a una catena di risultati previsti o osservati; un approccio sovente utilizzato nella valutazione d'impatto.

Per questo motivo, la *theory-based evaluation* è un approccio o un modello analitico-concettuale, un modo per strutturare e intraprendere l'analisi causale e non un metodo o una tecnica di analisi. L'approccio in esame sviluppa una teoria del cambiamento che spiega come un intervento produce i risultati attesi. La teoria inizia tipicamente con una sequenza di eventi, vale a dire i risultati immediati, i risultati intermedi e i risultati finali che dovrebbero verificarsi a causa dell'intervento. Ciò è comunemente indicato come logica del programma o modello logico. La teoria del cambiamento è una narrazione di ciò che dovrebbe accadere o che è già accaduto seguendo il tradizionale modello logico del programma. Ma a differenza del modello logico, la teoria del cambiamento delinea i meccanismi, i presupposti, i rischi e le risposte del contesto che possono facilitare o impedire ai risultati di manifestarsi. Una valutazione basata sulla teoria prova, quindi, a verificare empiricamente la presunta concatenazione di cause che collegano input, output, outcome e impatti, analizzando ciascun nesso causale ipotizzato.

In particolare, l'analisi si svolge a partire da una serie di ipotesi di una catena causale che sottende sia l'impianto concettuale su cui si basa il programma sia le premesse su cui si basa l'implementazione. L'analisi si articola secondo i passaggi che seguono:

- Ricostruzione della catena causale [attraverso la teoria del programma e la teoria dell'implementazione] come specificato nei paragrafi 4.1 e 4.2 che seguono;
- Comprensione del contesto;
- Analisi dell'eterogeneità dei risultati e dei contesti;
- Svolgimento di un ragionamento controfattuale credibile rispetto all'ipotetica

- situazione in assenza del programma;
- Rigorosa analisi fattuale;
- Uso di metodi misti.

4.1 Ricostruzione della teoria del programma

Ricostruire la teoria del programma è un'operazione particolarmente complessa che, come già sottolineato, non si riduce alla mera compilazione di un quadro logico. Piuttosto, con la teoria del programma, i valutatori mettono a fuoco la catena causale che collega gli input ai risultati e agli impatti. L'analisi dà corpo ad una teoria che spiega come l'intervento dovrebbe generare l'impatto previsto. La teoria che i valutatori ricostruiscono potrebbe essere già incorporata in un quadro o modello logico ma – come già sottolineato precedentemente – quest'ultimo in genere non esplicita le ipotesi su cui si basa un intervento. Diversamente, verificare le ipotesi è fondamentale in una valutazione basata sulla teoria.

Al fine di ricostruire le premesse teoriche su cui si basa il programma, conviene focalizzare l'attenzione su ciò che il programma intende realizzare. È ragionevole iniziare con gli obiettivi ufficialmente dichiarati, che esprimono le intenzioni esplicitate pubblicamente dai decisori o dai finanziatori del programma. A questo stadio, è utile reperire i documenti che contengono le descrizioni degli obiettivi. Nelle circostanze in cui gli obiettivi sono confusi o non adeguatamente messi a fuoco, i valutatori consultano i manager o lo staff del programma per venire a capo delle finalità delle iniziative da valutare. Quando la documentazione ufficiale dei programmi contiene dichiarazioni particolarmente ambiziose, nebulose o vaghe – come ad esempio aumentare l'occupazione o migliorare la qualità della vita; ampliare le opportunità di lavoro; rafforzare i processi di diffusione della conoscenza – i valutatori sono chiamati ad un lavoro di approfondimento. Le affermazioni altisonanti e alquanto astratte dei manager o dei decisori, variamente contenute nei documenti ufficiali, veicolano informazioni preziose.

Come suggerisce Carol Weiss, i manager e tutti gli altri attori direttamente impegnati nella gestione operativa sviluppano un approccio intuitivo più che analitico alla programmazione. Si concentrano sulle attività da realizzare piuttosto che sugli obiettivi da articolare coerentemente in relazione ai bisogni del contesto. L'ambiguità che pervade gli obiettivi dei programmi permette di dissimulare eventuali divergenze politiche: i conflitti rimangono sotto traccia, se l'intento è quello di migliorare le condizioni di vita di gruppi sociali più o meno ampi. Talvolta la divergenza nelle

intenzioni dei programmi opera a livelli differenti della gerarchia amministrativa. Sovente accade che i vertici dell'organizzazione hanno delle finalità che il frontline non condivide. Ad esempio, i dirigenti rispondono ad obiettivi di razionalizzazione della spesa pubblica e, quindi, mirano a ridurre gli abusi, gli sprechi o ad evitare frodi nei programmi di assistenza sociale. Diversamente, gli operatori sul campo sono più inclini a venire incontro alle richieste dei beneficiari.

Quando gli obiettivi sono poco chiari o ambigui, occorre ricostruire le finalità dell'intervento a partire dall'osservazione dei comportamenti degli operatori e delle condotte dell'organizzazione. I valutatori dovranno sforzarsi a cogliere le divergenze e i conflitti impliciti perché possano essere risolti a beneficio del processo di attuazione del programma e della ricostruzione di una plausibile teoria del programma. Un problema simile emerge allorché le condizioni del contesto mutano o evolvono nel tempo. Le finalità ufficiali erano coerenti all'inizio dell'attuazione del programma ma nel tempo, il tipo di beneficiari, il personale, il budget, le idee su come il programma dovrebbe operare possono cambiare anche radicalmente. I valutatori sono chiamati a riconoscere l'obsolescenza delle dichiarazioni ufficiali e discuterne con lo staff, che può talora non percepire i cambiamenti intervenuti: ciò può essere fonte di tensioni che i valutatori devono affrontare.

Altre complicazioni emergono allorché il programma prova a raggiungere le finalità ufficialmente dichiarate, ma ottiene effetti inattesi o non intenzionali (negativi). Un programma di formazione a favore dell'occupabilità delle donne con bassa qualificazione permette loro di avere accesso al mondo del lavoro, ma può generare tensioni familiari per la cura dei minori e degli anziani dipendenti. Le donne impiegate ottengono un salario con cui migliorano le condizioni materiali di vita, ma possono incorrere in situazioni di povertà di tempo e i loro figli possono cadere in comportamenti a rischio. Anche un programma che sostenga i processi di ricerca e sviluppo all'interno delle PMI può, da un lato, favorire l'innovazione attraverso l'interazione con esperti di scienza e tecnologia, ma, dall'altro, esporre l'azienda agli elevati costi della ricerca e al rischio di disvelare preziosi segreti aziendali.

Quando si ricostruisce la teoria del programma, i valutatori si focalizzano sulle idee, gli assunti e le premesse concettuali che collegano logicamente le risorse ai risultati attesi. Si tratta di un processo interattivo ed iterativo che si perfeziona man mano che i valutatori comprendono il funzionamento del programma insieme allo staff e agli attori coinvolti nell'intervento e nel processo valutativo. Con il termine teoria non si intende evidentemente una teoria universalmente valida, ma un insieme di credenze che orientano l'azione. La teoria del programma non deve essere uniformemente

accettata e non deve essere sempre valida. La teoria consiste in un insieme di ipotesi che giustificano l'azione orientata al cambiamento dello status quo. Si tratta di una spiegazione dei nessi causali che mettono in relazione gli input con gli output e gli outcome del programma. È un modello plausibile – più o meno sofisticato – che propone un'interpretazione sul modo in cui un intervento possa funzionare nel contesto ove esso viene attuato.

Comprendere il contesto diventa fondamentale per caratterizzare gli effetti e l'impatto di lungo termine del programma e, quindi, per progettare la valutazione. Come già accennato precedentemente, il contesto fa riferimento all'ambiente sociale, politico ed economico in cui il programma ha luogo e a tutto ciò che può influenzare il modo in cui si svolge la catena causale innescata dal programma. L'impatto di uno stesso programma può differire in contesti diversi. Tuttavia, pensare che esistano programmi identici è una vera e propria forzatura, in quanto irriducibile è la diversità dei contesti. Comprendere il contesto aiuta, quindi, ad anticipare l'eterogeneità degli impatti e a comprendere in che misura si possono generalizzare le risultanze della valutazione. Gli effetti del programma possono variare in base al disegno dell'intervento, alle caratteristiche dei beneficiari o alle condizioni contestuali di natura politico-istituzionale, organizzativa o socio-economica. In tal senso, la ricostruzione della teoria del programma contribuisce ad identificare le molteplici eterogeneità associate ad un intervento e consente ai valutatori di incorporare tali considerazioni nella formulazione del disegno di valutazione.

Anticipare le probabili eterogeneità associate ai programmi, agli impatti e ai contesti è importante per almeno due ordini di ragioni. In primo luogo, occorre considerare le dimensioni dei campioni in cui indagare gli effetti ipotizzati; il che riflette la capacità di disaggregare le informazioni che verranno utilizzate nell'analisi: maggiore è il grado di disaggregazione e granularità dell'analisi più grande sarà la dimensione del campione richiesto per la verifica degli effetti. In secondo luogo, la ricostruzione della teoria del programma permette di identificare gruppi di beneficiari in cui esplorare impatti differenziali. Come mostrato nell'analisi che segue, ciò è particolarmente rilevante nella valutazione delle politiche attive del lavoro.

4.2 Ricostruzione della teoria dell'implementazione

Se la teoria del programma accorda una significativa importanza alle reazioni dei beneficiari rispetto alle azioni messe in campo, occorre considerare un altro tipo di ipotesi che vengono formulate durante la programmazione. Si tratta di assunti che investono l'implementazione dell'intervento al fine di verificare come realmente il programma è stato attuato. In altri termini, occorre comprendere se il programma funziona secondo il piano previsto per l'attuazione. In tali circostanze, i risultati dovrebbero materializzarsi così come sono stati pianificati.

Secondo l'ampia letteratura sui modelli di implementazione delle politiche pubbliche, si possono classificare almeno tre tipi di teorie, vale a dire: (1) teorie di implementazione top-down, (2) teorie bottom-up secondo il noto modello degli *street-level bureaucrats* di Lipsky, ovvero il modello di mappatura prospettico dal basso verso l'alto e il modello di struttura, (3) teorie ibride che includono teorie sul decentramento, il modello razionale, il modello di gestione, il modello di sviluppo dell'organizzazione, le partnership pubblico-private e i modelli di governance multi-livello. Le classificazioni cui si è appena fatto cenno possono orientare i valutatori nella ricostruzione della specifica teoria dell'implementazione del programma da valutare, confrontando logiche organizzative differenti, che non esauriscono la pluralità delle condizioni e delle forme osservabili nella realtà dei programmi e delle organizzazioni.

Al fine di procedere alla ricostruzione della teoria dell'implementazione, occorre procurarsi i piani e i crono-programmi nonché gli organigrammi che illustrano chi fa cosa, come e quando. Ma al fine di verificare la bontà della teoria, occorre comprendere le premesse su cui si basa l'attuazione. Per esempio, in un programma di formazione a favore dell'occupabilità dei giovani e dei soggetti svantaggiati, le attività del programma prevedono l'insegnamento di competenze e abilità utili al lavoro. Le premesse su cui si articola l'implementazione del programma potrebbero essere:

- la pubblicizzazione del programma attraverso forme di comunicazione pubblica,
- la selezione dei partecipanti alla formazione che soddisfano i criteri di eleggibilità,
- la formazione dei partecipanti in aula e in azienda,
- l'erogazione di un compenso ai partecipanti,
- il counselling,
- l'incontro tra domanda e offerta di lavoro.

Se il programma viene attuato seguendo tutti i passaggi appena esposti, l'aspettativa è che i partecipanti avranno buone chance di trovare un'occupazione. Se invece il programma dovesse incontrare ostacoli di qualsiasi sorta durante l'attuazione,

le reazioni attese da parte dei beneficiari potrebbero non materializzarsi. Il programma potrebbe fallire non a causa di un vizio nell'impostazione logica (teoria del programma), ma a causa delle difficoltà impreviste incontrate nell'attuazione (teoria dell'implementazione).

La verifica della teoria dell'implementazione è un'informazione preziosa a vantaggio dei pianificatori che potranno correggere le soluzioni pratiche adottate che si sono rivelate inadeguate. Diversamente, se le problematiche riscontrate investono l'impianto concettuale dell'intervento, i pianificatori dovranno nuovamente mettere mano alla teoria del programma. Si pensi ad esempio ad un programma di valorizzazione di un bene culturale come un museo che prevede una serie di partnership per realizzare mostre su un tema concordato con gli attori del territorio. Se le attività da co-organizzare dovessero venir meno per disfunzioni legate a uno o più degli attori coinvolti, occorrerebbe verificare se gli ostacoli incontrati nell'attuazione si possono caratterizzare come problemi risolvibili o se, invece, manifestano vizi nell'impianto del programma: l'idea di realizzare l'azione culturale sulla cooperazione tra le istituzioni del territorio – non riscontrata sul piano empirico – potrebbe caratterizzare una teoria del programma poco plausibile e praticabile.

4.3 La teoria del cambiamento

Integrando la teoria del programma con la teoria dell'implementazione, i valutatori giungono alla ricostruzione della teoria del cambiamento, basata su:

- un modello logico/catena dei risultati;
- le ipotesi, i rischi e, in alcuni casi, i meccanismi associati a ciascun anello della catena di causalità;
- i fattori esterni che possono influenzare i risultati attesi; e
- qualsiasi evidenza empirica a supporto delle ipotesi, dei rischi e dei fattori esterni.

Le ipotesi sono eventi o condizioni assunte che devono verificarsi affinché si generi il nesso causale. I rischi sono influenze o eventi al di fuori dell'intervento che possono inibire il verificarsi del nesso causale. I meccanismi sono i processi causali che consentono al programma di produrre risultati. I fattori esterni sono circostanze al di fuori del controllo del programma, come il contesto sociale, politico o economico, che possono influenzare la capacità del programma di raggiungere un risultato previsto. Come già accennato, al fine di verificare la validità o la congruenza della teoria del cambiamento, il valutatore esaminerà:

- la teoria del programma, osservando se le reazioni all'intervento sono in linea rispetto alle assunzioni ipotizzate;

- la teoria dell'implementazione, osservando se le attività del programma sono realizzate secondo il piano dell'attuazione e che le stesse premesse su cui si basa l'implementazione sono logicamente fondate e verificate nella realtà.

Le due teorie diventano strettamente connesse man mano che il programma evolve. Ogni stadio di attuazione dell'intervento presuppone una risposta attesa da parte dei beneficiari e le risposte degli stessi condizionano i passaggi successivi del programma. Se l'agenzia responsabile dell'attuazione del programma non completa tutte le attività previste, i beneficiari non potranno reagire alle opportunità offerte dal programma e pertanto il programma non realizzerà i progressi attesi rispetto agli obiettivi. D'altronde se i beneficiari non sono interessati o motivati a partecipare, il programma non riuscirà ad attivare i meccanismi causali ipotizzati.

4.4 La valutazione realistica: una specifica forma di *theory-based* evaluation

La valutazione realistica è una forma di valutazione basata sulla teoria sviluppata da Pawson e Tilley (1997). Secondo i due eminenti studiosi, il funzionamento degli interventi dipende dai meccanismi sociali che entrano in gioco in un contesto. Per Pawson e Tilley, il risultato emerge dall'interazione dei meccanismi con il contesto di riferimento secondo la configurazione nota con l'acronimo CMR, vale a dire:

Meccanismo + Contesto = Risultato

Come già accennato precedentemente, i meccanismi sono leve di cambiamento che un intervento può o meno attivare. In un programma di lotta contro il fumo, ad esempio, i meccanismi potrebbero includere la pressione dei pari a smettere o non smettere di fumare, il timore dei rischi per la salute o le considerazioni economiche per il costo delle sigarette. I meccanismi sono reazioni di natura cognitiva, emotiva/motivazionale o razionale responsabili del cambiamento comportamentale osservabile negli individui e nei contesti sociali più ampi. Per i valutatori realisti, le domande valutative chiave sono: cosa funziona? Per chi? In quali circostanze? Come? I valutatori realisti sono meno preoccupati di verificare i risultati attesi a livello macro e molto più interessati ad esplorare il contesto per comprendere i fattori che possono facilitare o ostacolare l'azione dei meccanismi.

Secondo tale approccio, i meccanismi intervengono e mediano l'erogazione delle risorse, l'implementazione delle attività del programma e la ricezione dell'intervento

con l'emergere dei risultati d'interesse. I meccanismi causali non sono le attività del programma ma la risposta che le attività del programma possono generare da parte dei destinatari e degli attori a vario titolo coinvolti nel programma e nel contesto di riferimento. Ad esempio, l'*open innovation* prova a creare piattaforme di innovatori che liberamente generano innovazioni non immediatamente commerciabili, che possono essere utilizzate dalle imprese per migliorare il proprio prodotto o processo. Il meccanismo causale responsabile del cambiamento atteso, vale a dire, l'innovazione e la trasformazione digitale non è la piattaforma o la disponibilità delle soluzioni innovative offerte dagli sviluppatori ma l'interesse da parte dei fruitori ad incorporare le soluzioni proposte nel processo produttivo. Oppure potrebbe essere la capacità da parte degli sviluppatori di adattare le soluzioni innovative alle esigenze specifiche degli utilizzatori.

Come già accennato precedentemente, le reazioni al programma possono essere di natura cognitiva/culturale, emotivo/affettiva/motivazionale, o razionale. In altri termini, il comportamento degli attori individuali e collettivi cambia in quanto:

- si acquisiscono nuove conoscenze che trasformano in profondità il *modus pensandi et operandi* degli agenti;
- la motivazione, l'aspirazione o l'attaccamento ad un ideale sollecitano una trasformazione del comportamento individuale o delle pratiche organizzative/istituzionali (Appadurai e Hirschman tunnel effect);
- le convenienze (calcolo costi-benefici) giustificano l'assunzione di nuove condotte;
- l'interazione sociale e le preferenze pro-sociali trasformano le priorità dei soggetti per cui invece dell'auto-interesse, gli agenti aumentano il proprio benessere quando gli altri soggetti della comunità aumentano il loro benessere.

I meccanismi appena richiamati operano singolarmente ma molto più frequentemente attraverso mix differenziati al loro interno. I mix di meccanismi hanno potere causale e influenzano l'agire dei soggetti individuali e collettivi, dando luogo a cambiamenti di scala variabile. Una valutazione basata sui meccanismi prova, quindi, ad intercettare e ad isolare le leve di cambiamento per comprenderne l'incisività nei contesti di riferimento a diversi livelli o unità di analisi. La portata del cambiamento che il meccanismo o il mix di meccanismi possono generare dipende dal contesto e dai fattori di freno o dalle barriere che potrebbero depotenziare o annullare l'azione dei meccanismi.

La valutazione realista formula e esplora le ipotesi che investono i meccanismi responsabili di incidere sui risultati di cambiamento osservati nei contesti. L'approccio realista si concentra sulla teoria dell'intervento, sui suoi presupposti comportamentali e sulle condizioni che sostengono l'intervento. L'attenzione si focalizza sulle configurazioni contesto-meccanismo-risultato, che mostrano come gli interventi funzionano o dovrebbero funzionare rispetto a quali popolazioni e in quali condizioni. Si tratta di mettere a fuoco mini-teorie del cambiamento o specifici collegamenti in una teoria più generale del cambiamento associata ad un intervento. Ciascuna configurazione contesto-meccanismo-risultato è oggetto di verifica rispetto alle prove empiriche disponibili. Ciascuna configurazione CMR fornisce la base per una valutazione orientata dalla teoria in cui i meccanismi sostanziano un impianto concettuale da verificare nella teoria del programma. L'interazione dei meccanismi con il contesto suggerisce i passaggi chiave da verificare nella teoria dell'implementazione - secondo i presupposti della *theory-based evaluation*.

4.5 Limiti e vantaggi associati al *theory-based*

Una critica comunemente avanzata all'approccio della teoria del programma consiste nel fatto che essa si presenta sovente come una teoria lineare, sia nel senso di unidirezionale sia nel senso di deterministico. In verità, ciò non è corretto sebbene sia verosimile che i decisori o manager dei programmi possano esprimere idee anche abbastanza semplici su come gli input di un intervento si colleghino alle attività, agli output ai risultati e agli impatti dei programmi. La valutazione basata sulla teoria ha proprio l'obiettivo di esplorare e verificare le intuizioni che guidano le azioni dello staff del programma al fine di dimostrare che i risultati osservati siano proprio il risultato delle attività del progetto e non viceversa. La causalità inversa, o bidirezionale, è al centro dei dibattiti sulla valutazione dell'impatto: l'auto-selezione, ad esempio, è un pregiudizio legato al posizionamento dei beneficiari rispetto al programma per cui le variabili di risultato influiscono su chi partecipa, piuttosto che il contrario. Ad esempio, è altamente probabile che le comunità, che già presentano elevati livelli di capitale sociale, richiedano fondi da destinare a programmi di sviluppo a favore della comunità. I programmi in esame hanno l'obiettivo di costruire capitale sociale, ma semplicemente osservando gli effetti *ex post* è più probabile che gli impatti riflettano il livello di capitale sociale già presente nelle comunità di intervento.

Una critica più incisiva riguarda la staticità dell'approccio rispetto agli aggiustamenti, agli adattamenti e all'evoluzione che gli interventi subiscono nei contesti reali. L'evoluzione dei programmi può talora rendere irrilevanti i disegni valutativi adottati

precedentemente. Ad esempio, se per venire incontro alle difficoltà negli spostamenti incontrate dai partecipanti ad un corso di formazione, i manager erogano loro sussidi per il trasporto pubblico, la valutazione basata sulla teoria dovrebbe segnalare che il modello originario su cui si fondava il programma è superato e richiede di essere modificato per riflettere le nuove condizioni di vita e di lavoro. Occorrerà allora ricostruire i micro processi che collegano l'intervento con i risultati attesi e parzialmente osservati, attraverso una nuova analisi teorica.

La teoria del programma deve riflettere i cambiamenti intervenuti nello stesso e la valutazione deve dar conto del processo di apprendimento che ha condotto alla riformulazione del disegno dell'intervento, rispondendo, ad esempio, alla domanda: perché emergono differenze significative rispetto al piano originale del programma e come tali differenze influiscono sulle performance dell'intervento? La teoria del programma è uno strumento di valutazione dinamico durante tutto l'arco di attuazione della valutazione ed è finalizzata a facilitare l'apprendimento direttamente a partire dall'osservazione dell'esperienza. Attraverso la ricerca sul campo, si potrà affinare la teoria del programma sulla base delle informazioni che emergono durante il processo di raccolta dei dati, in maniera iterativa. È un approccio molto diverso rispetto alle metodologie statistiche che si basano su modelli statici e sovente immutabili. Nell'analisi statistica, i valutatori verificano quanto i dati riescono a spiegare del modello con il rischio di giustapporre una teoria ad una realtà in continua evoluzione.

Un'altra critica anche valida all'approccio *theory-based* è che concentrandosi sulla concatenazione causale, i valutatori finiscono per trascurare l'analisi degli effetti indesiderati. Si tratta di una potenziale debolezza che si può attenuare in due modi. In primo luogo, un'attenta applicazione della teoria del programma può identificare le possibili conseguenze indesiderate. In secondo luogo, un preliminare lavoro sul campo, compresa un'analisi partecipativa, può far emergere i risultati non intenzionali del programma, che possono così essere incorporati nella ricostruzione della catena di causazione.

Gli effetti indesiderati sono anche collegati ad un'altra questione rilevante nella ricostruzione della teoria del programma. Un valido disegno *theory-based* tiene conto delle diverse teorie che possono caratterizzare il funzionamento dei programmi, anche concorrenti o rivali. Come già accennato precedentemente, i decisori o i manager ai vertici dell'organizzazione possono esprimere una visione, che non necessariamente viene condivisa dal personale che lavora direttamente sul campo o dai beneficiari e dagli attori portatori di interesse, i quali possono serbare punti di vista anche molto differenziati rispetto al programma e agli esiti desiderabili dello stesso. In tali

circostanze, occorre considerare sia le teorie del programma ufficiali, sia le teorie implicite o tacite da disvelare attraverso la valutazione. Le teorie concorrenti o rivali non sono meno incisive rispetto alla teoria ufficiale del programma nell'influenzare le modalità di implementazione.

I vantaggi dell'approccio *theory-based* sono molteplici. Un primo vantaggio nel ricostruire la teoria del cambiamento consiste nel fatto che i valutatori possono impostare la valutazione sulla base delle premesse su cui si fonda il programma, raccogliendo dati e informazioni a supporto delle ipotesi da verificare. Le domande valutative saranno, quindi, formulate al fine di indagare i meccanismi causali individuati nella teoria del cambiamento. I valutatori potranno utilizzare le teorie del programma e dell'implementazione per pianificare le operazioni di raccolta dei dati in corrispondenza di ciascun passaggio cruciale da spiegare e verificare empiricamente. In tal modo, i valutatori potranno anche spiegare le ragioni sottese alle conseguenze inattese, non pianificate e indesiderate che possono scaturire da processi causali che conducono ad eventi imprevisti, con esiti anche controproducenti.

Un altro vantaggio legato al disegno *theory-based* risiede nel fatto che la ricostruzione della teoria del programma può segnalare precocemente il grado di efficacia del programma. Non occorre attendere i risultati d'impatto di lungo termine per capire già nelle fasi intermedie dell'attuazione se il programma presenta anomalie che possono minare la capacità di incidere nel contesto e di generare il cambiamento desiderato. Se dovessero emergere ostacoli o barriere durante il processo di attuazione, la teoria del cambiamento può orientare gli interventi di rimodulazione utili a correggere la rotta. Ad esempio, se le prime reazioni dei partecipanti ad un corso di formazione presentano problematiche di apprendimento, ciò potrebbe segnalare che le basi logico-concettuali su cui il programma è stato formulato non sono state adeguatamente sviluppate e richiedono una rivisitazione.

Generare informazioni rilevanti nelle prime fasi dell'attuazione è strategico allorché il programma ha una durata estesa. Ad esempio, un programma di rigenerazione urbana impiegherà una decina d'anni o più per raggiungere gli obiettivi prefissati. Anche un programma di *Smart Specialization* potrà impiegare un tempo relativamente lungo per sviluppare processi di diversificazione economica nell'ambito di regioni con diversi gradi di sviluppo. Nei casi appena menzionati, una valutazione basata sulla teoria raccoglierà informazioni logicamente collegate alla capacità del programma di generare una trasformazione di lungo periodo. I dati, infatti, non offrono solo informazioni sui risultati intermedi – ad es. quanti progetti vengono finanziati durante la fase di scoperta imprenditoriale della strategia di specializzazione (v. più

oltre); i dati sono essi stessi i risultati intermedi logicamente connessi agli impatti di lungo termine. In altri termini, i dati raccolti sono parte delle ipotesi del processo di causazione.

La ricostruzione puntuale della teoria del cambiamento permette di comprendere se gli effetti empiricamente osservati sono attribuibili al programma, escludendo altri possibili fattori e variabili di interesse. Il modo tradizionale di valutare se il programma è responsabile degli effetti osservati consiste nel comparare unità che hanno ricevuto l'intervento con altre unità equivalenti che non lo hanno ricevuto. Per tanti programmi, però, come ad esempio i programmi di rigenerazione urbana appena menzionati, non è possibile effettuare comparazioni tra unità di analisi equivalenti. L'esistenza di quartieri simili è estremamente rara se non addirittura impossibile. Al fine di attribuire gli effetti osservati ai programmi avviati, la ricostruzione della teoria del cambiamento fornisce la prova che i risultati già dispiegatisi sono collegati al programma, in quanto scaturiscono dai meccanismi attivati dalle operazioni del programma.

In ultimo, la valutazione basata sulla teoria può essere integrata con i tradizionali approcci sperimentali e quasi sperimentali. Nella letteratura sulla valutazione, si tende a separare gli approcci che si focalizzano sul programma o sull'assenza del programma (controfattuale) come disegni alternativi (Funnell e Rogers 2011; Pawson e Tilley 1997). Nell'approccio *theory-based*, è possibile, anzi auspicabile, combinare più approcci. Diversi studiosi suggeriscono che gli esperimenti controllati randomizzati dovrebbero essere perfezionati in combinazione con metodi non sperimentali, mentre Rogers (2008), come già menzionato precedentemente, sostiene che anche quando si scelgono disegni sperimentali, occorrerebbe sempre e comunque ricostruire la teoria del programma.

4.6 L' approccio *theory-based* nelle politiche attive del lavoro

Le politiche del lavoro promuovono l'inserimento lavorativo attraverso misure di sostegno dell'occupazione dal lato dell'occupazione. Le misure di inclusione attiva mirano ad alleviare gli effetti negativi della disoccupazione sulla salute e sul benessere, provando a modificare le condizioni di vita dei disoccupati e di gruppi sociali vulnerabili (come ad es. i giovani e le donne), offrendo loro opportunità di formazione e socializzazione per rafforzare le reti sociali e migliorare l'occupabilità. Nonostante esista un'elevata variabilità tra le politiche attive del lavoro, si possono individuare almeno quattro tipologie di misure da considerare:

- *l'orientamento al lavoro*, attraverso l'analisi delle competenze rispetto alla situazione del mercato del lavoro locale e qualificazione e riqualificazione professionale;
- la *formazione di competenze e abilità tecniche* specifiche per il lavoro, ma anche alfabetizzazione e calcolo, nonché misure che migliorano le cosiddette competenze trasversali;
- i *servizi per l'impiego* come consulenza e assistenza nella ricerca di lavoro e/o servizi di tutoraggio, che possono essere integrati da tirocini e/o assistenza tecnica o finanziaria;
- gli *incentivi al lavoro autonomo, alla mobilità territoriale e all'avvio d'impresa* cui si possono associare misure passive per l'occupazione attraverso sussidi salariali, occupazione pubblica ad alta intensità di manodopera: i programmi in esame mirano a ridurre i costi del lavoro per i datori di lavoro e fornire opportunità di lavoro, infrastrutture sociali e progetti comunitari.

La Commissione europea classifica le politiche attive del lavoro secondo le categorie illustrate nel box che segue. L'ambito della banca dati sulle politiche del lavoro dell'Unione Europea è limitato principalmente agli interventi mirati su gruppi target che presentano difficoltà nel mercato del lavoro. Ciò facilita la distinzione tra politiche attive del mercato del lavoro rispetto alla denominazione più ampia che fa riferimento a politiche per l'occupazione, per la protezione sociale o politiche fiscali che possono avere obiettivi simili in termini di promozione del lavoro, ma che agiscono in modo non selettivo sulla popolazione.

Servizi

- 1. Servizi del mercato del lavoro
 - 1.1. Servizi al cliente
 - 1.1.1. Servizi di informazione
 - 1.1.2. Gestione del caso individuale
 - 1.1. Altre attività dei centri per l'impiego
 - 1.1.1. Amministrazione delle misure per l'impiego
 - 1.1.2. Amministrazione del supporto delle misure per l'impiego
 - 1.1.3. Altri servizi/attività

Misure attive del mercato del lavoro

- 2. Formazione
 - 2.1. Formazione istituzionale
 - 2.2. Formazione sul posto di lavoro
- 1.2. Apprendistato
- 2.3. Sostegno speciale per l'apprendistato

3. Job rotation e job sharing (ora incluso nella categoria 4)

- 4. Incentivi all'occupazione

4.1. Incentivi all'assunzione

- 4.1.1. Permanente
 - 1.1.1. Temporaneo
- 4.2. Incentivi per il mantenimento del posto di lavoro
- 4.3. Job rotation e job sharing
 - 4.3.1. Rotazione delle mansioni
 - 4.3.2. Condivisione del lavoro

5. Lavoro protetto e assistito e riabilitazione

- 5.1. Lavoro protetto e assistito
- 5.2. Riabilitazione

6. Creazione diretta di posti di lavoro

7. Incentivi alla creazione d'impresa (start-up)

Fonte: European Commission (2018) e European Training Foundation (2022)

Secondo la Commissione europea (2018), i destinatari delle politiche attive del lavoro sono, in primo luogo, i disoccupati, vale a dire, le persone senza lavoro, disponibili al lavoro e in cerca attiva di impiego. Ci sono poi gli occupati a rischio, che sono potenziali beneficiari in quanto persone occupate ma a rischio di perdita involontaria del lavoro per motivi economici, circostanze che investono il datore di lavoro, ristrutturazione o altro⁹. I giovani, le donne, i disoccupati di lunga durata e i disoccupati con bassa qualificazione, inclusi i migranti, sono gruppi target delle politiche attive del lavoro. L'esito atteso delle politiche del lavoro è l'impiego dei destinatari delle misure e la domanda valutativa più ricorrente è se il beneficiario accresce la probabilità di trovare un lavoro come effetto del programma che riceve (Koning e Peers, 2007). Eppure gli studi disponibili evidenziano altri risultati che investono l'occupazione, il reddito, le ricadute sociali e gli effetti cosiddetti *soft*, di natura non solo occupazionale. La tabella 6 classifica le politiche attive del lavoro rispetto ai risultati attesi, agli effetti specifici da osservare (nel breve, medio e lungo termine), agli indicatori utilizzati per verificare i progressi compiuti e alle prospettive o alle teorie che sottendono gli interventi realizzati. A seconda delle teorie di riferimento, cambiano gli effetti e le metriche utilizzate per verificare gli impatti.

⁹ *Le persone inattive non fanno parte della forza lavoro in quanto non sono né occupate né disoccupate ma vorrebbero entrare nel mercato del lavoro e sono svantaggiate in qualche modo.*

TAB. 6 LE PROSPETTIVE DI ANALISI, I RISULTATI ATTESI E LE METRICHE DELLE POLITICHE ATTIVE DEL LAVORO

Prospettiva	Risultati attesi	Risultati specifici		Metriche
Occupazione	Maggiore probabilità d'impiego	Impiego come effetto dell'intervento	Effetto occupazione di breve termine	Tasso di occupazione entro 12 mesi dall'intervento
			Effetto occupazione di medio termine	Tasso di occupazione entro 12-24 mesi dall'intervento
			Effetto occupazione di lungo termine	Tasso di occupazione dopo 24 mesi dall'intervento
		Occupabilità	Maggiore capacità di trovare lavoro	Dati micro su base individuale basati sulle informazioni amministrative dei servizi per l'impiego
	Minore distanza dal mercato del lavoro o maggiore prontezza nel rispondere alle esigenze del mercato del lavoro			
	Altri effetti di impiego	Qualità del lavoro	Contenuto del lavoro. Requisiti di competenza e qualificazione. Tipologia dei contratti	Requisiti bassi, medi, elevati. Valutazioni individuali. Rapporto tra tempo determinato e indeterminato
Produttività individuale			Aumento del prodotto	Dati a livello di impresa e dati aggregati a livello nazionale
Effetti di welfare e di reddito	Lotta alla povertà	Diminuzione del tasso di povertà		Dati amministrativi
	Livelli e tassi di disoccupazione	Diminuzione del tasso di disoccupazione		
	Dimensione dell'economia informale	Riduzione del peso delle attività riconducibili all'economia informale		
Ricadute sociali	Tasso di criminalità	Riduzione del tasso di criminalità		Dati amministrativi
	Sanità pubblica	Riduzioni dei costi sanitari		
	Istruzione	Riduzione della dispersione scolastica		
Risultati soft	Benessere dei beneficiari	Aumento del benessere soggettivo		Dati micro sui beneficiari sia di natura amministrativa sia di natura osservazionale
	Auto-valutazione delle carriere e delle prospettive di impiego	Maggiore fiducia in sé stessi		
	Imprenditorialità	Aumento dell'autonomia e avvio di forme di auto-impiego o creazione di impresa		

Fonte: European Training Foundation (2022).

Sulla base di una veloce ricognizione delle valutazioni condotte in Italia nell'ambito della programmazione del Fondo Sociale Europeo si può notare che una buona parte di queste si concentrano sugli effetti occupazionali d'interventi di inserimento lavorativo dei disoccupati di lunga durata e dei giovani. Le valutazioni disponibili adottano prevalentemente un approccio controfattuale e provano a dare risposta alla domanda: il programma ha funzionato?

Le valutazioni controfattuali documentano la variazione in aumento o in diminuzione del tasso medio di transizione individuale dalla disoccupazione all'occupazione, specialmente in relazione ai tirocini e alla formazione. Rispetto agli effetti sulle diverse tipologie dei destinatari, le valutazioni controfattuali considerano effetti significativi sulle donne, i giovani e gli stranieri, ma non differenziano gli impatti rispetto ai diversi gruppi di destinatari all'interno delle categorie appena richiamate.

Eppure le ricerche più recenti osservano che l'analisi dell'impatto delle politiche attive del lavoro dovrebbe essere disaggregata per le diverse componenti dei programmi, a seconda dei contesti in cui sono attuati gli interventi e rispetto al tipo di beneficiari¹⁰. La Commissione europea (2017) sostiene che la progettazione, la copertura e il targeting delle politiche attive del mercato del lavoro nonché le modalità di attuazione incidono sull'efficacia e sull'efficienza dei programmi. Uno studio di Ecorys e dell'Institute of Labor Economics (Ecorys/IZA, 2012), ad esempio, conclude che non esiste un'unica politica che possa universalmente migliorare le prospettive di lavoro dei disoccupati. Per questo motivo, occorre esaminare le interdipendenze esistenti tra le misure attuate, in quanto gli interventi non sono isolati (Immervoll e Scarpetta, 2012). La loro efficacia dipende da condizioni di contesto che non sempre programmi possono influenzare – siano queste condizioni dipendenti da altre politiche o da fattori istituzionali e socioeconomici (BIRS/Banca mondiale, 2020). Ad esempio, l'efficacia della formazione può dipendere dalle misure di sostegno al reddito, dalle indennità erogate a favore del trasporto o per la casa. In alcuni casi, gli effetti presunti delle politiche del lavoro a livello macro (come ad es. l'aumento del reddito) potrebbero non essere raggiunte a causa di fattori di contesto come ad esempio l'elevata incidenza del lavoro povero.

Di recente, gli interventi di promozione dell'occupazione si concentrano sempre più sui percorsi in grado di rispondere alle esigenze dei destinatari; esigenze individuate tramite la profilazione e la consulenza (Nivadurupola e Esposito, 2021). Risulta allora cruciale analizzare il quadro istituzionale e il contesto socioeconomico di riferimento

¹⁰ V. Yayati et al. (2019).

per comprendere i fattori che possono incidere sull'impatto delle politiche a favore dell'occupazione. In tal caso, la ricostruzione della teoria dei programmi contribuisce ad approfondire l'analisi degli esiti delle misure rispetto a diversi gruppi di beneficiari, mettendo in evidenza perché, come e in quali circostanze gli interventi conducono a risultati più o meno soddisfacenti in termini di inserimento lavorativo ad esempio. Con un approccio *theory-based* si possono indagare gli effetti differenziali rispetto a diversi target, esplorando come e perché si osservano risultati variabili a seconda delle diverse categorie di disoccupati, come donne e giovani. Ad esempio, si possono approfondire gli effetti osservati rispetto al livello di istruzione, al tipo di formazione e alle competenze dei beneficiari. Si possono approfondire le differenze negli impatti osservati in relazione ai contesti regionali o in relazione alle specializzazioni produttive degli ecosistemi di imprese a livello locale. Si possono, inoltre, ricostruire le interdipendenze tra più programmi, individuando le complementarità che possono emergere tra formazione e tirocini, incentivi e orientamento. In tutti gli esempi appena citati, l'idea è che con la TBE si possano verificare teorie differenti che sottendono gli interventi a favore dell'occupazione, proprio in considerazione della loro non standardizzabilità.

L'analisi causale da effettuare dipende dall'insieme di ipotesi che formano una teoria che, implicitamente o esplicitamente, sottende il programma o la politica di attivazione. Il valore aggiunto di un simile approccio risiede nell'analisi della variabilità degli effetti e nell'esame dell'eterogeneità dei programmi, dei gruppi target e dei contesti. Vale la pena ribadire che gli effetti osservabili variano perché variano i contesti e perché gli stessi programmi vengono attuati in maniera differente a seconda dei contesti. L'approccio valutativo orientato alla teoria risponde, quindi, all'esigenza di indagare le ragioni alla base dei successi o dei fallimenti degli interventi che, interagendo con sistemi amministrativi e socioeconomici diversi, possono produrre esperienze di maggiore o minore partecipazione al mercato del lavoro.

Il presupposto più convincente secondo cui le politiche per l'occupazione e per il lavoro richiedono un approccio *theory-based* risiede nel fatto che si tratta di misure che prevedono l'attivazione dei beneficiari. I programmi in esame presentano una forte componente motivazionale e cognitiva, specialmente sul piano della formazione o dell'aggiornamento delle competenze. Ciò implica un'elevata differenziazione delle reazioni dei destinatari degli interventi. La logica che sottende i programmi di attivazione assume una reazione ad un tempo *needing and enabling* - citando l'espressione inglese di Niyadurupola e Esposito (2021). Il raggiungimento dell'obiettivo di partecipazione al mercato del lavoro è una responsabilità condivisa tra l'individuo e lo Stato e si concretizza in una serie di prestazioni a favore del beneficiario,

ma subordinate all'adempimento verificabile di determinati obblighi da parte dei disoccupati – ad esempio, la ricerca del lavoro, la partecipazione alla formazione e l'accettazione di offerte di lavoro.

Gli studi disponibili sulla condizione di disoccupazione evidenziano, ad esempio, che la formazione professionale ha successo nelle circostanze in cui genera tra i partecipanti la percezione del loro coinvolgimento in una condizione alternativa all'essere disoccupati. Le evidenze disponibili suggeriscono che formare competenze trasversali come, ad esempio, scrivere un curriculum e imparare a rispondere a un colloquio di lavoro, possono aumentare tra i disoccupati la sensazione di una maggiore capacità di controllo sulla propria vita. Alcuni studi hanno indagato i cambiamenti comportamentali dei destinatari delle misure di attivazione. Ad esempio, in uno studio sugli effetti delle politiche di attivazione sui giovani disoccupati, Niyadurupola e Esposito (2021) hanno esaminato ciò che motiva i giovani a scegliere di partecipare o meno al mercato del lavoro e quali meccanismi attitudinali e comportamentali generano una risposta di attivazione in termini di partecipazione alla formazione professionale e di ricerca del lavoro. I cambiamenti comportamentali si possono sintetizzare in sette fasi processuali di durata e intensità variabile non generalizzabile, vale a dire:

- **Disponibilità professionale:** il processo che vede il disoccupato passare dalla completa inattività alla disponibilità professionale, ovvero la volontà di partecipare alla formazione, cominciandola ricerca del lavoro con la definizione di obiettivi di carriera.
- **Auto-verifica:** una volta disponibile professionalmente, il beneficiario si prepara ad esplorare il mercato del lavoro e le proprie competenze attraverso workshop, stage e tirocini.
- **Conoscenza di sé:** il beneficiario acquisisce una maggiore consapevolezza dei propri punti di forza, debolezze, barriere, interessi, preferenze, possibilità e obiettivi di carriera personali.
- **Fiducia in sé stessi:** una migliore conoscenza di sé conduce ad una visione più promettente del futuro professionale ed una maggiore auto-efficacia e speranza riguardo al raggiungimento degli obiettivi personali.
- **Orientamento agli obiettivi:** il beneficiario focalizza l'attenzione sulle proprie decisioni e azioni verso il raggiungimento degli obiettivi di carriera prefissati e affronta attivamente le barriere personali precedentemente identificate.
- **Attività professionale:** il beneficiario dà inizio ad esperienze professionali concrete, come la ricerca attiva di un lavoro, l'avvio di un'attività lavorativa, il ritorno alla formazione o la partecipazione a programmi specifici volti alla

rimozione delle barriere personali, all'accesso all'istruzione o al lavoro.

- Perseveranza: il beneficiario è capace di continuare le attività menzionate nella fase precedente nonostante le esperienze di frustrazione, sviluppando resilienza rispetto alle esperienze di rifiuto o di insuccesso.

Gli studi sugli effetti delle misure attive del mercato del lavoro hanno sottolineato, inoltre, che il coinvolgimento dei destinatari migliora il benessere psicofisico, specialmente in condizioni di disoccupazione. Una recente meta-valutazione di 36 studi quantitativi e qualitativi condotti a partire dagli anni '90 ha documentato un impatto positivo sulla salute dei disoccupati nei paesi ad alto reddito (Puig-Barrachina et al., 2020). Tra tutte le tipologie di misure di attivazione valutate, l'assistenza offerta nella ricerca di lavoro accompagnata da una componente di sostegno psicologico ha prodotto un effetto positivo sulla salute mentale – attraverso il miglioramento della fiducia in sé stessi, dell'immagine di sé e dell'auto-efficacia¹¹. Le analisi quantitative dei programmi che offrono assistenza nella ricerca di lavoro con componenti psicologiche, hanno documentato risultati positivi in linea con la teoria di Fryer per cui si acquisisce maggiore controllo sulla propria vita e un aumento generale dell'impegno sul lavoro. Le indagini qualitative sull'assistenza nella ricerca di lavoro a favore dei giovani (in particolare svolte in Finlandia) hanno evidenziato che le misure in esame contribuiscono a creare un'organizzazione strutturata di vita, trasformando la condizione di disoccupazione in un'occasione di migliore utilizzo del tempo, maggiore consapevolezza e senso di padronanza delle proprie scelte, autonomia economica e capacità di cogliere le opportunità nel mercato del lavoro (Puig-Barrachina et al., 2020).

L'impatto delle misure di attivazione dipende anche da molteplici fattori sociali e politici, come il "modello di attivazione", i ruoli di genere, la famiglia come istituzione sociale e la cultura del lavoro. Il "modello di attivazione" si riferisce alla strategia di policy volta a ridurre la disoccupazione, nonché al corrispondente sistema di protezione sociale a favore dei disoccupati. Esistono due principali modelli di attivazione, il tipo *work-first*, d'ispirazione liberale, e il tipo universalistico altrimenti noto come *life first*. Nell'approccio liberale, il ruolo delle politiche del lavoro e delle politiche sociali

¹¹ Il risultato appena riportato è lo stesso per i diversi paesi considerati (Australia, EUA, Finlandia). Gli altri programmi hanno mostrato risultati controversi a causa dell'eterogeneità dei contesti in cui questi programmi sono stati implementati e dell'eterogeneità del contenuto dei programmi. Ad esempio, Strandh et al. (2022) mostrano che in Svezia alla fine degli anni '90, i disoccupati coinvolti in iniziative di supporto per la ricerca del lavoro riportavano condizioni di salute mentale migliore rispetto a coloro che non prendevano parte alle iniziative valutate, ma non inclusive della formazione professionale.

si limita a incentivare gli individui a cercare lavoro, fornire informazioni e servizi di intermediazione, investire in formazione professionale a breve termine e indurre i beneficiari ad essere attivamente coinvolti nella vita sociale delle comunità locali, accettando qualsiasi lavoro offerto. Diversamente il tipo universalistico di attivazione garantisce standard di vita relativamente più elevati e non impone ai disoccupati di accettare qualsiasi lavoro disponibile per scongiurare potenziali effetti di dipendenza e opportunismo. Il modello liberale è tipico dei paesi anglosassoni e quello universalistico della Scandinavia, mentre la maggior parte degli altri paesi dell'Europa continentale e meridionale presentano varianti dei due approcci (Puig- Barrachina *et al.*, 2020). Da un punto di vista valutativo, le analisi proposte suggeriscono di ricostruire il modello di attivazione cui il programma si ispira, considerando le varianti sperimentate nella pratica delle organizzazioni territoriali.

Il recente studio di Starke e Hollertz (2022), ad esempio, ha adottato un disegno valutativo realista al fine di indagare l'impianto delle misure di attivazione messe in campo da istituzioni di governo territoriale – principalmente comuni e municipalità – e da organizzazioni del terzo settore (nella figura che segue identificate come organizzazioni non governative, con l'acronimo inglese di NGOs), al fine di comprendere in che modo decisori e attuatori di diversa estrazione socio- politico-istituzionale formulano e realizzano interventi per l'occupazione a livello locale. Lo studio ha messo a fuoco le analogie e le differenze negli elementi delle catene di causazione che collegano input, output e outcome, formulando le seguenti domande valutative:

- Quali sono le componenti chiave relativamente a: le attività, la popolazione target, la durata del programma e le esperienze pregresse dei comuni e delle organizzazioni non governative?
- Quali sono le dinamiche di interazione quando le misure di attivazione sono realizzate dai comuni?
- Quali sono le dinamiche di interazione quando le misure di attivazione sono realizzate dalle organizzazioni non governative?
- Quali sono le analogie e le differenze nelle catene di causazione tra i programmi attuati dai comuni o dalle organizzazioni non governative?

Le figure 4 e 5 che seguono illustrano sinteticamente i meccanismi di cambiamento – un elemento imprescindibile della valutazione realistica – indagati negli studi in cui operano le municipalità (M1, M2 e M3) e le organizzazioni del terzo settore (NGO1, NGO2 e NGO3). È particolarmente interessante notare che in corrispondenza degli stessi obiettivi, emergono sequenze differenziate che si rifanno agli approcci *life first* o *work first* o a varianti delle due categorie. In altri termini, le teorie dei programmi

e dell'implementazione assunte dalle municipalità e dalle organizzazioni del terzo settore sono sostanzialmente differenti, benché esse siano riconducibili agli stessi programmi di attivazione. La prospettiva del *life first* è condivisa nell'ambito del terzo settore che rappresenta i partecipanti come gruppi sociali vulnerabili che richiedono un ambiente protetto, incluso il supporto dei pari. In tal caso, l'attenzione si rivolge all'empowerment dei partecipanti attraverso il sostegno alla formazione del capitale umano e sociale. Diversamente l'approccio *work first* è più comune tra le municipalità indagate ed è associato alla percezione e alla rappresentazione dei beneficiari come competenti, capaci e particolarmente motivati a partecipare al mercato del lavoro. La catena di causazione, in tal caso, riflette l'interesse da parte dei beneficiari di ottenere un lavoro retribuito al fine di acquisire indipendenza economica. Lo studio in esame fa luce, quindi, sulle variazioni nella teoria del cambiamento e ne sottolinea le differenze che attraverso l'approccio *theory-based* possono essere ulteriormente messe a fuoco, indagate empiricamente e discusse con diversi gruppi di stakeholder.

FIG. 4 SEQUENZA DEI PROCESSI DI CAMBIAMENTO COMPORTAMENTALE IN RISPOSTA ALLE MISURE ATTUATE DALLE MUNICIPALITÀ.

M1

Acquisire esperienza per gestire puntualmente i tempi di lavoro	Acquisire una conoscenza specifica collegata al posto di lavoro	Acquisire competenze linguistiche	Comprendere come i mercati del lavoro funzionano	Ottenere un impiego nel mercato del lavoro
---	---	-----------------------------------	--	--

M2

Affrontare il timore del fallimento	Accedere a diverse forme di supporto e a reti professionali	Gestire il tempo di vita e di lavoro	Essere in grado di adattarsi ad assumere incarichi di lavoro specifici	Comprendere le dinamiche e le richieste del lavoro svolto	Migliorare il rapporto con il datore di lavoro	Sviluppare capacità di iniziativa	Completare incarichi di lavoro in maniera indipendente	Ottenere reddito adeguato e un lavoro remunerato
-------------------------------------	---	--------------------------------------	--	---	--	-----------------------------------	--	--

M3

Sviluppare la speranza per una vita migliore	Comprendere le proprie esigenze e di che tipo di supporto si può aver bisogno	Aver accesso alle forme e alle strutture di supporto	Sviluppare una organizzazione di vita quotidiana	Garantire puntualità sul posto di lavoro o in un contesto di formazione	Apprendere le condizioni che caratterizzano un certo lavoro	Apprendere a gestire le relazioni interpersonali sul posto di lavoro	Padroneggiare le proprie capacità e competenze professionali
						Sentirsi coinvolto e parte della società	Ottenere un reddito adeguato (e un lavoro remunerato)

Fonte: Starke e Hollertz(2022), trad. dell'autore.

FIG. 5 SEQUENZA DEI PROCESSI DI CAMBIAMENTO COMPORTAMENTALE IN RISPOSTA ALLE MISURE ATTUATE NELL'AMBITO DEL TERZO SETTORE.

ORGANIZZAZIONE DEL TERZO SETTORE 1

Acquisire competenze professionali	Acquisire una conoscenza delle regole del lavoro per bilanciare tempi di vita e tempi dedicati alla professione	Comprendere le cause delle difficoltà incontrate	Sviluppare la capacità di aprirsi al dialogo per esternare le difficoltà incontrate sul lavoro	Escogitare strategie per affrontare le difficoltà incontrate sul lavoro	Aver fiducia nella possibilità di trovare un'impiego	Avere uno scopo e sviluppare la percezione di poter entrare nel mercato del lavoro anche mantenendo il supporto psicologico
------------------------------------	---	--	--	---	--	---

ORGANIZZAZIONE DEL TERZO SETTORE 2

Avere la capacità di organizzare una vita equilibrata	Sviluppare capacità manuali o capacità artigianali	Contenere il senso di vergogna e di inadeguatezza	Comprendere i propri punti di forza e di debolezza	Sviluppare un senso di fiducia in se se stessi e la percezione di poter recuperare e guarire	Sviluppare maggiore tolleranza e capacità di considerare se se stessi e gli altri	Sviluppare la capacità di superare le difficoltà cambiando lavoro o intraprendendo altre attività
---	--	---	--	--	---	---

ORGANIZZAZIONE DEL TERZO SETTORE 3

Sviluppare una capacità di gestione del tempo	Sviluppare una conoscenza del negozi dell'usato	Superare il senso di solitudine provando ad interagire con amici e familiari	Apprendere a relazionarsi in una comunità e a sviluppare social skills	Migliorare la gestione delle emozioni	Sviluppare una capacità introspettiva	Aumentare la stima e la fiducia in se stessi
---	---	--	--	---------------------------------------	---------------------------------------	--

Fonte: Starke e Hollertz (2022), trad. dell'autore.

Ricapitolando, l'approccio *theory-based* può integrare l'analisi controfattuale con una valutazione circostanziata dell'eterogeneità dei programmi e degli effetti delle politiche attive del mercato del lavoro. Nel caso di programmi di inserimento lavorativo e formazione al lavoro (come i tirocini), la ricostruzione della teoria dei programmi e dell'implementazione degli stessi può evidenziare le misure efficaci ad accrescere le competenze e la motivazione dei disoccupati nella ricerca del lavoro, indagare per quali tipi di disoccupati le misure sperimentate funzionano, perché esse funzionano per alcuni gruppi e non per altri, e esplorare se esistono fattori contestuali che promuovono o inibiscono la ricerca del lavoro o la formazione delle competenze richieste dal mercato del lavoro. Con il *theory-based approach* si può verificare in che

modo il programma permette ai disoccupati di trovare un posto di lavoro adeguato alle proprie competenze e conoscenze, migliorare le candidature per le posizioni per cui fanno domanda (rendendolo così più probabile la selezione), offrendo loro una visione realistica delle opportunità di lavoro o di studio o mostrando loro possibili alternative per cui essi possano aggiustare le proprie aspettative.

Con un approccio *theory-based*, si integrano le analisi di processo e di risultato, per cui ogni aspetto del programma di assistenza alla ricerca di lavoro può essere caratterizzato come misura da specificare in obiettivi più circoscritti e risultati intermedi in vista di un esito di lungo termine. Sviluppare un disegno basato sulla teoria diventa, quindi, rilevante non solo per valutare i risultati ma ancor di più per considerare le specificità dell'intervento rispetto ai contesti. Ad esempio, la formazione per la ricerca del lavoro può essere standardizzata con una chiara serie di obiettivi o risultati di apprendimento ma dovrà essere adattata alle esigenze del mercato del lavoro locale. In base alla valutazione orientata alla teoria, è possibile, quindi, pianificare e verificare in anticipo le modalità di attuazione e i risultati attesi.

Per concludere, al fine di selezionare il disegno di valutazione più appropriato, è essenziale analizzare e comprendere la natura delle misure a favore dell'occupazione. Gli studi sperimentali esaminano gli effetti delle misure in esame attraverso modelli causali lineari compatibili con disegni valutativi controfattuali. Tuttavia, gli approcci standard potrebbero non affrontare adeguatamente l'irriducibile imprevedibilità dei contesti né tanto meno considerare l'eterogeneità delle modalità di attuazione. Si tratta di considerare gli aspetti complicati o complessi delle politiche in esame che si adattano alle mutevoli circostanze dei contesti sociali, generando effetti emergenti (Funnell e Rogers 2011; v. più oltre). Alcuni studi, d'altra parte, tendono a caratterizzare le politiche attive del mercato del lavoro come interventi complessi e adattativi rispetto al problema della disoccupazione, talora, considerato come recalcitrante e ingovernabile (Rothstein 1998). Anche questa valutazione non è del tutto precisa. Occorre effettuare un'analisi del problema specifico e dell'intervento da indagare al fine di selezionare il disegno di valutazione più appropriato.

5. COMPLESSITÀ E VALUTAZIONE

5.1 I principi della scienza della complessità rilevanti per la valutazione

Nell'ambito degli approcci *theory-based*, si possono annoverare teorie che richiamano i principi e i costrutti della scienza della complessità e del pensiero sistemico – altrimenti noto in inglese come *systems thinking* – che hanno avuto crescente applicazione nel campo della valutazione. La scienza della complessità si occupa dei fenomeni in cui le interconnessioni tra elementi non sono lineari e le proprietà emergenti dalle interconnessioni sono soggette ad incertezza. L'approccio *systems thinking*, nell'ambito della scienza della complessità, ha origine da tre tradizioni di pensiero: vale a dire, la ricerca filosofica tesa a sviluppare una visione più ampia dei fenomeni di interesse (pensiero olistico), la ricerca pratica tesa a includere prospettive multiple, ciascuna riconducibile a giudizi specifici (pensiero pluralistico e pratica partecipativa) e il pensiero propositivo che si pone alla ricerca di situazioni migliorative (ricerca operativa e ricerca-azione). Nel campo della valutazione, l'adesione alle idee e ai costrutti della scienza della complessità e del pensiero sistemico significa:

- superare il riduzionismo dei modelli razionali che sottendono le politiche, allorché incorporano assunzioni causali di natura lineare e adottano disegni sperimentali di valutazione che possono limitare la comprensione dei fenomeni complessi;
- incoraggiare una comprensione olistica ma 'situata' dei fenomeni d'interesse assecondando la capacità dei valutatori di gestire processi deliberativi partecipativi in maniera democratica¹²;
- promuovere il pluralismo e l'inclusività nel senso di incorporare prospettive differenti che responsabilizzino alla pratica valutativa incoraggiando l'uso di più metodi;
- evidenziare la necessità di affrontare l'emergenza e il cambiamento dei sistemi, accogliendo l'idea di una causalità generativa aperta a considerare situazioni imprevedibili;

¹² Con la terminologia *systems thinking* ci si riferisce ad approcci che provano ad essere più olistici rispetto alle metodologie scientifiche che concentrano l'attenzione su un insieme ristretto di variabili predefinite. Benché sia impossibile per il pensiero umano realizzare analisi esaustive (ad es. Churchman 1970; Bunge 1977; Ulrich 1983), molti ricercatori e valutatori hanno sviluppato approcci sistemici per contestualizzare o integrare le premesse e le risultanze che si possono ricavare attraverso approcci più tradizionali (es. Midgley 2000).

- sottolineare la co-evoluzione dei programmi e dei contesti per la pratica della valutazione e per le pratiche oggetto di valutazione;
- individuare le problematiche più salienti nel contesto in evoluzione ed orientare l'attenzione degli attori verso problemi trascurati o emergenti, man mano che si presentano.
- dare importanza alle questioni che riguardano il valore e il merito dei programmi.
- orientare l'attenzione sui processi, in particolare quelli che facilitano partecipazione e dialogo.

5.2 Le dimensioni di complessità nei programmi e nei contesti

La complessità dei programmi e dei contesti in cui si realizzano le misure di policy ha conferito una crescente rilevanza ai disegni valutativi in grado di esaminare causalità multiple e non lineari al fine di rappresentare più adeguatamente effetti dinamici, talvolta ambigui in condizioni di elevata incertezza. Valutare la complessità non implica necessariamente nuovi strumenti e metodi di analisi, quanto piuttosto nuovi modi di pensare alle relazioni tra politiche, attori e contesti. I principi e costrutti della teoria della complessità possono orientare i valutatori in tale impresa e possono essere incorporati nei modelli logici utili a ricostruire la teoria dei programmi e le dinamiche che caratterizzano i contesti con cui i programmi interagiscono. Le questioni più rilevanti della teoria della complessità si possono sintetizzare secondo i seguenti principi brevemente presentati nel prosieguo del paragrafo e nella tabella 8:

- Sistemi complessi adattivi;
- Auto-organizzazione;
- Interdipendenze;
- Emergenza;
- Radicamento o *embeddedness*.

Secondo la teoria della complessità, qualsiasi intervento teso al cambiamento è immerso in una serie di sistemi complessi adattivi. I sistemi complessi adattivi – altrimenti noti con l'espressione inglese *complex adaptive systems* – sono caratterizzati da molteplici agenti che apprendono, si auto-organizzano e co-evolvono con l'ambiente in cui operano secondo traiettorie non lineari. Il cambiamento si genera nell'interazione tra gli attori, non necessariamente diretti da un centro o dall'esterno. L'interazione può condurre all'emergere di elementi di novità che non si aggregano nella somma delle parti originarie. Pertanto, non è possibile prevedere le trasformazioni che si possono verificare anche se le organizzazioni operano in contesti relativamente stabili. Ad

esempio, la diversificazione delle economie regionali come effetto di strategie di specializzazione produttiva è legata alle decisioni di investimento di singole imprese ma anche alle dinamiche di collaborazione e competizione che emergono all'interno dei sistemi produttivi territoriali. L'esito di siffatti processi non è né deterministico né prevedibile.

Il concetto di auto-organizzazione ci aiuta a comprendere le variazioni dei contesti locali come, cioè, gli attori si organizzano per soddisfare i bisogni e realizzare le attività che essi reputano necessarie e rilevanti, date le risorse disponibili. I processi di auto-organizzazione sono incisivi quando sono radicati e capaci di assicurare la realizzazione delle azioni più salienti per la sopravvivenza dei sistemi locali. Il concetto di auto-organizzazione fa luce, quindi, sui limiti delle strutture imposte dall'alto, come i piani, i crono-programmi di attuazione, gli organigrammi e le gerarchie formali che i programmi giustappongono talora acriticamente sulle relazioni organizzative e istituzionali locali.

I sistemi complessi adattativi possono intensificare le interdipendenze, vale a dire le strutture e le relazioni che caratterizzano l'interazione degli attori. Questi ultimi cooperano, scambiano informazioni e interpretano le proprie osservazioni, generando conseguenze reali in maniera iterativa. Le interdipendenze si sviluppano nelle reti che si costruiscono nel tempo per cui occorre comprendere in che modo i programmi riescono a potenziare le economie di rete e la scalabilità degli effetti di interesse.

TAB. 8 LE DIMENSIONI DELLA COMPLESSITÀ DA SPECIFICARE RISPETTO AI CONTESTI E AGLI INTERVENTI

Dimensioni della complessità	Complessità del contesto	Complessità del programma
Adattamento e auto-organizzazione	Comportamento delle istituzioni politiche Dinamiche delle imprese e delle associazioni di categoria Azione dei gruppi sociali	Molteplicità degli obiettivi strategici Meccanismi formali e informali di coordinamento Strumenti di politica pubblica (più o meno volontari)
Interdipendenze	Alta densità delle reti sociali, produttive e organizzative Effetti di rete dal lato della domanda e dell'offerta Complementarietà Competizione	Programmi che rinforzano economie di rete (sia dal lato dell'offerta che della domanda) Programmi che rinforzano le economie di scala Scalabilità dei programmi e degli effetti Diversificazione degli effetti
Emergenza	Polarizzazione delle posizioni politiche rispetto ai problemi da affrontare (es. cambiamento climatico) Crisi (es. pandemia, guerra, migrazioni) Vincoli strutturali allo sviluppo Molteplicità di fonti di informazione (conoscenza scientifica, expertise, esperienza degli utenti, citizen's science, ecc.)	Impatti di lungo termine poco conosciuti I programmi presentano dimensioni di ambiguità rispetto agli effetti I cambiamenti osservati sono difficili da misurare
Radicamento o embeddedness	Stratificazione sociale Agglomerazione delle attività produttive Circolazione della conoscenza Governance multi-livello	Programmi multi-attore, multi-settoriale, multi-obiettivo e multi-sito Mix di meccanismi causali Molteplici teorie del cambiamento (rivali, complementari, co-esistenti, ecc.)

Con il concetto di emergenza si fa riferimento ai processi che generano componenti nuove nell'ambito del sistema, che non corrispondono alla mera somma delle parti. Nel cambiamento emergente le nuove proprietà di un fenomeno non esistono ad un'unità di analisi inferiore e l'esito osservato non è prevedibile *ex ante*. La governance dei complessi programmi europei è sovente un processo emergente, legato alle interdipendenze che si generano tra le istituzioni, i gruppi e le reti locali, regionali, nazionali e sovranazionali. L'emergenza è radicata nelle forme in cui gli attori sociali interpretano le situazioni, dando luogo a diverse forme di cambiamento, come ad esempio, la partecipazione dei cittadini o la diversificazione produttiva delle imprese nei contesti regionali. Ciò che 'emerge' può essere in continuità con i processi già in corso, oppure radicalmente differente come nuovi sistemi o istituzioni; risultati inattesi sia positivi che negativi.

Il concetto di *embeddedness* evidenzia il ruolo delle relazioni e delle strutture o delle reti in cui si dispiegano le relazioni che generano fiducia e scoraggiano comportamenti opportunistici. La premessa è che ciascun attore sceglie le proprie azioni in considerazione delle scelte degli altri per cui le relazioni sociali e interpersonali sono importanti tanto quanto le risorse materiali esistenti in un contesto. Le esperienze pregresse possono facilitare la cooperazione nel presente e influenzare le attese future. Occorre, quindi, comprendere la stratificazione dei sistemi sociali, l'agglomerazione delle attività produttive, i canali in cui circolano le informazioni e la conoscenza e, in ultimo, ma non per importanza, occorre ricostruire le modalità di coordinamento nell'assunzione delle decisioni strategiche a più livelli del sistema. Da un punto di vista valutativo, ciò significa che i tradizionali approcci micro o bottom-up e macro o top-down non permettono di intercettare le pratiche e i processi di cambiamento che emergono nelle interazioni socioeconomiche e politico-istituzionali. Nel monitoraggio e nella valutazione delle politiche di specializzazione produttiva, la dimensione del radicamento suggerisce, quindi, di individuare i bisogni di molteplici gruppi sociali attraverso la mappatura dei principali stakeholder in relazione a specifiche aree di ricerca e innovazione al fine di coinvolgerli nella raccolta e nell'analisi dei dati.

Secondo la teoria della complessità, i meccanismi di cambiamento caratterizzano rapporti ibridi, sia di natura causale sia di natura costitutiva, cioè, delle diverse parti di un sistema rispetto al sistema nel suo complesso. La valutazione esamina ciò che il sistema produce in termini di risultato (relazioni causali), ma anche le interazioni tra i meccanismi esistenti all'interno del sistema (relazioni costitutive). Gli elementi costitutivi e causali si esprimono attraverso diverse unità di analisi e indicatori, che ne misurano la variabilità. Si tratta di misure di causalità, ma anche di rappresentazioni di parti di sistemi che co-evolvono nel tempo (Marra, 2017).

Ad esempio, nell'ambito delle politiche di innovazione, attraverso il partenariato istituzionale e socioeconomico si formulano e si attuano interventi di sviluppo imprenditoriale. Il successo di un partenariato è legato a quell'insieme di meccanismi cognitivi, economici e motivazionali che riescono a suscitare e cementare il consenso tra i partecipanti a prescindere dalla loro numerosità. Il consenso esprime una relazione causale per progettare in maniera cooperativa iniziative di sviluppo e aumentare le possibilità di intraprendere azioni coerenti con le priorità selezionate nei territori. Il consenso o l'allineamento tra gli attori è anche un elemento costitutivo degli ecosistemi di innovazione (v. Adner, 2019), e agisce come un meccanismo, capace di innescare sequenze di azioni più o meno incisive in contesti multilivello. Nella valutazione delle politiche d'innovazione, distinguere le relazioni costitutive e

causali può contribuire ad individuare le trasformazioni che intervengono a livello di imprese, sistemi territoriali, regioni o paesi in relazione a diversi interventi attuati.

5.3 La politica di *Smart Specialization*: questioni valutative

Valutare gli effetti di una politica di innovazione è tanto complesso quanto formulare una strategia di innovazione secondo i dettami della cosiddetta strategia di specializzazione intelligente – *Smart Specialization* (S3). Sviluppare una strategia S3 è indubbiamente un programma complesso che prevede l'identificazione di un numero di aree prioritarie di investimento al fine di sostenere corrispondenti attività trasformatrice. Il portafoglio di attività da supportare è gestito a livello regionale e può essere modificato man mano che si presentano nuove opportunità di trasformazione strutturale delle economie regionali. Un'attività di trasformazione ha due proprietà utili ad orientare e catalizzare l'innovazione perseguita da parte di aziende, fornitori, partner di ricerca e clienti verso la trasformazione strutturale desiderata.

La prima proprietà riguarda l'idea che tutti i progetti selezionati tramite un approccio S3 siano correlati. La correlazione è cruciale per cogliere gli effetti di complementarità e di agglomerazione derivanti dal fatto che i progetti individuati convergono verso la direzione del cambiamento atteso. La correlazione tra i progetti può essere un indizio di complementarità che si possono sviluppare tra gli interventi ma suggerisce anche la formazione di relazioni tra gli attori coinvolti nel processo di trasformazione. I progetti selezionati dovrebbero, infatti, integrarsi e rafforzarsi vicendevolmente per stimolare l'emergere di un modello duraturo di cambiamento sul piano della produzione e dello sviluppo regionale. Da un punto di vista valutativo, le complementarità tra i progetti richiedono di esplorare le interdipendenze che si sviluppano tra progetti e attori.

La seconda proprietà delle politiche in esame riguarda l'idea che tutti i progetti che fanno parte di un'azione trasformatrice possono generare ricadute delle conoscenze tecno-scientifiche utili alla produzione. Quando si diffondono le informazioni sull'esito favorevole di progetti d'investimento, altri agenti possono essere indotti a partecipare alle attività trasformatrice, influenzando non solo le scelte di natura economica, ma anche la circolazione delle conoscenze. Pertanto, i programmi di trasformazione dei sistemi economici regionali sono anche e soprattutto interventi che promuovono la diffusione della conoscenza all'interno di un'azienda e a livello territoriale. Da un punto di vista valutativo, le relazioni tra dimensioni economiche e dimensioni cognitive suggeriscono di esplorare l'emergenza dell'innovazione tecnologica in radicati processi produttivi.

La progettazione e la realizzazione di una strategia di specializzazione prevedono tre passaggi fondamentali:

- l'individuazione delle aree tematiche prioritarie;
- la traduzione delle aree prioritarie in attività trasformative;
- l'implementazione delle attività trasformative con un piano d'azione.

La traduzione delle priorità in attività trasformative è la fase più cruciale del processo, che implica la definizione di una *roadmap* da seguire attraverso un insieme di progetti e attori che convergono verso una condivisa direzione del cambiamento. Il passaggio da area prioritaria a *roadmap* di trasformazione è un processo complesso. La complessità di una strategia S3 ruota intorno alla transizione di un determinato settore industriale, con un determinato livello di tecnologia, occupazione, qualifiche, modello di business e performance, verso un livello più elevato di tecnologie, qualifiche e performance economiche – per cui occorre domandarsi perché non è stato già raggiunto il livello cui si intende giungere e quali vincoli, ostacoli e fallimenti di mercato e di coordinamento hanno impedito l'evoluzione desiderata. Al fine di rispondere ai quesiti appena esposti, occorre ricostruire analiticamente i problemi, i progetti e gli attori identificati come parte dell'attività di trasformazione che dovranno affrontare i vincoli incontrati. Ciò significa non solo realizzare investimenti in ricerca e sviluppo ma anche formare competenze, mettendo in campo nuove forme di gestione, nuove tecnologie di diffusione e beni o servizi pubblici specializzati. E per valutare tali processi di trasformazione, è del tutto evidente l'importanza della lente dell'emergenza per comprendere come avviene l'innovazione.

I progetti individuati come parte dell'attività di trasformazione seguiranno una logica di sostegno dell'ecosistema dell'innovazione – attraverso Ricerca e Sviluppo (R&S), infrastrutture di R&S specializzate, formazione scientifica, acceleratori e incubatori, ecc.- ma anche una logica di supporto degli attori ancora in ritardo rispetto agli standard dell'Industria 4.0. In tal caso, i progetti prevedono la diffusione della tecnologia, delle capacità manageriali e di innovazione, servizi specializzati rivolti alle piccole imprese, formazione professionale, ecc. I progetti sono, quindi, diversificati e costituiscono l'insieme delle attività trasformative come mix di capacità, attività e soggetti collegati ad una o più strutture esistenti – cui si possono aggiungere capacità extra-regionali – orientati verso la direzione di cambiamento prescelta. In tal caso, il pensiero sistemico può aiutare ad intercettare le interazioni tra i sistemi di R&S e i sistemi di produzione e di commercializzazione.

In particolare, la trasformazione delle strutture dell'economia regionale si articola intorno alle "cinque D" dell'approccio *Smart Specialization*, vale a dire, (i) la direzione di cambiamento, (ii) la densità relazionale, (iii) la differenziazione regionale, (iv) la scoperta imprenditoriale - in inglese *entrepreneurial discovery process* (EDV) e (v) la distribuzione delle capacità. La scoperta imprenditoriale è un processo di trasformazione in cui si individuano le caratteristiche e le proprietà dei progetti selezionati per intraprendere un percorso di specializzazione regionale, con una buona dose di intenzionalità strategica, ma anche con il riconoscimento di un'elevata incertezza, che richiede una sperimentazione decentrata e imprenditoriale, non facile da gestire da parte dei responsabili politici. Sul piano della valutazione, ciò suggerisce di puntare l'attenzione sulla capacità di adattare le strategie S3 in relazione ai processi di auto-organizzazione dei sistemi produttivi territoriali.

Un'altra finalità rilevante della S3 consiste nell'incoraggiare le regioni a costruire i propri vantaggi competitivi in base ai punti di forza reali e potenziali e in base alle opportunità esistenti e future in netta antitesi rispetto ai tradizionali processi di imitazione dei modelli vincenti o performanti delle economie più avanzate. Per farlo con successo, ciò implica una profonda conoscenza delle specificità regionali e del radicamento dei sistemi produttivi industriali e richiede altresì interventi con un elevato grado di specificità. Ad esempio, il processo d'innovazione biotecnologica richiede la fornitura di input pubblici e l'istituzione di meccanismi di coordinamento molto diversi da quelli necessari per innovare l'industria automobilistica o turistica: più specializzata e specifica è la strategia di innovazione e più sono granulari le conoscenze richieste ai governi e ai loro partner.

5.4 Le carenze delle valutazioni realizzate

Nel quadro sinteticamente delineato, il ruolo della valutazione è cruciale nel fornire informazioni tempestive perché i responsabili politici possano apprendere e reagire rispetto ai risultati emergenti di una strategia S3. Eppure, gli studi valutativi disponibili sono esigui e si limitano a verificare gli indicatori della spesa erogata, descrivendo le principali fonti di finanziamento delle strategie S3, le priorità metodologiche d'intervento e le aree di specializzazione nelle diverse filiere produttive, come l'agroalimentare, la meccanica, il sistema casa, le tecnologie marittime o la *smart health*.

Il piano della valutazione del PON Ricerca e Innovazione 2014-2020¹³, ad esempio, a partire dalla documentazione relativa ai progetti finanziati e dai dati di monitoraggio, integrati da un'indagine diretta, ha delineato una versione preliminare delle azioni di verifica da mettere in campo. Il piano si è proposto di esaminare la capacità di realizzare sistemi integrati e coerenti di ricerca- formazione-innovazione in grado di innescare un processo di crescita competitiva e sostenibile dell'economia. Il documento distingue due livelli dell'attività valutativa, vale a dire, la valutazione di processo/operativa, non tematica, funzionale a supportare la regolare attuazione del Programma, e la valutazione d'impatto (quantitativa o qualitativa) diretta a far emergere, secondo l'analisi controfattuale e/o *theory based*, gli effetti dei contributi erogati rispetto agli obiettivi del Programma, ai beneficiari target e alle ricadute esterne (efficacia esterna degli interventi). Il piano in esame si limita a contemplare gli approcci e i metodi in maniera generica senza sviluppare disegni specifici e senza entrare nel merito delle questioni più critiche da valutare.

A livello europeo, emergono analoghe carenze sia sul piano degli approcci sia sul piano delle strumentazioni. Come documentato nell'analisi di Hegyi e Prota per il Joint Research Centre (JRC) della Commissione europea nel 2021, lo stato dell'arte della valutazione delle strategie S3 segnala la necessità di definire puntualmente il disegno complessivo delle valutazioni e le specifiche operazioni di monitoraggio. L'analisi di Hegyi e Prota (2021) solleva i seguenti quesiti di ricerca:

1. In che misura i sistemi di monitoraggio e valutazione delle strategie di *Smart Specialization* alimentano un processo ciclico di apprendimento al fine di migliorare la progettazione e i meccanismi delle politiche e, inoltre, i sistemi in essere sono in grado di misurare l'impatto socio-economico-ambientale degli interventi realizzati?
2. I sistemi di monitoraggio e valutazione sono adeguati a raccogliere, organizzare e trasmettere informazioni sullo sviluppo degli interventi in modo tale da fornire, gestire e utilizzare i dati per politiche informate dalle evidenze già emerse?
3. L'approccio metodologico adottato nelle specifiche strategie di specializzazione intelligente è adeguato a valutare globalmente la politica S3?

La ricognizione delle pratiche valutative effettuate durante il periodo di programmazione 2014-2020 ha utilizzato due principali fonti d'informazione. L'indagine si basa su un questionario strutturato e una serie di studi di caso che includono regioni e territori a diversi livelli di sviluppo e con differenti assetti istituzionali. Delle 120

¹³ v. http://www.ponricerca.gov.it/media/395693/pianodivalutazione_testo-aggiornato.pdf

strategie di specializzazione analizzati dal JCR, il questionario è stato compilato da 79 autorità nazionali o regionali, da 19 paesi membri, mentre i casi di studio coprono 14 strategie regionali e 4 nazionali con un approfondimento sulle pratiche di attuazione (Hegy e Prota, 2021).

L'analisi condotta ha raccolto esperienze e percezioni sull'attuazione delle strategie e sul futuro dell'agenda politica di specializzazione intelligente: l'89% degli intervistati provengono da amministrazioni regionali (25 regioni meno sviluppate; 39 regioni più avanzate; 7 regioni in transizione), mentre l'11% da ministeri (compresa la strategia delle sei città della Finlandia). Gli studi di caso si basano su analisi dei dati secondari e interviste a funzionari e attori portatori di interesse coinvolti nella progettazione, realizzazione, monitoraggio e valutazione di strategie di specializzazione intelligente. I casi coprono 13 regioni (2 regioni meno sviluppate; 3 regioni in transizione; 8 regioni più sviluppate) e 4 paesi (Polonia, Ungheria, Slovenia e Spagna) (Hegy e Prota, 2021).

Il rapporto appena citato riporta che le capacità di apprendimento delle regioni più avanzate sono promettenti, ma ancora da definire ed integrare sono i sistemi di monitoraggio e valutazione nei contesti regionali meno innovativi. Le questioni più salienti ruotano intorno alla carente caratterizzazione dell'approccio *place-based* che richiede la definizione delle priorità regionali in settori, campi o tecnologie innovativi attraverso il processo di scoperta imprenditoriale. Benché una strategia di specializzazione debba far emergere ciò che una regione sa fare meglio in termini di dotazioni scientifiche e tecnologiche (Foray, David e Hall, 2009), le valutazioni finora condotte mostrano alcuni problemi comuni come ad esempio:

- la sindrome del vincitore, per cui prevalgono gli approcci più tecnologicamente avanzati sperimentati con i risultati più lusinghieri;
- una carente comprensione del contesto istituzionale e produttivo;
- l'imitazione, altrimenti nota come 'il copia e incolla' delle politiche di innovazione progettate nelle regioni con le migliori performance.

Le principali sfide legate al sistema di monitoraggio riguardano la mancanza di dati adeguati e tempestivi per l'elaborazione e l'aggiornamento di indicatori significativi, come pure l'assenza di connessioni coerenti tra obiettivi e indicatori. La disponibilità di dati tempestivi e affidabili sull'implementazione di una strategia di specializzazione è un prerequisito per la valutazione che non può fare a meno di un sistema integrato di monitoraggio su scala regionale e nazionale. Secondo gli studi di caso condotti da Hegyi e Prota (2019), esempi di sistemi di monitoraggio ben definiti sono quelli dell'Emilia-Romagna e della Catalogna - regioni classificate come moderate innovator

secondo lo Scoreboard dei sistemi regionali d'innovazione nel 2021. In Emilia Romagna, ad esempio, la Regione ha specificato gli "indicatori di specializzazione" e gli "indicatori di transizione" al fine di comprendere gli aspetti strutturali coinvolti nelle attività di trasformazione¹⁴. L'esperienza della Catalogna è degna di nota anche per il tentativo di monitorare il processo di scoperta imprenditoriale attraverso l'evoluzione delle principali aree settoriali e l'identificazione delle attività emergenti che utilizzano sistemi di sorveglianza delle tecnologie e dei mercati nonché un dialogo costante con le imprese e gli stakeholder nel campo della ricerca e dell'innovazione.

Sul piano della valutazione dell'implementazione delle politiche di specializzazione, la ricognizione condotta evidenzia che la formulazione degli obiettivi e della logica degli interventi è una debolezza comune a molte strategie. Nella fase di progettazione l'attenzione è posta sulla definizione delle aree prioritarie che non vengono adeguatamente articolate in obiettivi specifici al fine di mettere a fuoco la natura e la dinamica del cambiamento socio-economico atteso. L'analisi d'impatto delle strategie stenta, quindi, ad avanzare sul piano delle operazioni di verifica da predisporre ma anche e soprattutto sul piano della cultura della valutazione che è ancora improntata, sia a livello nazionale che europeo, ad approcci ancora troppo tradizionali (Hegyí e Prota, 2021).

Benché secondo il rapporto del JCR poc'anzi citato esista uno stretto collegamento tra monitoraggio e valutazione per le politiche di coesione come nel caso della Regione Toscana (Hegyí e Prota, 2021), l'avvento delle politiche S3 non pare abbia modificato le pratiche di valutazione più invalse. Le prove raccolte da Hegyí e Prota (2021) suggeriscono che non vi sia stata alcuna discontinuità in termini di approcci e metodologie adottati rispetto agli esercizi di valutazione precedentemente condotti. Il punto è che i tradizionali approcci di valutazione orientati alla verifica delle responsabilità di attuazione sono poco adeguati a sostenere le capacità di

14 *Gli indicatori di specializzazione hanno l'obiettivo di misurare il livello di specializzazione dei 5 sistemi produttivi regionali con riferimento ai 19 orientamenti tematici individuati dalla s3. Gli IS fanno parte degli Indicatori di Cambiamento, che hanno l'obiettivo di valutare la direzione dei cambiamenti in corso, la loro intensità ed il grado di coerenza rispetto alle direttrici individuate nella strategia. Gli indicatori di specializzazione sono ad esempio: start-up innovative, assegni di ricerca, contratti di ricerca (dati draft), PMI innovative (dati draft) e brevetti. Gli indicatori di transizione hanno l'obiettivo di misurare la direzione e l'intensità del cambiamento atteso nei sistemi produttivi rispetto agli obiettivi tecnologici della strategia. Gli IT fanno parte degli Indicatori di Cambiamento, quelli che hanno l'obiettivo di valutare la direzione dei cambiamenti in corso, la loro intensità ed il grado di coerenza rispetto alle direttrici individuate nella strategia. V. Esempi al link [https://www.provincia.bz.it/politica-diritto-relazioni-estere/europa/downloads/Parte_III_Monitoraggio_e_valutazione_S3\(1\).pdf](https://www.provincia.bz.it/politica-diritto-relazioni-estere/europa/downloads/Parte_III_Monitoraggio_e_valutazione_S3(1).pdf)*

apprendimento dei sistemi territoriali. Se l'obiettivo consiste nel contemperare le crescenti complessità socioeconomiche con i requisiti di efficacia e efficienza delle politiche pubbliche, la tempistica delle valutazioni dovrebbe quanto meno garantire l'avvio degli studi contemporaneamente allo sviluppo e all'implementazione dei programmi.

5.5 Come valutare una politica di innovazione?

È ormai noto che per valutare qualsiasi politica o programma, occorre formulare domande valutative pertinenti che possano investire la logica della strategia, il processo di attuazione, gli esiti e gli impatti già verificabili nei contesti regionali. Anche per la valutazione delle politiche di innovazione, le domande scaturiscono dalle esigenze cognitive dei committenti e si affinano nell'interlocuzione dei valutatori con gli stakeholder. La tabella 9 riporta le domande e gli approcci valutativi adottati finora nell'ambito degli studi già condotti in tema di specializzazione intelligente in Italia rispetto all'obiettivo strategico T01 della politica di coesione europea.

TAB. 9 SINTESI DELLE DOMANDE VALUTATIVE TRATTE DALLA RASSEGNA DEGLI STUDI DI VALUTAZIONE DELLE STRATEGIE DI SPECIALIZZAZIONE INTELLIGENTE IN ITALIA

Obiettivo specifico	Domande valutative	Approccio valutativo/Strumenti di raccolta e analisi dei dati
Incremento dell'attività di innovazione delle imprese	<p>a. Che tipo di progetti e imprese sono stati finanziati?</p> <p>b. Lo strumento dei servizi di innovazione può costituire uno strumento efficace nel favorire atteggiamenti più pro-attivi delle PMI verso la ricerca ed innovazione?</p>	Analisi desk
Aumento dell'incidenza di specializzazioni innovative in perimetri applicativi ad alta intensità di conoscenza	<p>Chi sono i soggetti che hanno beneficiato dei finanziamenti?</p> <p>Che rapporti di collaborazione si sono instaurati? Con quale interazione e con che soddisfazione?</p> <p>Che sviluppi concreti stanno avendo le attività di R&S?</p> <p>Quali effetti più generali sono prefigurabili nel medio-lungo termine per le strutture di ricerca?</p> <p>Quali effetti più generali sono prefigurabili nel medio- lungo termine per le imprese?</p>	Campionamento stratificato e interviste telefoniche
	<p>Esistono tratti comuni tra le traiettorie definite nelle diverse aree di specializzazione? È possibile individuare uno schema condiviso tra le diverse aree di specializzazione entro il quale sia possibile individuare delle traiettorie comuni o almeno parzialmente sovrapponibili?</p> <p>In che misura tale (eventuale) schema comune trova corrispondenza nello schema nazionale del Piano Nazionale della Ricerca?</p>	Analisi desk e questionari online rivolti ai referenti sia delle imprese che delle altre tipologie di strutture (dipartimenti universitari, enti di ricerca e di trasferimento tecnologico, ecc.) coinvolte nei progetti finanziati.
	<p>Quali progetti stiamo finanziando? Quali traiettorie stiamo seguendo?</p> <p>In che misura è coinvolto il sistema regionale? È migliorata (sta migliorando) la capacità innovativa del sistema regionale? È migliorata (sta migliorando) la competitività del sistema economico regionale sia in generale che nelle aree di specializzazione della S3?</p> <p>Si sta assistendo allo sviluppo di nuovi settori/ ambiti produttivi in grado di favorire una trasformazione positiva del sistema economico regionale?</p>	<p>Analisi desk</p> <p>Monitoraggio delle aree di specializzazione</p> <p>Definizione di una "situazione controfattuale" e interviste con stakeholder</p> <p>Approccio <i>theory-based</i></p>

	<p>Quali azioni hanno avuto un ritardo o un'un'incompleta attuazione e quali sono state le ragioni dei ritardi o della mancata attuazione di quanto previsto?</p> <p>In quale misura le azioni hanno contribuito al raggiungimento degli obiettivi strategici previsti nel dal PON REC 2007-2013 e dal PON Ricerca e Innovazione 2014-2020?</p> <p>Gli indicatori utilizzati per il monitoraggio e la valutazione del PON REC hanno la capacità di cogliere i risultati delle azioni? In che misura le azioni hanno contribuito a rafforzare e valorizzare l'intera filiera della ricerca e le reti di cooperazione tra il sistema della ricerca e le imprese?</p> <p>In quale misura le azioni sono state capaci di sostenere la diffusione e l'utilizzo di nuove tecnologie e servizi avanzati presso i beneficiari e nell'intero sistema locale?</p> <p>Il livello delle competenze e delle conoscenze scientifiche e tecniche nel sistema produttivo e nelle Istituzioni ha beneficiato delle azioni realizzate? Ci sono state ricadute in termini di creazione e di qualificazione dell'occupazione?</p>	<p>Controfattuale</p> <p><i>Theory-based</i></p>
--	---	--

Come evidenziato dalla tabella, le domande valutative sollevate si riferiscono sia al processo di attuazione che agli esiti delle strategie già sviluppate. Rispetto agli approcci, predominano ancora le analisi documentali e le analisi dei dati di monitoraggio mentre la complessità delle politiche in esame rende evidente la necessità di realizzare valutazioni di sistema che verifichino l'efficacia delle strategie complessive. A tal fine, occorre individuare le unità di analisi da considerare a livello micro, meso e macro. Si tratta di una scelta che può far emergere una serie di altri quesiti particolarmente spinosi, come ad esempio: nella selezione delle aree di priorità, ogni progetto di investimento è da considerarsi separatamente, oppure in stretta connessione - o in rete - con altri progetti? In che modo bisogna analizzare le interazioni e le interdipendenze tra i progetti? Nel caso in cui vi siano più misure che presentano obiettivi simili, occorre valutare separatamente ciascuna di esse o considerarle come un tutto unico?

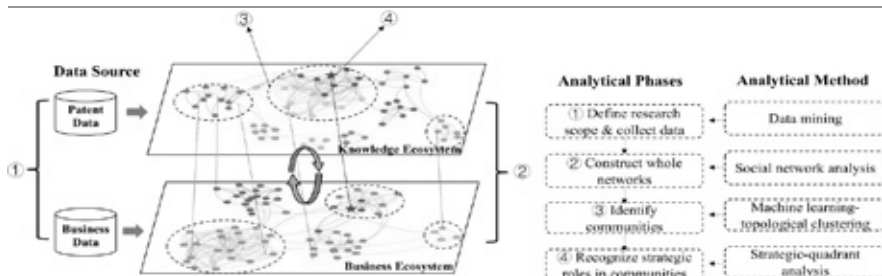
I quesiti sollevati richiamano l'importanza di tenere conto dei cambiamenti comportamentali che intervengono nei gruppi target ma anche tra gli attori che non sono stati sostenuti dalla politica. Occorre allora impostare uno studio che esplori il contesto e si concentri sui rapporti tra i beneficiari e i gruppi di stakeholder collegati secondo un pensiero sistemico aperto, in cui gli attori – individui, organizzazioni e istituzioni – interagiscono nel tempo. La valutazione è chiamata ad esaminare il cambiamento osservabile là dove i programmi sono attuati al fine di apprendere dall'esperienza e non deve semplicemente descrivere attività, realizzazioni e risultati conseguiti. Occorre, inoltre, considerare la presenza di molteplici interventi che, interagendo vicendevolmente, non permettono di isolare gli effetti attribuibili ad un intervento rispetto a un altro. Ciò in considerazione del fatto che la maggior parte della ricerca sui cosiddetti ecosistemi dell'innovazione si è concentrata su specifiche dimensioni, sottovalutando le eterogeneità e le connessioni tra diversi livelli di interazione, come ad esempio, nelle relazioni tra ecosistemi della conoscenza e ecosistemi di business (Adner, 2019; Xu *et al.* 2018).

Nonostante esistano numerose prove scientifiche sull'impatto della conoscenza nello sviluppo dell'innovazione, non sono ancora stati valutati gli effetti economici/cognitivi/tecnologici di interventi mirati alla costruzione di ecosistemi di imprese, nell'ambito della definizione delle priorità delle strategie di specializzazione intelligente (Sebestyén e Varga, 2019). La definizione delle priorità ha l'obiettivo di selezionare determinati campi o attività della conoscenza che potrebbero diventare la base per una nuova specializzazione dell'economia di una regione. Un aspetto chiave nella definizione delle priorità è il potenziale di diffusione della conoscenza e delle tecnologie più importanti, che possono favorire la localizzazione di nuove imprese e il conseguente cambiamento strutturale nell'economia della regione.

Nell'ambito delle tecniche più avanzate di analisi dei dati, da cui la valutazione ha molto da apprendere, nuovi strumenti e informazioni di diverse fonti possono essere utilizzati per profilare gli ecosistemi relativamente alle scelte di investimento e alle performance delle imprese sia sul piano della diffusione della conoscenza sia sul piano dell'innovazione della produzione. L'analisi delle reti multilivello è un approccio analitico che permette di mappare le interdipendenze esistenti tra diversi sistemi. Si possono verificare le complementarità o le economie di rete che le collaborazioni tra imprese, università, centri di ricerca, agenzie di governo locale e regionale possono generare, come suggeriscono i modelli dell'innovazione regionale (Etzkowitz e Leydesdorff, 2000). L'analisi di rete multilivello risponde a domande tipo: Qual è la

visione di un ecosistema dell'innovazione che consente di ricostruire le connessioni che si attivano tra gli agenti sul piano della conoscenza e sul piano delle collaborazioni economico-commerciali? Quali sono gli attori più importanti in un ecosistema e a quali livelli operano? Quali sono i ruoli strategici delle imprese rispetto a più livelli del sistema? L'analisi di rete multi-livello – come già menzionato precedentemente – prova a mappare un ecosistema dell'innovazione caratterizzato da diversi strati e indaga le complementarità emergenti tra i diversi strati.

FIG. 6 MAPPATURA DELLE RETI DI CONOSCENZA E DI BUSINESS IN UN ECOSISTEMA DELL'INNOVAZIONE



Fonte: Xu et al. 2018.

L'analisi delle reti multilivello – una tecnica di analisi che sempre più di frequente viene adottata negli studi valutativi – richiede l'osservazione dei comportamenti degli attori e dei sistemi di interesse nonché l'utilizzo di strumenti econometrici per computare le informazioni qualitative secondo un'analisi dinamica delle interazioni. Si tratta di una tecnica di analisi che presenta un certo grado di difficoltà il cui vantaggio risiede nella ricostruzione delle interdipendenze che esistono tra diversi livelli di intervento e di interazione che coinvolgono molteplici gruppi di attori. Come rappresentato nella figura 6, ad esempio, le reti che si sviluppano nelle collaborazioni di R&S – caratterizzate dai dati sui brevetti – si collegano alle reti che si formano a partire dagli scambi commerciali e dalle relazioni commerciali, secondo metriche economiche. La tecnica delle reti multi-livello permette di scattare una fotografia delle complementarità esistenti ma anche di simulare potenziali connessioni da attivare attraverso interventi di politica pubblica di livello meso.

Esistono anche altri modelli noti come *Agent-based Models* (ABM) in grado di catturare le dinamiche di innovazione esaminando le reti di cooperazione tra gli agenti (v. Sebestyén e Varga, 2019). I modelli basati sull'agente sono una classe di modelli

computazionali finalizzati alla simulazione di azioni e interazioni di agenti autonomi ed eterogenei (tanto individuali quanto collettivi, come organizzazioni e gruppi) al fine di valutare i loro effetti sul sistema nel suo complesso. L'ABM combina elementi di teoria dei giochi, sistemi complessi, comportamento emergente, sociologia computazionale, sistemi multi-agente. Si sostiene che i modelli in esame siano utili in un contesto di valutazione delle politiche *ex ante* (Sebestyén e Varga, 2019); tuttavia, l'obiettivo di queste analisi non è descrivere la formazione della rete nel suo complesso, ma di esaminare alcune proprietà dinamiche dei sistemi di innovazione regionali. Alcuni modelli basati sugli agenti si concentrano, infatti, sul comportamento e sulla motivazione delle imprese nel processo d'innovazione, in cui la formazione della rete è di secondaria importanza. Pertanto, i modelli in esame si rivelano troppo complessi per essere integrati in un modello di analisi dell'impatto economico complessivo.

Gli ABM, rispetto alle reti multi-livello, sono più appropriati nell'analisi delle interazioni di livello micro (ad es. consumatore, impresa, centro di ricerca, ecc.). D'altronde, le interazioni dinamiche possono generare implicazioni di livello macro, e gli ABM possono mettere a fuoco comportamenti rilevanti, traducendo quei comportamenti osservati empiricamente in regole su cui viene modellato il comportamento degli agenti nell'ambiente micro, meso, o macro al fine di verificare l'esito non deterministico delle interazioni aggregate. I modelli di livello meso hanno il vantaggio di poter considerare tanto i micro-fondamenti dei processi di causazione quanto le interazioni sociali che, propagandosi nel tempo e nello spazio, possono generare impatti di livello aggregato (macro)¹⁵. Ad esempio, la lente valutativa di livello meso permette di intercettare il cambiamento individuale (micro) a partire dalle dinamiche di diffusione dell'innovazione, di adattamento e co-evoluzione dei sistemi di produzione verso i cambiamenti strutturali desiderati.

La diffusione delle innovazioni, in particolare, è un ambito di ricerca che riflette le decisioni di singoli produttori, ma è anche frutto dell'interazione tra le caratteristiche dell'impresa e le tecnologie abilitanti dei territori nonché i tratti socioeconomici e politico-istituzionali dei contesti regionali. Come le reti multi-livello, anche gli ABM sono approcci simulativi utili nelle circostanze in cui i modelli comportamentali non possono essere descritti analiticamente ma possono essere rappresentati attraverso scenari *what-if*.

¹⁵ *La lente meso non è né una categoria amministrativa né una categoria statistica. L'approccio di analisi di livello meso si concentra sui meccanismi che modellano le interazioni sociali e la conoscenza, attraverso l'innovazione, l'adattamento e il coordinamento tra gli agenti. Man mano che gli agenti apprendono e applicano regole di comportamento a livello micro, l'interazione sociale genera un processo di conoscenza di livello macro.*

Ricostruzione delle teorie di cambiamento

Nella valutazione delle politiche d'innovazione occorre sviluppare la teoria del cambiamento, che è un compito niente affatto facile se lo stato dell'arte della conoscenza sul programma o sul contesto è limitato oppure se il programma si articola lungo estese catene di implementazione. La condizione appena richiamata è abbastanza frequente quando i programmi vengono attuati con la partecipazione di molteplici soggetti attuatori che svolgono il ruolo di 'agenti' o nelle circostanze in cui esiste una pluralità di misure. La sfida consiste, allora, nel ricostruire le teorie del cambiamento che fanno capo agli agenti coinvolti e alle diverse misure introdotte nei programmi. Si tratta di ricostruire e collegare una pluralità di teorie plausibili e logicamente coerenti. Come già suggerito precedentemente, è importante raccogliere dati sui processi di innovazione – es. investimenti e localizzazione delle imprese, produzione di brevetti, collaborazione tra agenti e università, ecc... È necessario avvalersi di strumenti analitici e informativi (big data, semantica web, ecc.) in grado di esaminare diverse tipologie di dati e fornire una rapida restituzione delle informazioni.

Esplorare i processi di radicamento territoriale (embeddedness)

La ricostruzione delle teorie del cambiamento non può svilupparsi come un esercizio di logica avulso dal contesto di riferimento. Piuttosto, per formulare un giudizio sulla strategia di specializzazione intelligente, occorre considerare il radicamento dei sistemi produttivi territoriali e le specializzazioni basate sui vantaggi competitivi esistenti. Gli studi sull'innovazione finora disponibili evidenziano che le regioni europee differiscono per la loro capacità innovativa a causa delle condizioni iniziali, delle dotazioni di risorse e dei livelli di sviluppo e innovazione già raggiunti. Numerosi studi convergono intorno all'idea secondo cui le economie regionali si muovono lungo traiettorie tecnologiche collegate al contesto produttivo, dove la ricerca scientifica e l'innovazione sono improntati alle capacità di creazione di conoscenza già esistenti e alle routine consolidate di diffusione delle informazioni tra sistema delle imprese e agenti della ricerca scientifica (Balland *et al.*, 2019). I c.d. sistemi di innovazione regionale (Regional innovation systems - RIS), come sono stati mappati in tutt'Europa, riflettono, tuttavia, un approccio top-down che non ha sufficientemente considerato i modelli di apprendimento emergenti a livello locale, le specificità dei contesti regionali, la distribuzione diseguale dell'imprenditorialità nei territori e le complesse dinamiche che caratterizzano i processi di apprendimento interattivo (Cooke, 2001, Becattini *et al.*, 2009; Quintero *et al.*, 2013). Solo di recente, come già sottolineato precedentemente, la mappatura degli ecosistemi ha adottato tecniche di rete multi-

livello per rilevare le connessioni tra imprese e altri agenti – come università e centri di ricerca – su diversi piani e in base a molteplici caratteristiche di interesse (Xu *et al.*, 2020, Vasilyeva *et al.*, 2021).

Un altro aspetto cruciale da considerare nel disegno di una valutazione ha a che fare con il radicamento dei sistemi di innovazione e l'eterogeneità delle imprese. Tali dimensioni suggeriscono che la variabilità del comportamento degli ecosistemi e delle singole imprese è un fattore di complessità nella verifica degli effetti delle politiche di innovazione. Occorre, infatti, tenere insieme più unità di analisi, centrate a livello dell'impresa, ma anche a livello territoriale e settoriale. Nell'analisi delle strategie di innovazione di un ampio campione di imprese italiane, ad esempio, Antonioli *et al.* (2013) mostrano che le dimensioni e il settore dell'impresa catturano una significativa eterogeneità aziendale ma anche reazioni innovative rispetto alle crisi. Altre caratteristiche a livello di imprese associate alle performance innovative sono la spesa pubblica e privata per R&S e condizioni di lavoro favorevoli e/o sfavorevoli (Antonioli *et al.*, 2013). Hervás- Oliver *et al.* (2021) esplorano anche l'innovazione nelle PMI che collaborano con altre aziende o istituzioni scientifiche e tecnologiche. Gli autori osservano che le PMI hanno una tendenza e una capacità di gran lunga inferiori rispetto alle aziende di più grandi dimensioni di avviare progetti di R&S e collaborare con le università. Le PMI sono più orientate al mercato e seguono processi di innovazione non basati su R&S (Hervás-Oliver *et al.*, 2021). Le attività interne di R&S possono esporre le PMI ad elevati rischi e costi fissi collegati ai requisiti minimi di investimento e ai vincoli nell'erogazione del credito. Soprattutto all'interno dei sistemi di innovazione in ritardo di sviluppo, le PMI preferiscono investire in attività come l'innovazione manageriale (ovvero, marketing e innovazione organizzativa), collaborando con fornitori, clienti o concorrenti (Hervás - Oliver *et al.*, 2021). Gli studi disponibili, tuttavia, non spiegano se e come le interazioni collaborative con le università e altri agenti di ricerca e sviluppo possano accelerare la trasformazione digitale anche all'interno di territori meno innovativi (Marra *et al.*, 2022). In tal senso, nel disegno di uno studio valutativo occorre non solo differenziare tra imprese a basso e alto rendimento, piccole e medie imprese e grandi gruppi industriali e tecnologici ma anche tra imprese e regioni che ricevono sostegni finanziari attraverso politiche di innovazione o di welfare. Ciò in quanto esiste una relazione tra le capacità di assorbimento delle imprese e il potenziale imprenditoriale di livello regionale che può influenzare processi di innovazione micro e macro (Marra *et al.*, 2022). In tal senso, la valutazione delle politiche di specializzazione intelligente verifica risultati di scala variabile a seconda del radicamento territoriale e regionale delle attività imprenditoriali all'interno e tra i settori produttivi.

Occorre affrontare l'imprevedibilità e l'emergenza dei processi di scoperta imprenditoriale: nei casi in cui i programmi hanno una traiettoria d'impatto di lungo periodo e nei casi in cui essi operino in contesti ancora poco esplorati, i disegni di valutazione devono essere flessibili, eventualmente sviluppati per fasi, in cui operare aggiustamenti iterativi nel tempo. In particolare, il concetto di emergenza è cruciale al fine di predisporre una strategia valutativa in grado di osservare gli effetti di innovazione che man mano trascendono l'entità che li genera. Come già sottolineato a più riprese, nella strategia di specializzazione la fase di c.d. *entrepreneurial discovery* è un processo 'emergente', interattivo e inclusivo in cui gli attori identificano potenziali nuove attività e informano le istituzioni di governo. Le istituzioni di governo regionale valutano, a loro volta, le informazioni ricevute e devolvono agli agenti il potere di perseguire gli investimenti più promettenti. Si tratta di un processo di natura sperimentale che non è solo collegato alla definizione delle priorità di investimento nel campo della ricerca e dell'innovazione, ma è anche e soprattutto un processo sociale e politico, ove il potere e gli interessi consolidati di diversi gruppi devono essere presi in considerazione (Magro e Wilson, 2019). Inoltre, come evidenziato da Marinelli e Perianez Forte (2017), concettualmente il processo di scoperta imprenditoriale non è relegato alla fase di progettazione di una strategia intelligente ma è un'attività continuativa. Sul piano della valutazione, ciò comporta il passaggio da approcci tradizionali verso approcci partecipativi in cui gli attori vengono coinvolti nella definizione e nell'attuazione delle strategie di specializzazione e consultati non solo al fine di estrarre informazioni rilevanti per la valutazione, ma anche per partecipare all'elaborazione del disegno valutativo. Ciò al fine di impostare domande pertinenti e orientare il processo di raccolta dei dati rispetto ai risultati attesi o considerati importanti per sostenere l'apprendimento del sistema, rafforzare i circuiti di feedback e migliorare l'accesso alle informazioni.

La lente dell'emergenza è anche rilevante nella comprensione delle "specializzazioni" diversificate secondo cui le regioni identificano le aree o i domini in cui esistono vantaggi competitivi anche potenziali al fine di specializzarsi in maniera differenziata rispetto ad altre regioni anche più avanzate. Ciò pone un problema di posizionamento del valutatore che è chiamato a giudicare sul valore e sul merito di strategie di specializzazione che provano a massimizzare il potenziale di sviluppo basato sulla conoscenza sia con sistemi di ricerca e innovazione avanzati e una struttura industriale high-tech, sia con sistemi più tradizionali ancora carenti sul piano delle connessioni tra ricerca scientifica e produzione industriale.

De Haan – uno studioso dei processi emergenti – mette a punto una tassonomia utile a comprendere la posizione del valutatore rispetto alla relazione tra la proprietà emergente e i suoi elementi costitutivi. Secondo De Haan, l'emergenza rispetto al valutatore può essere classificata come scoperta, esito meccanicistico o riflessivo. Nell'emergenza come scoperta, l'osservatore è esterno al processo e il cambiamento scaturisce come esito unidirezionale che dalle parti costituenti conduce alla proprietà emergente. In questa situazione, il valutatore ricostruisce le scelte di investimento produttivo a seconda delle caratteristiche materiali e culturali dei sistemi territoriali. Ma questo non è che il primo passo per cogliere i segni più evidenti dei processi di diversificazione produttiva. Nella seconda categoria di emergenza il valutatore è esterno al processo ma riconosce l'esistenza di una relazione tra le parti costituenti e le proprietà emergenti che possono incidere nel tempo sugli elementi costitutivi. L'osservatore, pur rimanendo esterno al processo, riconosce questo rapporto e ricostruisce le dinamiche emergenti negli ecosistemi dell'innovazione, i quali evolvono nel tempo con esiti imprevedibili. Diventa allora cruciale interrogarsi sul senso delle relazioni che si generano e rigenerano nella continua interazione tra gli attori per scongiurare il rischio di verificare solo gli esiti desiderati o desiderabili.

Diversamente, nell'emergenza riflessiva, il valutatore è interno al processo e partecipa al cambiamento a partire dalla sfera della propria esperienza soggettiva. Il valutatore è attivamente coinvolto e contribuisce a interpretare la realtà condivisa con il contesto osservato, ridefinendo gli elementi e le caratteristiche del sistema, oggetto della valutazione. Ciò significa che il valore e il merito delle strategie di specializzazione non possono essere considerate in fase *ex ante* ma solo in fase *ex post* secondo le teorie che gli attori coinvolti nel processo evidenziano come caratteristiche salienti e generative di ulteriore valore e sviluppo del territorio. Ciò pone anche un problema di posizionalità del valutatore rispetto agli attori coinvolti nella ricostruzione della teoria del cambiamento che richiede una continua riflessività sul ruolo svolto e la capacità di mettere in questione premesse e risultati del lavoro valutativo.

Ricapitolando: nella valutazione delle politiche a favore dell'innovazione, le domande poste dai committenti influenzeranno il disegno dello studio orientato all'analisi del processo di attuazione, delle performance o dell'impatto degli interventi. Questi ultimi si caratterizzano per la commistione tra incentivi pubblici e investimenti privati tesi alla specializzazione intelligente con variabili forme di condivisione e diffusione della conoscenza che non necessariamente può essere quantificata in termini monetari. La creazione o il consolidamento di una rete di attori suggerisce l'opportunità di ricostruire più teorie di cambiamento che possono co-esistere, rafforzarsi vicendevolmente o entrare in competizione. Se l'obiettivo consiste nel ricostruire le complementarità

tra imprese, centri di ricerca e istituzioni di governo del territorio, occorrerà esplorare le interdipendenze che possono intervenire a diversi livelli operativi e secondo diverse dinamiche d'interazione. In tal caso la modellistica delle reti multi-livello o degli agenti eterogenei (Agent-based models) offre strumenti teorici per confrontare le dinamiche verificabili nei contesti e metodologie econometriche e computazionali per trattare informazioni quali-quantitative complesse.

Negli studi in cui esplicitamente si affronta la complessità delle interazioni di natura sociale, cognitiva e produttiva, la scelta dell'unità di analisi non può non tenere insieme variabili a livello di impresa e fattori di contesto per intercettare il radicamento dell'innovazione nel comportamento delle aziende e degli ecosistemi. Proprio al fine di comprendere come emergono gli ecosistemi di innovazione, occorre esplorare i contesti produttivi in cui ricostruire le interdipendenze tra imprese, università, centri di ricerca e istituzioni di governo regionale che possono creare e consolidare specializzazioni intelligenti. Le trasformazioni economiche attese non comportano solo l'adozione di tecnologie avanzate, ma l'emergere di specializzazioni che assecondino le priorità scelte da ciascun territorio, diffondendo conoscenza scientifica e conoscenza applicata alle produzioni. In tali circostanze, il valutatore è chiamato a immergersi nei contesti regionali per evitare di omologare gli effetti osservati rispetto a standard non congeniali alle specificità dei sistemi produttivi locali sviluppando riflessività e pensiero critico.

CONCLUSIONI

Le complesse politiche pubbliche dei nostri tempi mettono alla prova le competenze e le abilità metodologiche dei valutatori nel verificare il cambiamento prodotto nei contesti (Rogers, 2008, 2011; Marra, 2011). I valutatori possono contribuire a migliorare le politiche e i programmi pubblici ricostruendo le motivazioni che muovono gli attori, i gruppi e le istituzioni a mutare il proprio comportamento. Si tratta di sviluppare analisi in grado di "far emergere" come, perché, per chi e in quali circostanze i programmi a favore dell'occupazione e dell'innovazione generano esiti trasformativi. Ciò non significa rinunciare all'esame puntuale dei dati al fine di intercettare i successi e i fallimenti delle politiche. Piuttosto, la ricostruzione delle teorie del cambiamento permette di sviluppare una conoscenza più approfondita delle finalità degli interventi, dei processi di attuazione e dei contesti ove i programmi vengono attuati. Si tratta di una conoscenza politicamente rilevante che può contribuire ad informare con evidenze scientifiche la programmazione di interventi complicati e complessi.

Nell'ambito dell'approccio *theory-based* vengono ricompresi diversi disegni valutativi accomunati da un'interpretazione generativa della causalità. Secondo tale interpretazione, i programmi possono attivare complessi processi di interazione sociale sollecitando le reazioni dei beneficiari e delle comunità in contesti via via più ampi. E proprio per le politiche che presuppongono un utente motivato – come le politiche attive del lavoro – che risponda alle opportunità offerte, ricostruire la teoria del cambiamento consente di valutare l'impatto sulla base delle reazioni dei beneficiari in relazione alle premesse su cui si basano gli interventi coerentemente con le esigenze e le specificità dei contesti. Particolare importanza assumono la teoria del programma e la teoria dell'implementazione nell'orientare i valutatori verso l'analisi di processi trasformativi piuttosto che verso la misurazione quantitativa degli impatti. In tal senso l'approccio *theory-based* può utilmente integrare le analisi controfattuali, affiancando alla quantificazione degli effetti informazioni qualitative sugli effetti differenziali e sulle eterogeneità dei processi di attuazione di interventi sempre più adattati alle realtà del mercato del lavoro locale e regionale.

In condizioni d'incertezza elevata e nelle circostanze in cui si sperimentano azioni innovative con effetti potenzialmente ambigui, la valutazione orientata alla teoria contribuisce ad informare le scelte gestionali di complessi processi di coordinamento multi-livello. I principi della scienza della complessità possono sostenere la

ricostruzione di teorie di cambiamento secondo un pensiero sistemico aperto in grado di intercettare fenomeni complessi frutto di situazioni interdipendenti che emergono, si adattano e si radicano in alcuni contesti. In tali circostanze, gli effetti osservabili non si aggregano semplicemente addizionando gli elementi costitutivi, ma attraverso la comprensione di processi emergenti. Ciò pone l'esigenza di sviluppare un diverso pensiero valutativo sensibile alla complessità e una cultura della valutazione come apprendimento dall'esperienza, prima ancora di applicare nuovi strumenti e tecniche di analisi che la scienza dei dati ci offre.

RIFERIMENTI BIBLIOGRAFICI

Abreu A., Urze P. (2016), "System thinking shaping innovation ecosystems" in *Open Engineering*, vol. 6, no. 1, <https://doi.org/10.1515/eng-2016-0065>.

Adner R. (2019), "Ecosystem as Structure: An Actionable Construct for Strategy" in *Journal of Management*, 43(1): 39-58.

Agenzia per la Coesione Territoriale/NUVAL (2017), Progetto Pilota di Valutazione Locale. Valutare dal locale per il locale. Un approccio valutativo per rafforzare le strategie locali e stimolare il dialogo sugli esiti delle politiche pubbliche: http://valutazioneinvestimenti.formez.it/sites/all/files/rapporto_progetto_pilota_reves_06_03_2017.pdf.

Bredgaard T. (2015), "Evaluating What Works for Whom in Active Labour Market Policies" in *European Journal of Social Security*, 17(4), 436-452 <https://doi.org/10.1177/138826271501700403>.

Callon M., Lascoumes P. e Barthe Y. (2009), *Acting in An Uncertain World: An Essay on Technical Democracy*, Cambridge, MA, MIT Press.

Csillag M., Krekò J., Scharle A. (2020), Valutazione controfattuale dell'impatto delle politiche sull'occupazione giovanile, Guida metodologica, INAPP.

Elster J. (1998), "A Plea for Mechanisms", in Hedström O., Swedberg R. (a cura di), *Social Mechanisms: An Analytical Approach to Social Theory*, Cambridge, Cambridge University Press.

Etzkowitz H., Leydesdorff L. (2000), "The Dynamics of Innovation: From National Systems and «Mode 2» to a Triple Helix of University-Industry-Government Relations" in *Research Policy*, 29(2):109-123.

European Commission (2020), European semester thematic factsheet: Active labour market policies: https://commission.europa.eu/system/files/2020-06/european-semester_thematic-factsheet_active-labour-market-policies_en_0.pdf.

European Commission (2018), Labour market policy statistics: Methodology 2018, Luxembourg, Publications Office of the European Union.

European Commission (2012), Analysis of costs and benefits of active compared to passive measures, Luxembourg, Publications Office of the European Union.

European Training Foundation (2022), Assessment of the effectiveness of active labor market policies in crisis and post-crisis situations: https://www.etf.europa.eu/sites/default/files/2022-02/almps_effectiveness_0.pdf

Foray D., Eichler M. e Keller M. (2021), "Smart specialization strategies—insights gained from a unique European policy experiment on innovation and industrial policy design" in Review of Evolutionary Political Economy 2, 83–103 <https://doi.org/10.1007/s43253-020-00026-z>.

Forss K., Marra M., Schwartz R. (2011), *Evaluating the complex: Beyond attribution and contribution*, Routledge, Francis & Taylor.

Funnell S., & Rogers P. J. (2011), *Purposeful program theory, Effective use of the theories of change and logic models*. Jossey Bass/John Wiley & Sons Inc.

Government of Canada, Theory-Based Approaches to Evaluation: Concepts and Practices: <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/theory-based-approaches-evaluation-concepts-practices.html#toc4>.

Hegyí F.B., Prota, F. (2021), Assessing Smart Specialisation: Monitoring and Evaluation Systems, Luxembourg, Publications Office of the European Union (online).

Hervás-Oliver J-L, Parrilli M.D., Rodríguez-Pose A., Sempere-Ripoll F. (2021), *The drivers of SME innovation in the regions of the EU* in Research Policy, 50 (9), <https://doi.org/10.1016/j.respol.2021.104316>.

Hidalgo C.A. (2021), "Economic complexity theory and applications" in Nature Review of Physics 3, 92–113 <https://doi.org/10.1038/s42254-020-00275-1>

Hwang, I. (2020), "An Agent-Based Model of Firm Size Distribution and Collaborative Innovation" in Journal of Artificial Societies and Social Simulation 23(1) 9, <https://doi.org/10.18564/jasss.4190>.

Kahneman D., Tversky A. (1992), "Advances in Prospect Theory: Cumulative Representations of Uncertainty" in *Journal of Risk and Uncertainty*, 5: 297-323.

Leone L., Giunta G. (2019), *Riqualificazione urbana e lotta alle diseguaglianze. L'approccio delle capacitazioni per la valutazione di impatto del programma messinese*, HDE Civic Economy.

Marra M. (2010), "Note sulla causalità. Gli indicatori tra micro e macro analisi", in C. Bezzi, L. Cannavò, M. Palumbo, *Costruire e usare indicatori nella ricerca sociale e nella valutazione*, Milano, Franco Angeli.

Marra, M. (2011) Micro, meso, and macro dimensions of change: a new agenda for the evaluation of structural policies, in Forss, K., Marra, M., Schwartz, R. (eds) *Evaluating the Complex: Beyond Attribution and Contribution*, Taylor&Francis, Routledge, NY.

Marra M. (2016), An in depth look at causality, dispense corso IPDET, Banca mondiale/ Carlton University, Ottawa, CA.

Marra M. (2022), *Connessioni virtuose. Come nasce (e cresce) un ecosistema dell'innovazione*, Bologna, Il Mulino.

Marra, M. (2023) *L'impatto socioeconomico del Museo Archeologico Nazionale di Napoli*, Edizione Scientifica, Napoli.

Moizer A.D., Lean J. (2022), "A systems approach to mapping UK regional innovation ecosystems for policy insight" in *Industry and Higher Education*, Vol. 37, Issue 2 <https://doi.org/10.1177/09504222221115977>.

Ministero dell'Università e della Ricerca (2020), PNR 2021-2027 Piano Nazionale per la Ricerca: https://www.gea.mur.gov.it/docs/PNR_2021-2027.pdf.

Naldini A. (2013), "Lo stato della valutazione della politica di coesione in Italia", in Vergani A. (a cura di), *Prove di valutazione. Libro bianco sulla valutazione in Italia*, Milano, Franco Angeli.

Niyadurupola V., Esposito L. (2021), "What gets them going? The effects of activation policies on personal change processes of unemployed youth" in *Journal of Education and Work*, 34:4, 590-609 <https://doi.org/10.1080/13639080.2021.1953694>.

Palumbo M. (2001), *Il processo di valutazione. Programmare, decidere, valutare*, Milano, Franco Angeli.

Pawson R. (2013), *The Science of Evaluation: A Realist Manifesto*, London, Sage.

Pawson R., Tilley N. (1997), *Realistic Evaluation*, London, Sage.

Puig-Barrachina V., Malmusi D., Artazcoz L., Bartoll X., Clotet E., Cortès-Franch I., Ventura L., Daban F., Díez È, Borrell C. (2020), "How does a targeted active labour market program impact on the well-being of the unemployed? A concept mapping study on Barcelona "Employment in the Neighbourhoods"" in *BMC Public Health*, Mar 17;20(1):345: [dhttps://doi.org/10.1186/s12889-020-8441-2](https://doi.org/10.1186/s12889-020-8441-2); <https://pubmed.ncbi.nlm.nih.gov/32183755/>; <http://www.ncbi.nlm.nih.gov/pmc/articles/pmc7076971/>.

Regione Autonoma Friuli Venezia Giulia (2018), 1° Rapporto Tematico: Il modello di governance della S3 del FVG (online).

Reynolds M., Gates E., Hummelbrunner R., Marra M., & Williams B. (2016), "Towards Systemic Evaluation, Systems Research and Behavioral Science" in *Systems Research and Behavioral Science* 33: 662-673.

Rogers P. J. (2008), "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions" in *Evaluation*, 14(1), 29-48: <https://doi.org/10.1177/1356389007084674>.

Sebestyén T., Varga A. (2019), "Knowledge networks in regional development: an agent-based model and its application" in *Regional Studies*, 53:9, 1333-1343, <https://doi.org/10.1080/00343404.2019.1622663>.

Starke M., Hollertz, K. (2022), "Theoretical underpinnings of active labour market programmes implemented by municipal and non-governmental organisations in Sweden" in *European Journal of Social Work*: <https://doi.org/10.1080/13691457.2022.2063809>.

Stern E. (2016), *La valutazione di impatto. Una guida per committenti e manager preparata per Bond*, Milano, Franco Angeli.

Taplin D.H., Clark H. (2013), *Theory of change toolkit*. Mental Health Innovation network: <http://mhinnovation.net/resources/theory-change-toolkit>.

Vasilyeva E., Kozlov A., Alfaro-Bittner K. et al. (2021), "Multilayer representation of collaboration networks with higher-order interactions" in *Scientific Reports*, 11, 5666.

Vooren M., Haelermans C., Groot W., Maassen van den Brink H. (2019), "The effectiveness of active labor market policies: A meta-analysis" in *Journal of Economic Surveys*, 33(1):125-149: <https://doi.org/10.1111/joes.12269>.

Weiss C.H. (1998), *Evaluation*, Upper Saddle River, NJ, Prentice-Hall.

Weiss C.H. (1995), Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families. in James Connell et al. (eds) *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, Washington, DC, Aspen Institute.

Weiss C.H. (1998), "Theory-based Evaluation: Past, Present, Future", in *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods*, n. 76 (Winter), pp. 41-55.

White H. (2009), Theory-Based Impact Evaluation: Principles and Practice, Working paper 3, International Initiative for Impact Evaluation (3ie): https://www.3ieimpact.org/sites/default/files/2017-11/Working_Paper_3.pdf.

Xu G., Hu W., Qiao Y., Zhou Y. (2020), Mapping an innovation ecosystem using network clustering and community identification: a multilayered framework in *Scientometrics*, 124:2057-2081 <https://doi.org/10.1007/s11192-020-03543-0>.

Yeyati E.L., Montane M., Sartorio L. (2019), What works for active labor market policies?, CID Faculty Working Paper, No 358, Harvard University.

I Working Papers nascono dall'attività di progetto e di studio del gruppo di ricerca della Fondazione Giacomo Brodolini. Sono uno strumento agile di informazione che permette la sistematizzazione e la diffusione dei lavori realizzati sulle principali tematiche d'interesse della Fondazione.