



Monitoraggio e valutazione degli interventi per prevenire e contrastare lo sfruttamento del lavoro e il lavoro forzato in agricoltura

Nota tecnica

“Supporto al rafforzamento della *governance* inter-istituzionale in materia di sfruttamento lavorativo in Italia”



Co-funded by
the European Union



Indice

Introduzione.....	1
1. Principali caratteristiche dei sistemi di monitoraggio.....	2
1.1. <i>Matrice di monitoraggio: Obiettivi, risultati e indicatori</i>	<i>2</i>
2. Analisi dei dati di monitoraggio.....	5
3. La valutazione ex post	6
4. La valutazione controfattuale degli effetti delle politiche pubbliche	7
4.1. <i>Valutazione controfattuale: disegni sperimentali.....</i>	<i>8</i>
4.2. <i>Valutazione controfattuale: disegni non-sperimentali (o quasi sperimentali)</i>	<i>9</i>
5. La valutazione basata sulla teoria	11
5.1. <i>L'approccio realista alla valutazione degli effetti.....</i>	<i>11</i>
Riferimenti bibliografici	15

Introduzione

L'obiettivo di questa nota tecnica è quello di esaminare le caratteristiche dei sistemi di monitoraggio e le principali metodologie di valutazione *ex post* circa l'impatto di interventi di politica pubblica, per offrire spunti di riflessione alle istituzioni responsabili di attuare misure di contrasto allo sfruttamento lavorativo e di protezione e assistenza delle vittime.

Il monitoraggio e la valutazione sono i processi che consentono di valutare come si evolve un intervento di politica nel tempo (monitoraggio); quanto sia efficace l'attuazione dell'intervento e se vi sono divergenze tra i risultati attesi e i risultati ottenuti (valutazione); e se i cambiamenti osservati siano determinati dall'intervento attuato oppure da altri fattori (valutazione d'impatto).

Il monitoraggio è un processo di raccolta e analisi continua delle informazioni sull'attuazione degli interventi e di confronto tra i risultati ottenuti e quelli attesi con lo scopo di verificare il raggiungimento dell'obiettivo finale. Il monitoraggio utilizza le informazioni generate durante l'attuazione dell'intervento (per esempio risorse utilizzate, caratteristiche dei beneficiari, situazione al termine dell'intervento) per confrontare i risultati ottenuti rispetto a quelli attesi e valutare le differenze (per esempio tra i beneficiari ultimi, il tipo d'intervento, le aree geografiche).

La valutazione è un processo che verifica - in modo sistematico e secondo criteri oggettivi - tutti gli elementi di un intervento (per esempio pianificazione, attuazione e risultati raggiunti) per determinarne il valore complessivo. Lo scopo è di fornire informazioni basate sull'evidenza per identificare le modalità per ottenere risultati sempre migliori. In linea di massima, esistono due tipi di valutazione:

- Le **valutazioni di performance** si concentrano sulla qualità dell'attuazione e sui risultati raggiunti da un intervento nel breve e medio periodo. Sebbene questo tipo di valutazione sia svolta sulla base delle informazioni raccolte dal sistema di monitoraggio, lo scopo è più ampio e teso a valutare se i cambiamenti osservati siano dovuti agli interventi attuati o ad altri fattori.
- Le **valutazioni d'impatto** si concentrano sulla differenza tra quanto è stato ottenuto attraverso la realizzazione dell'intervento e quanto si sarebbe ottenuto nella situazione ipotetica ("controfattuale") in cui l'intervento non fosse stato realizzato. Questo tipo di valutazione ha lo scopo di isolare i cambiamenti causati dall'intervento da quelli che si sono verificati in un determinato lasso di tempo nel gruppo di beneficiari o nell'area geografica oggetto dell'intervento stesso.

Il monitoraggio e la valutazione dovrebbero includere anche informazioni relative alle risorse finanziarie investite. Queste informazioni consentono di stimare i benefici apportati degli interventi rispetto ai loro costi e di identificare quale tra gli interventi attuati abbia il più alto tasso di rendimento. A questo scopo si possono utilizzare due approcci:

- L'**analisi costi-benefici** stima i benefici totali di un intervento rispetto ai costi. Questo tipo di analisi viene normalmente utilizzata *ex ante*, per decidere tra diverse opzioni di intervento. La principale difficoltà di tale approccio è la stima del valore monetario dei benefici cosiddetti "immateriali". Ad esempio, se lo scopo principale di un intervento di reinserimento socio-lavorativo è l'aumento delle opportunità occupazionali e salariali

per le vittime di sfruttamento lavorativo, questi sono considerati benefici tangibili ai quali può essere attribuito un valore monetario. Tuttavia, avere un lavoro e un salario dignitoso aumenta anche l'autostima e l'autodeterminazione delle vittime, fattori difficili da esprimere in termini monetari poiché assumono valori diversi a seconda degli individui.

- L'**analisi costi-efficacia** confronta i costi di due o più interventi nel raggiungere lo stesso risultato. Per esempio, l'istituzione di un numero verde per le vittime di sfruttamento lavorativo e le ispezioni nelle aziende agricole hanno entrambe lo scopo di identificare potenziali vittime di sfruttamento lavorativo. Il numero verde identifica potenziali vittime ad un costo di 10.000 euro per vittima, mentre l'ispezione ha un costo stimato di 6.000 euro per vittima. Le ispezioni hanno quindi un rapporto costo-efficacia migliore.

1. Principali caratteristiche dei sistemi di monitoraggio

Il monitoraggio e la valutazione sono strumenti fondamentali per acquisire dati e produrre informazioni affidabili sugli andamenti degli interventi e sugli attori istituzionali del sistema di prevenzione e contrasto allo sfruttamento lavorativo e sostegno alle vittime. Il monitoraggio mette in luce i punti di forza e di debolezza delle misure per attuare interventi correttivi e i miglioramenti necessari.

Un sistema di monitoraggio e di valutazione affidabile costituisce anche un supporto per le amministrazioni pubbliche centrali e locali responsabili delle politiche e dei finanziamenti ad esse associati per la verifica del rispetto degli adempimenti. Dotarsi di un sistema di monitoraggio consente inoltre di produrre parametri di riferimento (*benchmark*) per tutti i soggetti interessati, al fine di: (i) verificare l'evoluzione del fenomeno nel tempo; (ii) individuare possibili criticità sulle quali intervenire; (iii) permettere confronti su base territoriale; e (iv) misurare la qualità dei servizi forniti e la soddisfazione degli utenti.

Una condizione indispensabile per il funzionamento di un sistema di monitoraggio è il coinvolgimento di tutti i soggetti attuatori nella definizione degli elementi chiave dello stesso e nell'alimentazione delle informazioni per poter acquisire i dati all'origine. Tra gli elementi chiave di un sistema di monitoraggio va annoverato il coinvolgimento attivo dei partner istituzionali che hanno la responsabilità dell'attuazione degli interventi, in quanto modalità capace di soddisfare le diverse esigenze informative garantendo, allo stesso tempo, l'omogeneità dei contenuti.

1.1. Matrice di monitoraggio: Obiettivi, risultati e indicatori

La matrice di monitoraggio riassume, per ogni obiettivo, il traguardo principale da raggiungere ed il relativo metodo di misurazione, l'istituzione partner responsabile del coordinamento degli interventi, ed il quadro logico che sta alla base degli interventi.

Gli indicatori di *performance* sono misure concise, sia quantitative che qualitative, che possono essere facilmente misurate e che permettono di verificare l'andamento degli interventi a intervalli regolari. La formulazione degli indicatori determina le informazioni da raccogliere

e analizzare. Gli indicatori quantitativi misurano le variazioni di un valore specifico (numero, media o mediana) o di una percentuale. Gli indicatori qualitativi forniscono informazioni sui cambiamenti nelle attitudini, nelle convinzioni, nelle motivazioni e nei comportamenti degli individui. Per quanto importanti, le informazioni necessarie alla misurazione degli indicatori qualitativi richiedono più tempo per la raccolta, la misurazione e l'analisi, soprattutto nelle prime fasi di attuazione degli interventi.

È importante sottolineare che il successo di un sistema di monitoraggio e valutazione viene misurato non tanto in base alla quantità di dati e informazioni generate, ma piuttosto in termini dell'utilizzo che viene fatto delle informazioni messe a disposizione. A questo riguardo, uno dei principi del monitoraggio prevede che gli indicatori di performance siano SMART, ovvero:

S - *Specifici* (per misurare solo ciò che si deve misurare);

M - *Misurabili* (per raccogliere agevolmente le informazioni qualitative e quantitative necessarie),

A- *Disponibili* (fruibile e utile ai manager ad un costo accettabile);

R- *Rilevanti* (inerente al risultato e in grado di misurarne le variazioni nel corso del tempo) e

T- *Tempestivi* (circoscritti nel tempo).

La matrice di monitoraggio prevede tre diversi tipi di indicatori:

1. **Indicatori iniziali** (*baseline*): gli indicatori iniziali riportano il valore dell'indicatore di riferimento prima dell'attuazione dell'intervento. Tali indicatori sono determinati o sulla base di dati secondari (già disponibili) oppure possono richiedere la raccolta di dati primari.
2. **Indicatori di attuazione** (o di realizzazione o processo): questi indicatori misurano in modo continuativo la relazione tra gli input, le attività e gli output di un intervento o misura. In particolare, essi servono a verificare che gli input portino alle attività pianificate, e che le attività svolte portino agli output previsti.
3. **Indicatori di risultato**: Questi indicatori misurano i risultati di breve periodo che l'intervento raggiunge grazie alla produzione degli output. La loro verifica avviene in periodi precisi nel tempo (per esempio su base annua) per fornire un'istantanea dei risultati ottenuti a quella data. Questo tipo di indicatori sono i più rilevanti quando il monitoraggio riguarda interventi di politica pubblica, quali sono gli interventi per il contrasto dello sfruttamento lavorativo. Quando un indicatore di risultato specifica il valore (numero, media, mediana o percentuale) che questo dovrebbe raggiungere, l'indicatore diventa un traguardo (*target*). La funzione del sistema di monitoraggio, quindi, è quella di verificare a intervalli regolari se l'andamento dell'intervento sia sulla strada giusta per il raggiungimento del risultato.

Gli indicatori di attuazione e di risultato prevedono quasi sempre la disaggregazione dei dati da raccogliersi, indicata direttamente nella matrice. Le disaggregazioni più comuni sono le caratteristiche dei beneficiari dell'intervento (sesso, età, nazionalità); l'area geografica; e la tipologia di servizio reso (Riquadro 1).

Obiettivo da raggiungersi

Riquadro 1: Esempio di matrice di monitoraggio

Azione 5: Pianificazione e attuazione di soluzioni alloggiative dignitose per i lavoratori del settore agricolo in alternativa a insediamenti spontanei e altri alloggi degradanti				
Risultato principale	Indicatore	Definizione e misurazione	Baseline, traguardi e fonti	Unità responsabile e periodicità della raccolta dati
5.1. Definizione di criteri di qualità delle soluzioni alloggiative per la progettazione, gestione ed erogazione dei servizi a livello locale Definizione del risultato principale	Il 70 % dei lavoratori agricoli stanziali ha accesso ad alloggi in linea con gli standard minimi di qualità definiti a livello nazionale Valore dell'indicatore di risultato da raggiungersi entro il periodo Tipo di indicatore (di attuazione o di risultato) Livello di disaggregazione dei dati relativi all'indicatore	Numero di lavoratori agricoli stanziali che hanno accesso ad alloggi in linea con gli standard minimi di qualità sul totale dei lavoratori stanziali Definizione dell'indicatore di risultato Tipologia di indicatore Indicatore di risultato Disaggregazione - Caratteristiche dei lavoratori (sesso, età, nazionalità) - Tipologia alloggio - Area geografica La fonte delle informazioni necessarie a misurare il valore dell'indicatore	Baseline Informazioni sul numero di lavoratori agricoli in condizioni alloggiative degradanti non aggregate a livello nazionale Valore dell'indicatore iniziale (baseline) da cui inizia la misurazione Traguardo Entro il 2022, il 70% dei lavoratori agricoli stanziali ha accesso ad alloggi conformi agli standard minimi stabiliti a livello nazionali Fonti Dati del sistema informativo gestito dalla DG sistemi informativi del MLPS) Valore dell'indicatore (target) da raggiungersi nel periodo temporale indicato	Unità responsabile Associazione Nazionale Comuni Italiani (ANCI) Il partner istituzionale responsabile della raccolta e analisi dei dati di monitoraggio Periodicità: Annuale Frequenza della raccolta dati per il monitoraggio

Un altro gruppo di indicatori, da utilizzarsi in fase di analisi delle informazioni generate dal sistema di monitoraggio, sono gli indicatori di contesto. Essi rappresentano uno strumento indiretto per monitorare gli effetti dell'attuazione dell'intervento di politica, dato che descrivono il contesto in cui gli interventi vengono attuati e la sua evoluzione nel tempo.

La matrice di monitoraggio prevede, oltre agli indicatori, anche le unità responsabili per la raccolta e analisi dei dati, la periodicità della raccolta dati e le fonti da utilizzare. La periodicità della raccolta dati per il monitoraggio degli indicatori è quasi sempre annuale.

2. Analisi dei dati di monitoraggio

Vi sono due tipi di analisi, caratterizzate da modalità di attuazione e tempistiche diverse, che possono essere effettuate sui dati di monitoraggio.

La prima tipologia di analisi è uno strumento di gestione degli interventi che viene effettuata su base annua. L'analisi si basa sui valori degli indicatori contenuti nella matrice di monitoraggio ed ha lo scopo di accertare lo stato d'avanzamento degli interventi, valutare cosa sta funzionando e cosa no, al fine di fornire informazioni e dati puntuali durante l'attuazione. Il prodotto di tale analisi è un rapporto sintetico, corredato da una matrice con i dati del monitoraggio.

Il secondo tipo di analisi è una valutazione della *performance*. Questa si focalizza sulla qualità dell'attuazione e sui risultati raggiunti da un intervento. Questo tipo di valutazioni riguarda i risultati ottenuti nel breve e medio periodo (p.e. numero di vittime di sfruttamento lavorativo reintegrate nel mondo del lavoro), nonché la rilevanza (adeguatezza degli obiettivi ai problemi che si intendeva risolvere), l'efficacia (capacità dei risultati di contribuire al raggiungimento dello scopo), e l'efficienza (capacità di raggiungere i risultati previsti ad un costo ragionevole) degli interventi stessi. Questo tipo di valutazione ricomprende l'insieme delle attività analitiche ed interpretative volte ad esprimere un giudizio di valore sull'intervento, comparando i risultati ottenuti con quelli attesi, i tempi di realizzazione, i costi sostenuti e la qualità delle prestazioni. Il punto di partenza di questo tipo di valutazione sono le informazioni raccolte dal sistema di monitoraggio, con l'obiettivo di verificare se i cambiamenti osservati siano il risultato degli interventi attuati oppure di altri fattori di contesto.

3. La valutazione ex post

La valutazione è un processo che verifica in modo sistematico e secondo criteri oggettivi tutti gli elementi di un intervento di una o più politiche o programmi (per esempio pianificazione, attuazione e risultati raggiunti) per determinarne l'impatto complessivo. Lo scopo della valutazione è fornire informazioni basate sull'evidenza per identificare le modalità attraverso le quali ottenere risultati sempre migliori. Esistono due tipi principali di valutazione *ex post*, ovvero la valutazione di performance, già trattata nel paragrafo precedente, e la valutazione d'impatto.

La **valutazione d'impatto (o degli effetti)** di un intervento di politica pubblica ha lo scopo di verificare se, e in che misura, gli strumenti di *policy* adottati sono serviti a produrre i cambiamenti desiderati. La difficoltà principale che si riscontra in questo tipo di valutazione risiede nel distinguere quale parte del cambiamento osservato è direttamente attribuibile all'intervento pubblico e quanto invece è dovuto agli altri fattori che ne influenzano comunque l'evoluzione.¹ Esistono vari approcci che possono essere utilizzati per misurare gli effetti di un intervento di *policy*. La scelta dipende dal particolare intervento che si intende valutare e dalle specifiche domande di valutazione cui si vuole rispondere. Esistono diverse tipologie di valutazione d'impatto, ma gli approcci più consolidati sono:²

1. L'approccio *controfattuale*, che risponde alla domanda "si sono ottenuti i risultati voluti? gli effetti ottenuti sono stati causati dall'intervento?". Questo tipo di analisi si concentra sulla differenza tra quanto è stato conseguito con l'intervento e quanto sarebbe accaduto senza l'intervento, nella situazione ipotetica ("controfattuale") in cui l'intervento non fosse avvenuto. Questo approccio cerca di isolare i cambiamenti causati dallo specifico intervento tra tutti quelli che, in un determinato periodo tempo, si sono verificati nel gruppo o nel territorio oggetto dell'intervento. Il metodo controfattuale è adatto a valutare interventi circoscritti, con obiettivi chiari e meccanismi attuativi precisi.³
2. L'approccio *basato sulla teoria* analizza varie situazioni complesse dovute ai diversi contesti e meccanismi che vengono innescati dall'intervento. Questo approccio risponde alla domanda: "Cosa è successo? Cosa ha funzionato meglio, dove, per chi e perché?", concettualizzando l'intervento in termini di una teoria che ne spiega il funzionamento ed i meccanismi che hanno portato all'esito osservato.

Sin dalla programmazione del 2014-2020, il Fondo Sociale Europeo (FSE) ha previsto l'uso di valutazioni d'impatto per gli interventi co-finanziati attraverso le risorse europee. Il documento di orientamento del FSE suggerisce che i due approcci delineati sopra siano

¹ P.H. Rossi, M.W. Lipsey, H.E. Freeman: *Evaluation. A systematic approach* (7th Edition, Sage publication, 2020)

² Oltre all'approccio controfattuale e quello basato sulla teoria, la scienza sociale usa anche l'approccio *goal-free* (basato su criteri di qualità e sui bisogni delle persone a cui è rivolto l'intervento) e l'approccio *costruttivista/del processo sociale*, (basato sulla spiegazione di eventuali conflitti tra gruppi di partecipanti per arrivare a una conoscenza più accurata e condivisa). Questi approcci e i relativi metodi di utilizzo si trovano in G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale* (Ministero dello sviluppo economico, Unità di valutazione degli investimenti pubblici, 2011).

³ .H. Rossi, M.W. Lipsey, H.E. Freeman: *Evaluation. A systematic approach*, op.cit.

utilizzati in maniera complementare, al fine di fornire informazioni utili per la programmazione degli interventi del periodo successivo.⁴

4. La valutazione controfattuale degli effetti delle politiche pubbliche

La valutazione controfattuale si propone di verificare la capacità di una politica pubblica di modificare i comportamenti o le condizioni di un determinato gruppo di destinatari nella direzione desiderata. L'effetto viene definito come la differenza tra ciò che è accaduto *dopo* l'attuazione di una politica (situazione fattuale) e ciò che sarebbe accaduto se quella stessa politica non fosse stata realizzata (situazione controfattuale). Per esempio, la valutazione controfattuale di un intervento mirato a ridurre lo sfruttamento dei lavoratori del settore agricolo è tesa a misurare il numero di lavoratori che dopo l'intervento non sono più in una situazione di sfruttamento lavorativo rispetto al numero degli stessi lavoratori che non sarebbero più in una situazione di sfruttamento se l'intervento non fosse stato attuato. La differenza tra i due valori è l'effetto netto dell'intervento.

La difficoltà principale di questo approccio è la stima del secondo termine dell'equazione (quanti degli stessi lavoratori sarebbero emersi da una situazione di sfruttamento senza l'intervento), che non è possibile osservare direttamente. Da ciò deriva la necessità di produrre una stima affidabile dell'effetto, ricostruendo la situazione controfattuale con dati osservabili e in grado di approssimare ciò che sarebbe successo ai beneficiari in assenza dell'intervento.⁵ Queste stime vengono effettuate costruendo dei "gruppi di controllo" che sono a tutti gli effetti identici al gruppo oggetto di valutazione. Se i due gruppi si differenziano in modo sostanziale, si verifica una distorsione da selezione (*selection bias*). Ciò significa che i risultati osservati al termine dell'intervento di *policy* possono essere dovuti a queste differenze piuttosto che all'intervento attuato.

Questo approccio valutativo è particolarmente indicato per interventi ben circoscritti, applicati in maniera uniforme in tutti i casi e nei confronti di tutti i destinatari/utenti, con catene causali che collegano in modo lineare l'intervento all'effetto. In aggiunta, è necessario assicurarsi che:

- Sia possibile individuare *le variabili-risultato*, rispetto alle quali la presenza di un effetto possa essere verificata con gli strumenti analitici e le risorse a disposizione;
- *L'intervento sia chiaramente identificabile e circoscritto*. Se le componenti della *policy* non possono essere chiaramente identificate come singoli interventi, la portata complessiva della valutazione si riduce. Il trattamento deve essere identificabile nelle sue caratteristiche, nelle conseguenze che ci si aspetta esso produca, ed essere applicabile in modo uniforme a tutti i beneficiari.

Nella sua versione quantitativa, l'approccio controfattuale utilizza due diversi metodi, ovvero i *disegni sperimentali* e i *disegni non sperimentali*.

⁴ European Commission: *Monitoring and evaluation of European cohesion policy. European Social Fund guidance document* (June 2015)

⁵ G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale*, op.cit.

4.1. Valutazione controfattuale: disegni sperimentali

L'idea di fondo del metodo sperimentale nasce dalla sperimentazione clinica utilizzata per testare l'efficacia dei farmaci. Un certo numero di pazienti affetti da una patologia viene suddiviso in due gruppi mediante sorteggio (o randomizzazione). A un primo gruppo (gruppo sperimentale o di "trattamento") viene somministrato il farmaco, all'altro ("gruppo di controllo") una sostanza inerte. Il decorso della patologia che si osserva tra i membri del gruppo di controllo sarà quindi molto simile al decorso che la patologia avrebbe avuto tra i membri del gruppo sperimentale in assenza del farmaco.

Sebbene il disegno sperimentale (disegno sperimentale randomizzato o sperimentazione controllata randomizzata) offra la massima validità interna per valutare un intervento, questo metodo è spesso inutilizzabile. Ad esempio, non è possibile utilizzare un disegno sperimentale randomizzato quando : (i) l'intervento è già stato avviato e i partecipanti sono già stati selezionati con altri metodi; (ii) esistono risorse sufficienti per far partecipare all'intervento tutte le persone ammissibili e non sarebbe etico negare un servizio o un beneficio ai soli fini valutativi; e (iii) l'intervento riguarda un numero limitato di persone con caratteristiche uniche.⁶ Quando il disegno sperimentale non è realizzabile, la valutazione controfattuale si affida a disegni quasi-sperimentali.

Il problema principale del metodo sperimentale è il mantenimento dell'integrità dell'esperimento, cioè la separazione netta tra gruppo di controllo e gruppo sperimentale per l'intera durata dell'esperimento. Da un lato, vi è il problema della non partecipazione e della partecipazione incompleta (*no-show* e *drop-out*) da parte di coloro che sono stati assegnati al gruppo di trattamento; dall'altro vi è il problema dei *cross-over*, cioè quegli individui assegnati al gruppo di controllo che riescono a usufruire ugualmente dell'intervento. Questi fenomeni sono la principale fonte di distorsione delle stime che rendono impraticabile il semplice confronto tra partecipanti e non-partecipanti.

Il punto di forza di questo metodo di valutazione risiede nell'immediatezza della nozione di effetto come differenza tra osservato e controfattuale. Tuttavia, le difficoltà di applicazione sono tali da limitarne notevolmente la rilevanza pratica, persino nei paesi dove questo metodo ha raggiunto un alto livello di utilizzo (Regno Unito, Canada e Stati Uniti d'America). La randomizzazione riduce il campo di applicabilità di questo approccio a quelle situazioni, relativamente rare, in cui tale metodo non è ostacolato da ragioni etiche, politiche, o da difficoltà nell'ottenere la collaborazione degli operatori dei servizi. L'applicabilità del metodo sperimentale è anche limitata agli interventi a copertura parziale. Tutti gli interventi di natura universale (quali l'istruzione elementare, gli interventi di tutela ambientale o di promozione del lavoro dignitoso), non si prestano per definizione all'individuazione di un gruppo di controllo e quindi alla randomizzazione.

Se la randomizzazione offre il miglior metodo per assicurare la *validità interna* delle stime, il limite maggiore si riscontra sulla *validità esterna*. In particolare, gli effetti stimati su scala ridotta (tipica della sperimentazione) non tengono conto dei vincoli che emergono quando l'intervento viene applicato su larga scala.⁷

⁶ ILO: *Monitoring the performance and evaluating the impact of the Youth Guarantee: A learning package*, op.cit.

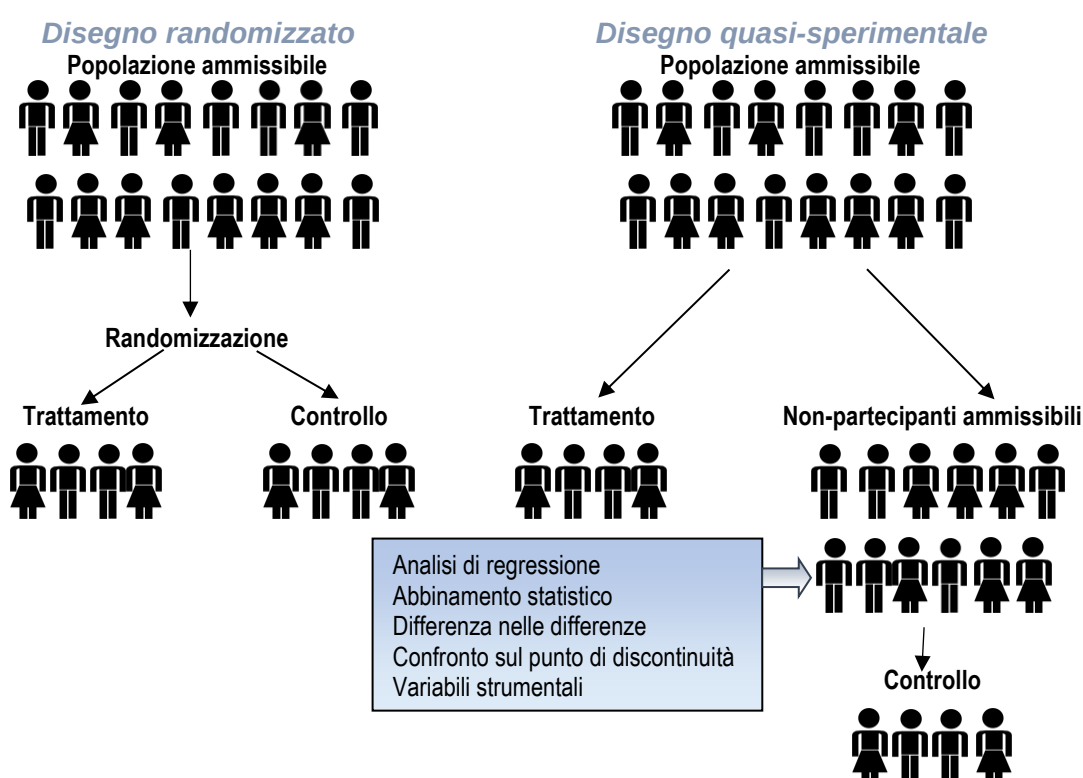
⁷ G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale*, op.cit.

4.2. Valutazione controfattuale: disegni non-sperimentali (o quasi sperimentali)

I metodi non-sperimentali (o quasi-sperimentali) sono i più diffusi, soprattutto in Europa. Essi vengono utilizzati in tutte le situazioni in cui non è possibile assegnare i destinatari di un intervento in modo casuale. In questi casi, il controfattuale può essere ricostruito soltanto facendo delle ipotesi che vanno verificate e tali verifiche dipendono dalla quantità e qualità dei dati a disposizione.

Alla base dei metodi quasi-sperimentali vi sono le diverse tecniche statistiche utilizzate per costruire un gruppo di controllo valido (differenza nelle differenze, l'analisi di regressione, l'abbinamento statistico, il confronto attorno al punto di discontinuità e l'impiego di variabili strumentali).⁸ Il grafico che segue illustra in modo schematico le differenze tra il disegno sperimentale randomizzato e il disegno quasi-sperimentale.

Grafico1. Disegni sperimentali randomizzati e i non-sperimentali



Fonte: ILO: *Monitoring and evaluation of youth employment programmes* (ILO, Geneva, 2003)

La tecnica della **differenza nelle differenze** (*difference-in-differences* o DD) è comunemente utilizzata nelle valutazioni d'impatto come alternativa alla randomizzazione. La tecnica DD si basa sul concetto che la differenza tra i risultati post-intervento dei partecipanti e del gruppo di controllo include sia il "vero" effetto di trattamento sia la distorsione da

⁸ Per un'analisi più puntuale delle varie tecniche quasi-sperimentali si veda G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale*, op.cit. Per un esempio pratico di cose utilizzare le varie tecniche si veda ILO: *Monitoring and evaluation of youth employment programmes: A learning guide* (ILO, Ginevra, 2013).

selezione causata dalle differenze tra i due gruppi. Questa distorsione può essere parzialmente corretta se si riescono ad ottenere dati sulla variabile risultato per entrambi i gruppi in almeno due periodi diversi. Sottraendo la differenza tra i due gruppi misurata in un periodo precedente si elimina la parte di distorsione legata alle differenti caratteristiche tra i due gruppi che non si modificano nel tempo. Se ciò che differenzia destinatari e non destinatari non cambia nel tempo, il metodo della differenza nelle differenze elimina la distorsione da selezione e produce stime corrette dell'effetto dell'intervento. Per esempio, l'applicazione della tecnica DD su un intervento che offre incentivi pubblici alle imprese per aumentare l'occupazione misura il numero medio di occupati in imprese che ricevono lo stanziamento e quello nelle imprese che non li ricevono, un anno dopo lo stanziamento. Questa differenza non rivela l'effetto degli incentivi perché è probabile che il gruppo delle imprese destinatarie si sarebbe comportato in modo diverso dal gruppo delle imprese non destinatarie anche in assenza di incentivi (il comportamento diverso è la distorsione da selezione). Se le differenze tra i due gruppi di imprese osservate in due periodi temporali rimangono costanti, è possibile sottrarre tali differenze dal risultato finale per raggiungere il valore dell'effetto netto.⁹ L'applicabilità di questo metodo richiede che la variabile risultato possa essere misurata ripetutamente in diversi periodi, sia tra i destinatari che tra i non destinatari. Altra condizione necessaria è che la politica da valutare sia identificabile con una variabile trattamento binaria, cioè che ci siano unità trattate e unità non trattate. Se vi sono trattamenti parziali o non uniformi, il metodo non è applicabile.

L'**abbinamento statistico** (*matching*) crea *ex-post* un gruppo di controllo composto da soggetti non trattati che è il più simile possibile nelle caratteristiche osservabili (sesso, età, livello di istruzione, impiego), ai partecipanti (per ogni persona che partecipa all'intervento si cerca un "gemello" non trattato). Una volta selezionato il gruppo di controllo, l'effetto del trattamento è la differenza tra le medie della variabile-risultato nel gruppo dei partecipanti e quelle del gruppo dei non-trattati abbinati.

Le tecniche di abbinamento statistico (e anche quelle di regressione) sono basate sull'assunto di "selezione delle variabili osservabili", in cui si ipotizza che le variabili che influenzano il processo di selezione siano osservabili e quindi possano essere utilizzate per eliminare le differenze di partenza. Quando non è possibile la selezione su variabili osservabili, si possono utilizzare **variabili strumentali** (*instrumental variables*). Il requisito è che esista un fattore esterno che influenza il processo di selezione, condizionando le scelte degli individui coinvolti, e che questo fattore esterno non abbia alcun effetto sulla variabile risultato. La strategia consiste nello "scomporre" la variabile-trattamento in due parti: una parte determinata dal fattore esterno al di fuori della sfera di controllo dell'individuo; e una parte determinata dalle decisioni dell'individuo legate a variabili inosservabili. Quindi si utilizza come trattamento solo quella parte che non dipende dalle variabili inosservabili (in econometria tale fattore esterno viene definito "variabile strumentale").

Il vantaggio principale della tecnica delle variabili strumentali è quello di sfruttare situazioni simili alla randomizzazione e quindi di ottenere stime dell'effetto con un alto grado di affidabilità, mentre la principale difficoltà è quella di trovare o costruire tali variabili, che,

⁹ ILO: *Monitoring the performance and evaluating the impact of the Youth Guarantee: A learning package*, op.cit.

per definizione, devono incidere sulla variabile risultato e devono essere indipendenti dal trattamento.¹⁰

5. La valutazione basata sulla teoria

La *valutazione basata sulla teoria* è utilizzata per l'attribuzione causale degli effetti attraverso la concettualizzazione dell'intervento sulla base di una teoria che individua i meccanismi che hanno permesso di arrivare a un certo risultato.

Nell'ambito di questo approccio, si ipotizza che gli interventi abbiano sempre un effetto (atteso o inatteso, positivo o negativo) e che innescino sempre dei meccanismi nelle categorie di attori coinvolti. Se gli attori sono i destinatari degli interventi si parla di "teoria del programma". Se invece gli attori fanno + parte della catena che va dalla decisione alla realizzazione dell'intervento (decisori, funzionari, manager, operatori), si parla di "teoria dell'implementazione". Le due teorie sono collegate perché le risorse prodotte nel corso dell'implementazione (economiche, sociali, cognitive, culturali, informative, relazionali, ecc.) vanno a integrare le risorse preesistenti del destinatario, influenzandone il comportamento.

La valutazione basata sulla teoria è adatta nel caso in cui si voglia indagare ciò che lega le diverse forme di evidenza empirica e ricostruire la "catena causale" che opera in profondità e che spiega la teoria dietro una correlazione o la variazione di un indicatore. Le domande tipiche della valutazione basata sulla teoria sono "Cos'è esattamente che fa funzionare un dato intervento? L'azione di quali attori è determinante? Quali sono le conseguenze di determinate interazioni e/o assetti organizzativi? Di cosa necessitano gli attori coinvolti nell'implementazione di un determinato intervento, affinché assumano uno specifico comportamento? Come è legato il comportamento dei destinatari alle azioni intraprese dai programmatori?". Nel campo delle scienze sociali, le valutazioni basate sulla teoria utilizzano principalmente l'approccio realista.

5.1. L'approccio realista alla valutazione degli effetti

L'approccio realista si propone di spiegare un esito (*outcome*) attraverso un meccanismo (*mechanism*) e il contesto (*context*). La configurazione "*Esito = meccanismo + contesto*" è l'elemento concettuale centrale nelle teorie nella valutazione realista. Il termine "teoria" ha qui un duplice significato che ricomprende sia teorie scientifiche, basate su ricerche pregresse, i cui risultati si trovano nella letteratura, sia le idee che gli attori (gli attuatori e i beneficiari dell'intervento) si costruiscono sui motivi per cui un intervento funziona. Tutte le teorie, indipendentemente dalla loro provenienza, dovranno poi passare al vaglio della ricerca empirica; dovranno, cioè, essere testate con i metodi, qualitativi e/o quantitativi più appropriati ai contenuti della teoria.¹¹ La valutazione basata sull'approccio realista comprende quattro fasi.¹²

¹⁰ G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale*, op.cit.

¹¹ R. Pawson, N. Tilley, "Introduzione alla valutazione scientifica realista", in N. Stame N. (a cura di), *Classici della valutazione* (Franco Angeli, Milano, 2007).

¹² Cft. A.Blamey, M. Mackenzie "Theories of change and realistic evaluation: Peas in a pod or apples and oranges." in *Evaluation*, N. 13(4), 2007, pp.439–455.

1. Attraverso il dialogo con gli esecutori del programma, il valutatore tenta di capire la natura dell'intervento: Qual è lo scopo del programma? Qual è la natura della popolazione destinataria del programma? In quali contesti e impostazioni opera? Quali sono le teorie prevalenti sul perché il programma funziona per alcune persone in alcune circostanze?
2. Il valutatore traccia una serie di potenziali mini-teorie che mettono in relazione i vari contesti di un intervento con i molteplici meccanismi attraverso i quali potrebbe operare.
3. Nella terza fase, il valutatore intraprende un'indagine sull'esito (*outcome inquiry*) di queste mini-teorie. Ciò comporta lo sviluppo di un modello quantitativo e qualitativo dell'intervento.
4. L'analisi di come si attuano le configurazioni *contesto-meccanismo-esito* all'interno di un programma, permette quindi al valutatore di affinare le teorie su ciò che funziona, per chi e in quali circostanze.

Il riquadro due fornisce un esempio pratico di configurazione contesto-meccanismo-esito utilizzata per la valutazione realista delle politiche di sostegno al reddito.

Riquadro 2: Relazione *contesto-meccanismo-esito* nei sussidi sociali

I sussidi sociali erogati a livello locale agli individui a rischio di povertà hanno successo se riescono a offrir loro gli strumenti per rendersi indipendenti dal sostegno economico pubblico (*outcome*). In alcuni casi, i servizi sociali raggiungono questo obiettivo perché offrono, oltre al contributo, misure integrative che rafforzano un atteggiamento proattivo del beneficiario e ne aumentano gli sforzi per cercare lavoro (*meccanismo*). In altri casi, l'esito positivo può essere dovuto al fatto che il servizio offerto espande le risorse a disposizione del destinatario, che aumenta il controllo sulla propria situazione (*meccanismo*), ad esempio nel caso in cui vengano, in aggiunta al contributo assistenziale, offerti servizi di cura dell'infanzia a madri single (*contesto*). Le madri single, controllando meglio la propria situazione, possono perseguire strategie di medio-lungo termine (*meccanismo*), come ad esempio partecipare ad un corso di formazione, che nel lungo periodo permettono loro di uscire dalla necessità di ricorrere a sussidi, cosa che risulterebbe difficile se la cura dei figli ricadesse interamente e senza alcuna risorsa aggiuntiva su di loro (*contesto*).

Nei casi in cui la misura di sostegno (il servizio di cura all'infanzia) è erogata in maniera non regolare (*contesto*), invece, le donne che beneficiano dell'intervento non la percepiscono come una risorsa su cui contare stabilmente, e sono disincentivate dall'adottare strategie di medio-lungo periodo (*meccanismo*). Altri casi di insuccesso si hanno quando il mercato del lavoro è in grado di offrire solo posizioni incerte o irregolari (*contesto*), spingendo i destinatari a preferire il reddito basso, ma certo e regolare, del sussidio, che, addirittura, può diventare un'integrazione del reddito principale (quello da lavoro informale) (*meccanismo*).

Contesto	Meccanismo	Esito
Oltre al sussidio, i servizi sociali offrono misure integrative	Queste misure inducono negli individui un atteggiamento proattivo, rafforzandone gli sforzi per la ricerca di un lavoro	Successo
Alle madri <i>single</i> , oltre al sostegno economico vengono offerti anche servizi di cura dell'infanzia	Aumentando il controllo sulla loro situazione, è possibile per le madri <i>single</i> impostare strategie di lungo periodo per l'uscita dall'assistenza economica	Successo
La misura è erogata in maniera tale che i beneficiari non la percepiscono come un diritto certo	I beneficiari sono disincentivati dall'implementare strategie di medio periodo per risolvere la loro situazione preferendo strategie di breve periodo	Fallimento
Le offerte di lavoro offrono un reddito incerto e irregolare	I beneficiari preferiscono il reddito basso, ma certo e regolare, dell'assistenza	Fallimento

Fonte: F. Biolcati Rinaldi, *Povert , teoria e tempo. La valutazione delle politiche di sostegno al reddito* (Franco Angeli, Milano, 2006)

L'approccio realista può essere utilizzato quando la valutazione d'impatto controfattuale non è possibile (attuazione diversa in varie aree geografiche, mancanza di omogeneità nell'attuazione) o quando non si conoscono bene i fattori che possono influenzare il risultato.

Nel contesto di cambiamento sociale tipico delle politiche pubbliche, le azioni sono differenziate, fortemente legate ai contesti in cui si realizzano e complesse (numerosi attori, relazioni causali non-lineari, interazioni tra gli elementi dell'intervento poco chiare). Diventa quindi difficile isolare i fattori che influenzano l'esito e utilizzarli come unità statistiche omogenee – non solo per la differenza dei contesti e dei meccanismi, ma anche per la diversità degli esiti, che variano di comunità a comunità.¹³ Il sistema dei valori ed il livello socio-economico hanno un'influenza importante nel definire sia le caratteristiche specifiche dell'intervento (e quindi gli obiettivi specifici e operativi), sia il modo di reagire dei beneficiari. In questi casi la valutazione realista può essere utile, poiché permette di chiedersi non solo se l'intervento funziona e quanto, ma cosa funziona dell'intervento, e in particolare cosa funziona meglio dove, per chi e in quali circostanze. Un possibile schema per condurre una valutazione realista è proposto nel Riquadro 3.¹⁴

Riquadro 3: Schema riassuntivo di una valutazione realista degli effetti di un intervento

Definire lo scopo della valutazione	Identificazione della domanda valutativa	- Qual è la natura e il contenuto dell'intervento? - Quali sono le circostanze e il contesto? - Quali sono gli obiettivi di politica? - Quale natura e forma hanno i risultati? - Effettuare ricerche esplorative per informare la discussione con i decisori
	Articolazione dello scopo della valutazione	- Integrità della teoria - l'intervento funziona come previsto? - Scelta della teoria - quali teorie sull'intervento sembrano adattarsi meglio? - Confronto - come funziona l'intervento in ambienti diversi e per gruppi diversi? - Test di verifica - come si traduce l'intento di <i>policy</i> nella pratica dell'intervento?
	Ricerca e definizione delle teorie dell'intervento	- Ricerca di teorie esplicative nella letteratura - Raggruppare, categorizzare e sintetizzare le teorie che emergono dalla letteratura - Definire un quadro di valutazione da popolarsi con le evidenze empiriche

¹³ G. Marchesi, L. Tagle, B. Befani: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale*, op.cit

¹⁴ R. Pawson et al.: *Realist synthesis: An introduction* (ESRC Research Methods Programme, University of Manchester, Paper 2/2004).

Riquadro 3: Schema riassuntivo di una valutazione realista degli effetti di un intervento (cont.)

Ricerca e valutazione delle evidenze	Ricerca delle evidenze	▪ Definire la strategia di campionamento ▪ Definire le fonti di ricerca, i termini e i metodi da utilizzare
	Valutazione delle evidenze	▪ Rilevanza del test - la ricerca affronta la teoria in esame? ▪ Rigore dei test - la ricerca sostiene le conclusioni tratte dai valutatori?
Estrazione e sintesi dei risultati	Estrazione dei risultati	▪ Sviluppare un sistema per la raccolta dati ▪ Raccogliere i dati per popolare il quadro di valutazione
	Sintesi dell'analisi dei risultati	▪ Confrontare i risultati con quelli della letteratura (sia in positivo che in negativo) ▪ Raffinare le teorie relative all'intervento alla luce delle evidenze empiriche raccolte
Conclusioni e raccomandazioni		▪ Coinvolgere i decisori nell'analisi dei risultati ▪ Elaborare e verificare le conclusioni e le raccomandazioni con gli attori principali ▪ Condivisione dei risultati, conclusioni e raccomandazioni

Fonte: Adattamento basato su R. Pawson et al.: *Realist synthesis: An introduction* (ESRC Research Methods Programme, University of Manchester, Paper 2/2004).

Riferimenti bibliografici

- Biolcati Rinaldi F., *Povert , teoria e tempo. La valutazione delle politiche di sostegno al reddito* (Franco Angeli, Milano, 2006)
- Blamey A., Mackenzie M. “Theories of change and realistic evaluation: Peas in a pod or apples and oranges.” in *Evaluation*, N. 13(4), 2007, pp.439–455.
- European Commission: *Monitoring and evaluation of European cohesion policy. European Social Fund guidance document* (June 2015)
- Fujiwara D. 2010. *The Department for Work and Pensions Social Cost-Benefit Analysis Framework* (Department for Work and Pensions working paper No. 86)
- Gertel, P.J. Martinez, S., et al. 2011. *Impact evaluation in practice* (Washington, DC, World Bank).
- Marchesi G., Tagle L., Befani B.: *Approcci alla valutazione degli effetti delle politiche di sviluppo regionale* (Ministero dello sviluppo economico, Unit  di valutazione degli investimenti pubblici, 2011).
- Mason P., Barnes M.: “Constructing Theories of Change”, in *Evaluation* Vol 13(2), 2007 pp. 151 – 170
- New Economics Foundation (NEF). 2013. *Social CBA and SROI* (Economics in policymaking, Briefing note 4)
- Organizzazione internazionale del lavoro (OIL): *Monitoring and evaluation of youth employment programmes: A learning guide* (ILO, Ginevra, 2013).
- *Monitoring the performance and evaluating the impact of the Youth Guarantee: A learning package* (ILO, Ginevra, 2017)
- *Social cost-benefit analysis of the Youth Guarantee: A learning package* (ILO, Geneva, 2017)
- Pawson R. et al.: *Realist synthesis: An introduction* (ESRC Research Methods Programme, University of Manchester, Paper 2/2004).
- Pawson R., Tilley N., “Introduzione alla valutazione scientifica realista”, in N. Stame N. (a cura di), *Classici della valutazione* (Franco Angeli, Milano, 2007).
- Rogers, P. (2007). Theory-based evaluations: Reflections ten years on. *New Directions for Evaluation*, 114, 63–67.
- Rossi P.H., Lipsey M.W., Freeman H.E.: *Evaluation. A systematic approach* (7th Edition, Sage publication, 2020)
- Stame, N.: “Theory-based evaluation and types of complexity” in *Evaluation*, 10(1), 2004, pp. 58–76